HYBRID
CONFERENCE

South African Genetics Society &
South African Society for Bioinformatics
JOINT CONFERENCE

BIO2022

Bioscience, Big Data & The 4th Industrial Revolution

24-27 APRIL 2022

STIAS Conference Centre
Marais Rd, Mostertsdrift, Stellenbosch

ABSTRACT BOOK

CONTENTS

WELCOME MESSAGE

Dear Colleagues and Fellow Researchers

Welcome to BIO2022!

It is our pleasure to welcome you to the joint conference of the South African Genetics Society and the South African Society for Bioinformatics: BIO2022; at Stellenbosch, in the heart of the Cape Winelands! This conference will also be the first hybrid conference to be hosted by these Societies, and we extend a special welcome to those joining via the online medium.

We are currently experiencing an exciting, revolutionary time in the Biosciences, with genetics and bioinformatics permeating every element of modern biology with far reaching implications for the broader society. And none so evident, as during the current global COVID-19 pandemic where genome sequencing has played a critical role in new variant detection, understanding viral spreading patterns and virus evolution. The development of sequencing- and other omic technologies, means that we are generating genetic sequence- and other biological data in quantities and speeds that are unprecedented in the history of scientific exploration!

And whilst we are gripped by the "Big Data" phenomenon, in the mists of the fourth industrial revolution (4IR), we are struck with awe, amazed by what we discover, and bewildered by it all…

Not only do we now have the potential capacity to answer research questions that have eluded science for centuries, but we are also increasingly struggling to effectively analyse and interpret this "Big Data". For this reason, the South African Genetics and Bioinformatics Societies thought it opportune to bring together researchers from across the spectrum under a central theme: *"Bioscience, Big Data & the 4th Industrial Revolution"* to facilitate dialog, networking, brainstorming and cross pollination of ideas. In this we hope that we will make research more impactful, tackling the many challenges and opportunities of our time, ranging from climate change and environmental degradation to sustainable economic growth, food security, and disease prevention. Where better to start this conversation, than between the vineyards and on the banks of the Eerste River, whilst enjoying a class of the Cape's finest produce in the picturesque historic town of Stellenbosch!

We are excited by the scientific programme that speaks to the innovative and encompassing nature of research being conducted in South Africa and abroad, and we thank all our presenters for their efforts and forthcoming contributions at the conference.

We look forward to hosting and engaging with you!

Clint Rhode Ph.D., Pr.Sci.Nat.
Conference Co-Chair (SAGS)

Gerard Tromp Ph.D.
Conference Co-Chair (SASBi)

COMMITTEES

Conference organising committee members:

Dr Clint Rhode (Co-Chair)

Prof Gerard Tromp (Co-Chair)

Prof Aletta Bester-van der Merwe

Prof Helena Kuivaniemi

Dr Ruben Cloete

Dr Cedric Werely

Scientific Committee:

Prof Aletta Bester-van der Merwe (Chair)

Dr Clint Rhode (SU)

Prof Irene Barnes (UP)

Dr Alisa Postma (UP)

Dr Fisayo Olotu (RUBi, Rhodes)

Dr Sue-Rica Schneider (UFS)

Ms Robyn Jacob (SASRI)

Dr Hocine Bendou (SANBI)

Prof Oliver Zishiri (UKZN)

Dr Juliana Klein (SU)

Dr Allan Sanyanga (RUBi, Rhodes)

Dr Olivier Amamuddy (RUBi, Rhodes)

Prof Sanuska Naidoo (UP)

Dr Thommas Musyoka (RUBi, Rhodes)

EMERGENCY NUMBERS

National Emergency Numbers

107 (Landline) / 112 (Cell) – Police, Ambulance, Fire

10177 – Ambulance/Metro Control Room

Ambulance /Metro control room

10177 – Ambulance

021 937 0500 – Metro

Fire / Rescue / Emergency

021 808 8888 / 0861 808 911 – Fire Station

021 887 4446 – Winelands Emergency

021 888 5275 – Mountain Fires,Berg Fires, Rural Fires

Animal Welfare

021 886 4901 Office

071 169 9922 Emergency

Vet

021 887 3052 – Stellenbosch Animal Hospital

082 578 1311 – Emergency

Locksmiths

021 887 6706 – Neelsie Sleutels & Tekens

021 886 6785 – N & G Locksmiths

Municipal Helpline

021 808 8111 – Stellenbosch

086 126 5263 – Cape Winelands District

Police

10111 – National

021-8095000/3/15 – Stellenbosch Police Station

021-8885940/47 – Cloetesville

Hospitals

021 808 6100 – Stellenbosch Hospital (Ambulance Service)

021 887 0310 – Stellenbosch Hospital

021 887 0305 – Medicross Stelkor

021 808 3496 / 3494 – Campus Health Service

021 861 2000 – Medi-Clinic (Ambulance Service)

021 809 6500 – Mediclinic (Private hospital)

Mountain Rescue

021 937 0300 – Mountain Club of South Africa

021 887 4446 – Winelands Emergency

021 937 0500 – Metro Emergency

10177 – Metro Control Room

Vehicle Breakdown

021 887 8596 – Tomson Motors

082 807 0901 – After Hours, GP Towing

KEYNOTE SPEAKERS

Prof Soraya Bardien

Soraya Bardien is a professor of human genetics in the Division of Molecular Biology and Human Genetics at Stellenbosch University, where she heads the Parkinson's disease Research Group. She did all her undergraduate and postgraduate studies at the University of Cape Town and obtained a PhD degree on the genetic aetiology of the eye disorder, retinitis pigmentosa. She later undertook postdoctoral training at Stellenbosch University, the University of the Western Cape and the University of Cape Town. Soraya served as the Chair of the Southern African Society for Human Genetics (2017-2019), and she currently serves on several international committees including the African Society of Human Genetics, GP2 Underrepresented Populations Working Group, International Parkinson and Movement Disorder Society's Task Force on Recommendations for Clinical Genetic Testing in Parkinson's Disease and Steering Committee of Genetic Epidemiology of Parkinson's Disease Consortium. She is also on the Editorial Boards of Frontiers in Neurology, Journal of Movement Disorders and Parkinson's disease and Related Disorders. Soraya has won awards for general performance and outstanding research outputs from Stellenbosch University, was the second runner up in 2018 for the Department of Science and Technology South African Women in Science Award (SAWiSA), and a finalist for the National Science and Technology Forum (NSTF)-South32 Awards in the Lifetime Award category in 2021. To date, she has co-authored 89 publications, 34 as senior (final) author, and three book chapters. When she's not working, Soraya likes to spend time working in her garden or hiking in the beautiful mountains in the Western Cape.

GENETICS OF PARKINSON'S DISEASE: THE SOUTH AFRICAN PERSPECTIVE

Parkinson's disease (PD) is a progressive, neurodegenerative disorder which occurs due to a loss of dopamine-producing neurons in a part of the brain known as the substantia nigra pars compacta. This loss of dopamine causes people with PD to suffer from a wide range of problems with movement as well as neuropsychiatric problems such as psychosis and depression. PD is the fastest growing neurological disorder worldwide in terms of prevalence, disability, and deaths. Addressing this impending health care challenge requires action aimed at preventing the disease, improving worldwide access to care and existing treatments (eg, levodopa), increasing funding for research (e.g. to understand the underlying causes), and development of new improved therapies.

Our research group focussed on PD, is the only one of its kind in South Africa, and we investigate two of the above-mentioned goals – i.e. studying the underlying causes and development of new therapeutic options. We currently have three main foci: (i) mutation screening, (ii) functional studies and (iii) therapeutic studies on curcumin. For the mutation screening work, we have recruited a unique collection of individuals with PD from diverse ancestries around South Africa for genetic studies. We use various mutation screening techniques starting from candidate gene screening and more recently moving to next-generation sequencing approaches (targeted gene panels, whole-exome sequencing). To date, these studies have revealed very few known mutations leading us to speculate that our populations may harbour novel PD-causing mutations. The talk will elaborate briefly on these approaches and our main findings. Moreover, we are currently embarking on a new collaboration with the Global Parkinson's Genetics Program (GP2) which will allow us to perform whole-genome sequencing on our study participants.

For the functional studies section, we use ex vivo (dermal fibroblasts from individuals with mutations) and in vitro (neuroblastoma cell line) approaches to study the effect of pathogenic variants on the cell. Our work to date has implicated mitochondrial dysfunction as a biological process involved in PD. Lastly, our work on curcumin has shown that it is a strong antioxidant and that it may be able to alleviate mitochondrial dysfunction in cellular models of PD.

In summary, this talk will provide a background to our studies on PD, an overview of our main research goals, as well as the approaches used and the main findings. The significance of our work is that understanding the genetic causes and disease mechanisms underlying PD in local populations is necessary for future development of precision medicine strategies for this debilitating disorder.

KEYNOTE SPEAKERS

Dr Simo Maduna

Simo Njabulo Maduna was born and educated in South Africa, although he lives and works in Norway. He earned all of his university degrees at Stellenbosch University in South Africa, from Bachelor (Molecular Biology & Biotechnology, 2011) to Doctorate (Genetics, 2017). He has trained in a broad range of subfields within the disciplines of Molecular Ecology and Aquaculture over the course of his academic career. His training has enabled him to participate in a variety of international collaborative projects spanning a range of research disciplines, including aquaculture genomics, genome assembly, landscape/seascape genomics, invasion genomics, molecular phylogenetics, reproductive biology, and wildlife forensics. Whenever possible, he combines genetic, phenotypic, and ecological techniques. His research has societal implications in terms of the biodiversity crisis, conservation of natural resources, primary production, and sustainable development. He is currently a Researcher in Molecular Ecology at the Norwegian Institute of Bioeconomy Research's (NIBIO) Department of Ecosystems in the Barents Region in Arctic Norway. He also serves as a Research Professor for Reel Science Coalition (RSC), a South African non-profit organization that serves as a link between academia, government, and the public in order to increase awareness and support for marine and coastal biodiversity conservation and management in southern Africa.

GENETIC MONITORING FOR BIODIVERSITY CONSERVATION AND ENVIRONMENTAL SUSTAINABILITY—EXAMPLES FROM THE NORWEGIAN BARENTS REGION

Continued anthropogenic environmental change is wreaking havoc on natural populations, with the stresses and pulses of induced ecological processes affecting a species' local habitat, resulting in inadvertent distribution shifts, hybridization events, and eventual biodiversity loss. It is more critical than ever to monitor the unintended consequences of human activity on not only natural populations, but also community structures and ecosystems. DNA-based (genetic and genomic) monitoring is a critical component of biodiversity monitoring because it allows for the tracking and quantification of temporal changes in population genetic metrics or other population data. Genetic/genomic monitoring enables the estimation of a variety of biological parameters, including demographic parameters (abundance, occupancy, hybridization, and disease status), population genetic parameters (genetic diversity, structure, and effective population size), and responses to anthropogenic selective pressures (exploitation, biological invasions, and climate change). This keynote address will highlight the practical implications of integrating genetic data into management, conservation objectives, and policymaking, as well as capacity building through international partnerships, using case studies from the Norwegian Barents Region.

KEYNOTE SPEAKERS

Prof Sanjib Panda

Prof. Sanjib Kumar Panda, Head, Department of Biochemistry, Central University of Rajasthan is among the top 2% of Scientists in the World from India in the field of Plant Biology in the list published by Stanford University, USA for whole careers in the year 2020 & 2021. He has done his M.Sc. (Gold Medalist), M.Phil. Ph.D. & D.Sc. from Utkal University, Bhubaneswar, Odisha, India and has a long 26 years of illustrious teaching & research career of working in various central Universities which includes eleven years as Professor and worked in various administrative positions as Dean, Head, Chairman, President, Governing body etc.. He has research trainings and Visiting Professorships in notable Institutions in University of California, University of North Carolina, Oklahoma State University, USA, University of Bonn, University of Dusseldorf, Germany, Okayama University, Gifu University, Shizuoka University, RIKEN, Japan, Stellenbosch University, South Africa and Russian Academy of Sciences etc.

He has published more than 160 research papers in International & National journals in his areas of expertise including Molecular Biology, Plant Functional Genomics, Genetic Engineering etc. with cumulative impact factors above 300, citations of more than 8040 and H index of 48. He is a fellow of the Royal Society of Biologists, London. He has supervised more than 21 Ph.D. students, and 150 + Master students. He has been the recipient of 19 research projects from DBT, DST, CSIR, UGC and 5 International projects with Japan, South Africa and Russia etc.

FUNCTIONAL GENOMICS FOR ABIOTIC STRESS TOLERANCE IN CROPS

Agricultural productivity and food security has remained a challenge in the era of global climate change, the Covid pandemic, and the shrinking of agricultural land because of extreme urbanization and an ever-increasing population. In this situation, adoption of technology and high-throughput platforms have become the choice to prepare climate smart crops. Functional genomics is one of the high potential post-genomic science technologies that help in addressing such biological questions that are posed to crops because of environmental stress. Using transcriptomic, proteomic and metabolomics platforms it is possible to identify novel genes, transcription factors, proteins and even metabolomics markers linked to abiotic stress. Our group attempts to look into the functional genomics perspectives in addressing realistic agricultural issues in terms of abiotic stress in major crops.

KEYNOTE SPEAKERS

Peter van Heusden

Peter van Heusden is a senior bioinformatician working at the South African National Bioinformatics Institute (SANBI) in Cape Town, South Africa. He has over 20 years experience in bioinformatics and the development of computing infrastructure to support data analysis in low resource settings. A specialist in the use of pathogen genomic data for disease surveillance and advisor to numerous public health and academic laboratories in Africa, he is passionate about improving access to diagnostics, especially for the benefit of those historically denied access to diagnostics and medical care.

REALISING THE POTENTIAL FOR PATHOGEN GENOMICS IN ADVANCING PUBLIC HEALTH IN AFRICA

The COVID-19 pandemic has ushered pathogen genomics and genomic epidemiology from the sidelines of bioinformatics to topics of mainstream attention (if not always understanding). The growth of pathogen genomics has happened alongside pressure to digitise case reporting and build data systems between the clinical, epidemiological and laboratory spaces that offer both potential for public health as well as challenges related to ethics, sustainability and equity. Addressing these challenges will require commitment from governments, donors and practitioners of public health and bioinformatics if their promise for improving surveillance and diagnostics are to translate into better public health outcomes for people on the African continent.

KEYNOTE SPEAKERS

Prof Chandra Verma

Chandra Verma carried out his undergraduate studies in the Indian Institute of Technology, Kanpur (India) followed by a D. Phil in York (UK). Subsequently he worked at the structural biology labs in York after which he assumed a group leader position in the Bioinformatics Institute (A*STAR, Singapore), leading efforts in applying physics based atomistic modelling to a diverse range of projects in collaboration with experimentalists, clinicians and pharma. Their work has resulted in various patents and two spin-off companies pursuing oncology, opthalmology and design.

PROGRAMME

SUNDAY 24 APRIL

16:00 - 18:00	**Welcome Reception** (STIAS)
	Registration, canapés, red & white wine, soft drinks

DAY 1 | MONDAY 25 APRIL

07:30 - 08:45	**Registration** (STIAS foyer)
08:45	**Welcome & Presidential Address SAGS/SASBi (Clint Rhode, Gerard Tromp, Sanushka Naidoo)**
09:15 - 10:30	**JOINT PLENARY SESSION** (Auditorium 1 & 2) | Chair: **Dr Clint Rhode**
09:15 - 10:00	Invited Plenary Speaker: **Prof Chandra Verma** | Bioinformatics Institute, The Agency for Science, Technology and Research, Singapore **Translating biomolecular modelling to industry and the clinic**
10:00 -10:30	Invited Keynote Speaker (Biomedical Sciences): **Prof Soraya Bardien** | Division of Molecular Biology & Human Genetics, Stellenbosch University **Parkinson's disease: a genetic conundrum**
10:30 - 11:00	TEA
11:00 - 12:45	**JOINT SESSION: Biomedical Sciences** (Auditorium 1 & 2) | Chair: **Prof Helena Kuivaniemi**
11:00 - 11:15	**Moela Pontsho** | Retinoblastoma Binding Protein 6 (RBBP6) expression elicits the sensitization of cervical cancer cells to cisplatin treatment
11:15 - 11:30	**Bolu Oladunjoye** | Structure-based virtual screening of selected malaria box compounds against falstatin, a multi-staged protein in *Plasmodium falciparum*
11:30 - 11:45	**Adebowale Emmanuel Aladejana** | In vitro evaluation of the anti-diabetic potential of Helichrysum petiolare Hilliard & B.L. Burtt using HepG2 and L6 cell lines
11:45 - 12:00	**Kathryn Step** | Parkinson's disease, plink, and putative disease susceptibility variants
12:00 - 12:15	**Fourie Joubert** | Germline sequence variants contributing to cancer susceptibility in South African breast cancer patients of African ancestry
12:15 - 12:30	**Graeme Ford** | Pharmacogenetics of CYP2a6, CYP2b6 and ugt2b7: relevance to HIV treatment in African populations
12:30 - 12:45	**Natasha Kitchin** | The gut microbiota in foetal alcohol spectrum disorders
12:45 - 14:00	LUNCH
14:00 - 15:45	**JOINT SESSION: Bio-Economy, Industries & Technologies** (Auditorium 1 & 2) | Chair: **Willem Botes**
14:00- 14:15	**Gold Sponsor presentation - DIPLOMICS**
14:15 - 14:30	**Oleg Reva** | Genomic and epigenetic comparison of *Bacillus strains* suitable for plant protection
14:30 - 14:45	**Dave Berger** | RNAi "vaccines" for plants: grey leaf spot disease control in maize
14:45 - 15:00	**Demissew Teshome** | Transcriptional re-programming during recovery from short term drought stress in *Eucalyptus grandis*
15:00 - 15:15	**Anneri Lotter** | Phased genome assembly and haplogenome comparison in an F1 hybrid of *Eucalyptus urophylla* and *E. grandis*
15:15 - 15:30	**Nanette Christie** | Data driven molecular breeding of tropical pines
15:30 - 15:45	**Masethabela Maphatsoe** | Exploring the role of microbial cars as potential biocatalysts, and its potential contribution to South Africa's bioeconomy
15:45 - 16:15	TEA
16:15 - 17:00	**JOINT SESSION: Poster Flash Presentations** (Auditorium 1 & 2) |
17:00 - 17:45	**SASBi AGM**

DAY 2 | TUESDAY 26 APRIL

07:30 - 08:45	**Registration** (STIAS foyer)	
08:45	**Welcome & housekeeping (Clint Rhode, Gerard Tromp)**	
09:00 - 10:20	**JOINT KEYNOTE SESSION: Ecology & Evolution** (Auditorium 1 & 2) | Chair: **Dr Ruben Cloete**	
09:00 - 09:35	Invited Keynote Speaker (Ecology and Evolution): **Peter van Heusden** | South African National Bioinformatics Institute **Realising the potential for pathogen genomics in advancing public health in Africa**	
09:35 - 09:50	**Olga Bochkareva** | Detection of genome rearrangements responsible for bacterial phenotype switching	
09:50 - 10:05	**Michal Slupski |** Phylogenetic analysis of circadian clock genes in *Chrysoporthe* species	
10:05 - 10:20	**Cassandra Bianca Schoeman** | The *Encephalartos natalensis*-cyanobacterial coralloid root partnership for nitrogen acquisition	
10:20 - 11:00	TEA	
11:00 - 12:45	**PARALLEL SESSION: Ecology & Evolution** (Auditorium 1 & 2) | Chair: **Prof Irene Barnes**	**PARALLEL SESSION: Biomedical Sciences** (Manor House) | Chair: **Dr Rencia Van Der Sluis**
11:00 - 11:15	**Sumari Venter** | Multi-omics approaches to studying host-pathogen interactions between the forest pest, *Sirex noctilio* and its biocontrol agent, *Deladenus siricidicola noctilio*	**Bonginkosi Shabangu** | Genomic characterisation of *Acinetobacter baumannii* associated with neonatal sepsis and stillbirths in a South African population
11:15 - 11:30	**Tanya Welgemoed** | Pangenome of African populations of the maize fungal pathogen *Cercospora zeina*	**Souiai Oussema** | Longitudinal and comparative analysis of Tunisian newborns gut microbiota according to delivery mode
11:30 - 11:45	**Khajamohiddin Syed** | Impact of lifestyle on cytochrome p450 monooxygenase evolution is clearly evident in bacteria	**Sikozile Ncembu** | Scalable production of h22(scfv)-eta' targeting cd64 in acute myeloid leukemia (AML)
11:45 - 12:00	**Leandri Klynsmith** | Tissue- and sex-specific odour coding for ecological niche adaptation in the forest pest, *Sirex noctilio*	**Nikita Simone Pillay** | Novel Gene Discovery in a Xhosa Family with Parkinson's Disease
12:00 - 12:15	**Shannon Bennet** | *Fusarium* soil survey from the Kalahari	**Chrystal Steyl** | Host genetic factors contributing to susceptibility to Covid-19
12:15 - 12:30	**Shailesh Joshi** | Evaluation of gene flow from commercial sugarcane hybrids to compatible wild relatives	**Caitlin Uren** | The Northern Cape tuberculosis case-control consortium (NCTBC3)
12:30 - 12:45	**Lindokuhle Emmanuel** | Interaction of *Paraburkholderia* species with different agricultural and acacia legumes in South Africa	**Kili James** | Identification of Genetic Variants in Human-Mouse Orthologous Genes in Patients Diagnosed with Non-syndromic Hearing Impairment
12:45 - 13:45	LUNCH	
13:45 - 15:05	**JOINT SESSION: Functional Biology** (Auditorium 1 & 2) | Chair: **Dr Marlon le Roux**	
13:45 - 14:20	Invited Keynote Speaker: **Prof Sanjib Kumar Panda** | Department of Biochemistry, Central University of Rajasthan **Functional genomics for abiotic stress tolerance in crops**	
14:20 - 14:35	**Julia Candotti** | Towards haplotype and structural variant based genetic dissection of complex traits in *Eucalyptus* hybrids	
14:35 - 14:50	**Medha Sood** | The ultrastructure and role of plastids in carbon partitioning during *xylogenesis* in VND7-inducible Arabidopsis	
14:50 - 15:05	**Robyn Jacob** | Comparing the resistant and susceptible defense response of sugarcane challenged with *Eldana saccharina*	
15:05 - 15:30	TEA	
15:30 - 16:30	**JOINT SESSION: Functional Biology** (Auditorium 1 & 2) | Chair: **Dr Cedric Werely**	
15:30 - 15:45	**Kelvin Hull** | Functional genomic processes of early domestication and artificial selection in black soldier fly colonies	
15:45 - 16:00	**Deborah Narh Mensah**| NRPS-dependent siderophore synthetase gene clusters and characteristics of NRPS siderophore synthetase genes in Armillaria and other species in the *Physalacriaceae*	
16:00 - 16:15	**Tassin Jackson** | Draft genome assembly and annotation of *Argyrosomus japonicus* in South Africa, provides insight into the evolutionary characteristics of this species	
16:15 - 16:30	**Darisia Moonsamy** | Expression profiling of interferon-stimulated genes in peripheral blood mononuclear cells from healthy and SARS-COV-2 infected individuals	
16:30 - 17:00	**JOINT SESSION: Poster Flash Presentations** (Auditorium 1 & 2) |	
17:00 - 17:45	**SAGS AGM**	
19:00	**Wine Pairing Social Event at Lanzerac Wine Estate**	

DAY 3 | WEDNESDAY 27 APRIL

07:30 - 08:45	**Registration** (STIAS foyer)		
08:45	**Welcome & housekeeping (Clint Rhode, Gerard Tromp)**		
09:00 - 10:30	**JOINT KEYNOTE SESSION: Statistical & Computational Methods** (Auditorium 1 & 2) | Chair: **Prof Gerard Tromp**		
09:00 - 09:45	Invited Keynote Speaker: **Prof Fyodor Kondrashov** | Evolutionary Genomics, Institute of Science and Technology Austria (IST Austria)		
09:45 - 10:00	**Ozlem Tastan Bishop** | Role of variations in modern drug research and development – Way forward via dynamic residue networks		
10:00 - 10:15	**Shae Swanepoel |** The in planta expression of Austropuccinia psidii in resistant and susceptible *Eucalyptus grandis*		
10:15 - 10:30	**Hannes Strydom** | Modelling the population dynamics of CRISP-CAS9 gene drive systems in Sirex noctilio		
10:30 - 11:00	TEA		
11:00 - 12:30	**PARALLEL SESSION: Statistical & Computational Methods** (Auditorium 1 & 2) | Chair: **Dr Alisa Postma**	**11:00 - 12:30**	**PARALLEL SESSION: Conservation & Environmental Sustainability** (Manor House) | Chair: **Prof Aletta van der Merwe**
11:00 - 11:15	**Tiego Mohlaba** | Development of comprehensive single-cell RNA sequencing pipeline: pre-processing, quality control and identification of outliers	11:00 - 11:30	Keynote address: **Dr Simo Maduna** | (NIBIO, Norway) **Genetic monitoring for biodiversity conservation and environmental sustainability – Examples from the Norwegian Barents Region**
11:15 - 11:30	**Ruben Cloete** | Molecular docking of zinc database natural compounds to SARS-COV-2 coronavirus proteins to identify novel inhibitors with antiviral activity	11:30 - 11:45	**Emma Hubble** | Population dynamics of two endemic catshark species inferred from molecular and tag-recapture data
11:30 - 11:45	**Matthew Adeleke** | Designing t- cell epitope-based vaccine against Eimeria from microneme protein 2 (MIC2) antigen using immunoinformatics approach	11:45 - 12:00	**Mondi Xaba** | Long read de novo assembly of the *Acacia mearnsii* genome
11:45 - 12:00	**Keaghan Brown** | Automated computational workflow to prioritize potential resistance variants identified in HIV Integrase Subtype C and CRF02_AG°	12:00 - 12:15	**Mia Groeneveld** | Towards resolving genetic differences and population distributions of rhino rays from the Southwest Indian Ocean region
12:00 - 12:15	**Carla Louw** | Wastewater-based epidemiology as a viable minimum data strategy for antimicrobial resistance surveillance in a peri-urban setting	12:15 - 12:30	**Anandi Bierman** | Phylogeography of the invasive Ambrosia beetle, *Euwallacea fornicatus* - a prequel to fine-scale landscape genomics for understanding invasiveness
12:15 - 12:30	**Shailesh Joshi** | Evaluation of gene flow from commercial sugarcane hybrids to compatible wild relatives		
12:30 - 13:45	LUNCH		
13:45 - 15:15	**JOINT SESSION: Functional Biology** (Auditorium 1 & 2) | Chair: **Prof Sanushka Naidoo**		
13:45 - 14:00	**Bulelani Sizani** | Structural and functional characterization of NLR proteins differentially expressed in Cassava plant Infected with SACMV		
14:00 - 14:15	**Lazarus Takawira** | Inferring R2R3-MYB family transcription factor gene targets in *Eucalyptus grandis* using DNA Affinity Purification sequencing and machine learning		
14:15 - 14:30	**Carla Buitendag** | The maize fungal pathogen *Cercospora zeina* ecp2 virulence effector causes necrosis in tobacco		
14:30 - 14:45	**Tendo Stanley Tshilate** | Genome comparison in abalone species to provides insights into growth related genes		
14:45 - 15:00	**Rencia van der Sluis** | Critical analyses of the literature: In isovaleryl acidemia, is isovalerylglycine a product fromed by glycine-n acyltransferase?		
15:00 - 15:15	**Kelda Perumal** | Differences and commonalities induced by pma and 1,25(OH)2d3 in THP-1 cells during monocyte-to-macrophage differentiation		
15:15 - 16:00	TEA		
16:00 - 16:30	**JOINT SESSION: E-posters Presentations** (Auditorium 1 & 2)		
19:00	**Gala Dinner at STIAS**		

ABSTRACTS

46

DESIGNING T- CELL EPITOPE-BASED VACCINE AGAINST EIMERIA FROM MICRONEME PROTEIN 2 (MIC2) ANTIGEN USING IMMUNOINFORMATICS APPROACH

Professor Matthew Adeleke[1], Ms Thabile Madlala[1], Dr Victoria Adeleke[1], Dr Abiodun Fatoba[1], Dr Moses Okpeku[1], Dr Adebayo Adeniyi[2]

[1]University of KwaZulu-Natal, Durban, South Africa, [2]Federal University Oye-Ekiti, Oye-Ekiti, Nigeria

Avian coccidiosis is an infectious parasitic disease globally recognised for incurring significant production loss in poultry industry. It is a consequence of single or multiple Eimeria spp. infection, characterized by malabsorption, enteritis, poor productivity, and compromised animal welfare in birds. Control measures against coccidiosis are dependent on chemoprophylaxis therapy and live anticoccidial vaccines. Outcome from these measures has not been satisfactory due to detection of drug resistance in parasites and the presence of drug residues in food, compromising food security. The industry's economic loss has led to an imperative search for novel strategies (vaccines or drugs) that can induce protection against multiple Eimeria species to control coccidiosis. This study aimed to explore Microneme Protein-2 (MIC2) antigen to predict and develop an epitope-based vaccine against coccidiosis by identifying antigenic T-cell epitopes using immunoinformatics. From the in-silico techniques employed, a total of 7 CD8+ and 12 CD4+ T-cell epitopes were successfully identified and merged along with an adjuvant (Monophosphoryl Lipid A) using EAAK, AAY and GPGPG linkers to produce a vaccine multiepitope construct. Physiochemical parameter assessment of proposed vaccine projected vaccine construct as thermostable (instability index of 25.18), hydrophilic, induce immunity through production of antibodies and cytokines, which may be vital in hindering the surface entry of parasite into host while boosting the immune system host, preventing infection. These findings serve as crucial information for experimental design of cost-efficient novel vaccines against Eimeria and have potential in ensuring food security through drastic reduction of treatments that often results in food contamination.

20

IMPLEMENTING A PIPELINE FOR ANALYSING SINGLE-CELL RNA SEQUENCING DATA

Mr. Kwame Ahiavi[1], Prof. Gerard Tromp[1]
[1]*Stellenbosch University*

A key feature of the scientific method is self-correction through independent verification, and this requires that data collection and analyses be performed in reproductive and robust manners. We are developing a robust pipeline for the analysis of single-cell RNA sequencing (scRNA-seq) data. scRNA-seq is a revolutionary tool which permits the dissection of gene expression at single-cell resolution and is providing fresh insights into the composition of apparently homogeneous cell types and transitions between cell states, thereby deepening our understanding of the cell as a functional unit. The data generated by scRNA-seq is characterised by sparsity, heterogeneity, and high-dimensionality as well as large scale. As a result of biological and technical limitations, scRNA-seq data are “noisier” and more complex than their bulk RNA-seq counterparts. Analysing scRNA-seq data demands new statistical and computational methods. Analytical algorithms employed in scRNA-seq pipelines are prone to producing different results depending on the state at the start of the analysis and the number of iterations of computation, complicating reproducibility. We are developing a highly robust, scalable, and reproducible analysis pipeline for scRNA-seq data, implemented in Nextflow, a workflow management system that complies with current best practices in bioinformatics. The pipeline will document all steps and transformations, record software packages and versions, and also incorporate ontological metadata annotation. Containerisation will ensure that software dependencies are satisfied, and contribute to consistent, robust, and reproducible science.

3

IN VITRO EVALUATION OF THE ANTI-DIABETIC POTENTIAL OF HELICHRYSUM PETIOLARE HILLIARD & B.L. BURTT USING HEPG2 AND L6 CELL LINES

Mr Adebowale Emmanuel Aladejana[2]

[1]Medicinal Plant And Economic Development (mped) Research Centre, University Of Fort Hare, Alice, South Africa, [2]Department of Biochemistry and Microbiology, University of Fort Hare, Alice, South Africa

Helichrysum petiolare Hilliard & B.L. Burtt has been listed in a survey of plants used in traditional medicine for the treatment of type 2 diabetes in the Eastern Cape of South Africa. In this study, the antidiabetic potentials of ethanol, cold aqueous (CAQ) and boiled aqueous (BAQ) extracts of H. petiolare were investigated. The cytotoxic and glucose utilization effects of the extracts were evaluated using L6 myocytes and HepG2 (C3A) hepatocytes. α- amylase, α-glucosidase and lipase inhibition assays were also carried out. The ethanol extract showed significant cytotoxic effects in the treated cells. Both BAQ and CAQ extracts significantly increased glucose uptake in L6 and C3A cell lines. The CAQ extract enhanced glucose uptake more in the L6 myocytes than in the C3A cell-lines hepatocytes. The BAQ extract showed higher levels of inhibition on α–amylase and α-glucosidase than CAQ. The activities were not significantly different from acarbose. However, BAQ showed lower lipase inhibition than acarbose ($p<0.05$). The BAQ and CAQ extracts of H. petiolare may, therefore, contain pharmacologically active and relatively non-toxic hypoglycaemic chemicals, which may be effective substitutes in the treatment of diabetes mellitus.

19

PIPELINE AND TOOLS FOR THE ANALYSIS OF MULTIPLEXED ELISA DATA

Mr Jesse Asimeng[1,2,3,4], Dr Elizna Maasdorp[1,2,3,4], Prof Gerard Tromp[1,2,3,4,5]
[1]SAMRC-SHIP South African Tuberculosis Bioinformatics Initiative (SATBBI) , [2]DST/NRF Centre of Excellence for Biomedical Tuberculosis Research, [3]SAMRC Centre for Molecular and Cellular Biology, [4]Division of Molecular Biology and Human Genetics, Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, [5]Centre for Bioinformatics and Computational Biology, Faculty of Science, ,

A cornerstone of scientific progress is independent data verification. It is, therefore, necessary to develop robust analysis pipelines that can ensure reproducible and verifiable analyses. The pipeline should also record all steps and software that generated the results. The analysis of multiplexed ELISA data (Luminex data) can be challenging due to its complexity and variability. In particular, the data preprocessing stage has many steps and is often ad hoc, leading to inconsistency, non-standard approaches and lack of reproducibility. Furthermore, there is the poor integration of ontological metadata necessary for producing data and results that are findable, accessible, interoperable, and reproducible (FAIR) . We will integrate the investigator, study, assay (ISA) metadata framework for annotating data and analyses to enhance the reproducibility of experimental results. An existing data preprocessing pipeline, the Luminex Pipeline, addresses some of the aforementioned challenges. There remains substantial work to improve its utility and overall generalisability. We are improving the summary statistic reports by using Rmarkdown and implementing unit testing of pipeline components using the R Testthat package. Unit testing will ensure greater robustness of the code, which will be compiled into an R package. Pipeline execution will be automated by implementing it in the Nextflow workflow management system. Finally, we will deploy the pipeline in a Singularity container for execution on any platform including high-performance computing clusters.

4

THE SECRETORY STRUCTURES OF C. ERYTHROPHYLLUM LEAVES AND STEMS: MICROMORPHOLOGY, AND ULTRASTRUCTURE

Ms Sahejna Bantho[1], Prof Yougasphree Naidoo[1]
[1]*University of Kwazulu-Natal, Durban, South Africa*

Combretaceae is a large angiosperm family currently prevalent in southern Africa, due to its extensive use in traditional medicine. Specifically, Combretum erythrophyllum is used to treat and prevent venereal diseases and abdominal pain, whilst the bark is used to treat sores. Medicinal properties are attributed to beneficial secondary metabolites found in certain plant species. These are secreted by specialized secretory structures such as trichomes. Due to limited knowledge of this species, this study aims to investigate the micromorphology and ultrastructure of the secretory structures found on the surface of leaves (emergent, young, and mature) and stems of Combretum erythrophyllum. The micromorphology was evaluated using light and electron microscopy. Two distinct trichome types were identified, characteristic peltate scales and non-glandular trichomes. The head cell count of peltate scales appeared to increase upon leaf maturation and ranged from 8 to 19 cells. In addition, the granulocrine pathway was identified as a possible mode of secretion due to the extensive presence of vesicles, vacuoles, and electron-dense material within the peltate scales. The peltate scale structures are seen to exude a gelatinous-like secretion, of varying amounts on leaf samples from different developmental stages. Research indicates the presence of many phytometabolites, hence, it can be concluded that C. erythrophyllum is indeed a plant worthy to be considered for its medicinal properties. The presence of phenols and steroids indicate possible, anti-microbial and antioxidant properties within the plant. The isolation and extraction of these beneficial compounds open avenues into their use in the pharmaceutical industry.

67

RNAI “VACCINES” FOR PLANTS: GREY LEAF SPOT DISEASE CONTROL IN MAIZE

Prof. Dave K. Berger[1,2], Ms Ingrid Marais[1,2], Ms Carla Buitendag[1,2], Dr. Tuan A. Duong[1,3], Dr Bridget G. Crampton[1,2], Prof. Jacques Theron[3]

[1]Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Hatfield, South Africa, [2]Department of Plant and Soil Sciences, University of Pretoria, Hatfield, South Africa, [3]Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Hatfield, South Africa

RNA interference (RNAi) is a phenomenon found across the tree of life, mainly as a mechanism to protect cells against foreign nucleic acids. Many fungi possess RNAi machinery. Recent studies have shown RNAi can be exploited for control of plant diseases caused by fungi. External application of dsRNA targeting fungal genes triggers RNAi resulting in reduced virulence on host plants. We hypothesized that the foliar pathogen Cercospora zeina, causing grey leaf spot disease of maize, possessed the RNAi machinery and could be controlled in this way. We identified orthologues for Dicer-like 1, Dicer-like 2, and two copies of Argonaute in the genome sequence of C. zeina. Phylogenetic analysis confirmed that these RNAi machinery genes were conserved in the Dothidiomycete class of fungi to which C. zeina belongs. Confocal microscopy showed that C. zeina protoplasts and mycelia took up fluorescently labelled dsRNA that was externally applied. GFP activity in a GFP-transgenic C. zeina line was reduced by 50% after treatment with a gfp-specific dsRNA. A dsRNA construct targeting three C. zeina genes reduced C. zeina cell viability by 40%. This was specific since the gfp-dsRNA had no effect on cell viability. Maize was inoculated with C. zeina and different dsRNA constructs were applied. GLS disease was reduced specifically and significantly by the three gene construct compared to the water or gfp-dsRNA treated inoculated plants. This study showed that C. zeina has a functional RNAi machinery, and it is the first step towards a RNAi-based “green” fungicide for GLS disease of maize.

5

PHYLOGEOGRAPHY OF THE INVASIVE AMBROSIA BEETLE, EUWALLACEA FORNICATUS ~ A PREQUEL TO FINE-SCALE LANDSCAPE GENOMICS FOR UNDERSTANDING INVASIVENESS

Dr Anandi Bierman[1], Prof Francois Roets[1], Prof John Terblanche[1]
[1]*Stellenbosch University, Stellenbosch, South Africa*

The Ambrosia beetle, Euwallacea fornicatus originated in southern Southeast Asia. The species complex of E. fornicatus has become an established pest in Israel, Southern California, and Southern Florida on avocado and other tree species. The damage caused by the beetle lies in the cultivation of its Ambrosia fungi which causes fusarium dieback in combination with mass attack by the beetle. E. fornicatus is now considered a high-risk quarantine pest of international concern. The dominant invasion pathways for E. fornicatus are shipment of solid wood packaging material and wood products. E. fornicatus (Polyphagous Shot Hole Borer; PSHB) was first detected in South Africa in Kwa-Zulu Natal in 2017. Since then, this invasive beetle has spread to every province except Limpopo and attacks, and often kills, a range of tree species. Dispersal patterns and relatedness of invasive species are often correlated, as one influences the other, and understanding dispersal and relatedness of invasive species is key to understanding ecology and evolution and allows for informed strategies on control measures. Variation within a population can drive differences in dispersal initiation and traveling distance, ultimately leading to greater or lesser success by spreading the risk. This study serves as the prequel to a SNP-based landscape genomics project and provides the first evidence of novel haplotypes and multiple introductions of E. fornicatus into South Africa. The work aims to provide insights as to the dispersal of this invasive beetle in the country and the relatedness between sub-populations locally and internationally.

53

DETECTION OF GENOME REARRANGEMENTS RESPONSIBLE FOR BACTERIAL PHENOTYPE SWITCHING

Dr Olga Bochkareva[1], Alexey Zabelkin[2], Natalia Natalia Dranenko[3], Dr Nikita Alexeev[2], Prof Mikhail Gelfand[3], Prof Fyodor Kondrashov[1]

[1]IST Austria, Klosterneuburg, Austria, [2]ITMO University, St Petersburg, Russia, [3]A.A. Kharkevich Institute for Information Transmission Problems, Moscow, Russia

Modern sequencing technology provides exceptional opportunities to investigate whole-genome organization and phenotype-genotype relationships. High plasticity of bacterial genomes is provided by numerous mechanisms including horizontal gene transfer and recombination via flanking repeats. Such genome rearrangements may provide parallel adaptation or intra-population phenotypic diversity. We implement an algorithmic solution for the identification of parallel rearrangements in bacterial populations as a tool PaReBrick. The tool takes a collection of strains represented as a sequence of oriented synteny blocks and a phylogenetic tree as input data. It identifies large-scale genomic variants, tests them for consistency with a tree, and sorts the events by their parallelism score. We demonstrated our approach's efficiency for prediction of new cases of phase variation in human pathogens as well as genomic regions responsible for pathogenicity and adaptation. First, new cases of phase variations via large-scale inversions were predicted in Streptococcus pneumoniae, Streptococcus pyogenes, Burkholderia pseudomallei and targeted for further experimental validation. Second, we showed independent acquisition of the genomic islands containing the type 3 secretion system effectors with highly conserved 5'-intergenic regions by different Shigella lineages. Nevertheless, fewer than half of Shigella genomes kept a complete set of effectors with frequent gene losses and duplications indicating that some proteins may affect the same pathway at different stages, working together to cause disease. Our research demonstrates the power of comparative genomics-based on synteny block composition and provides a framework for understanding new molecular mechanisms in pathogens.

80

AUTOMATED COMPUTATIONAL WORKFLOW TO PRIORITIZE POTENTIAL RESISTANCE VARIANTS IDENTIFIED IN HIV INTEGRASE SUBTYPE C AND CRF02_AG

Mr. Keaghan Brown[1], Dr. Ruben Cloete[2]
[1]South African National Bioinformatics Institute, Bellville, South Africa, [2]South African National Bioinformatics Institute, Bellville, South Africa

Mutations and variations plays a crucial role in the context of structural biology and the efficacy of treatment regimens. Several software tools have been developed such as PyMOL, Schrodinger, GROMACS and FoldX which aid in the understanding of mutation effects on protein structures and drug binding. However, most software tools require extensive knowledge and pre-processing on the users behalf. Although numerous software tools exist in order to predict the mutation effects on protein structure, few attempts have been made to integrate them into a single workflow. The Automated Mutation Introduction and Analysis (AMIA) – working title pipeline aims to automate the introduction of mutations, energy minimize the protein structures and analyse the generated mutant/ variant structures for the potential loss or gain of interactions with neighbouring residues and known Integrase Strand Transfer Inhibitors (INSTIs). Furthermore the stabilizing or destabilizing effects of the mutants on the HIV Integrase (IN) structure will be calculated using FoldX. Based on the results obtained from these analyses, the script will assess the relevance on binding and prioritize which mutations/ variants should be experimentally validated. This service aims to utilize the advantages of stand-alone and web-based applications for both wet-lab and dry-lab researchers by accelerating the process of mutation introduction and predicting their effect on drug binding to ultimately provide effective treatment regimens for patients

62

THE MAIZE FUNGAL PATHOGEN CERCOSPORA ZEINA ECP2 VIRULENCE EFFECTOR CAUSES NECROSIS IN TOBACCO

Ms Carla Buitendag[1,2], Mr Trystan Nadasen[1,2], Prof Dave Berger[1,2]
[1]Department of Plant and Soil Sciences, University Of Pretoria, Pretoria, South Africa, [2]Forestry and Agricultural Biotechnology Institute (FABI), Pretoria, South Africa

The fungal foliar pathogen Cercospora zeina is the causal agent of Grey leaf spot (GLS) disease of maize in South Africa and causes significant crop yield losses for smallholder and commercial farmers. Pathogens secrete effector proteins that contribute to their virulence and overcome plant defences. Investigating the C. zeina mode of infection at a molecular level contributes to the identification of virulence factors that can be targeted to combat GLS and increase resistance in maize. These studies contribute to the development of treatments alternative to chemical fungicides, e.g. RNA interference applications. Previous research demonstrated that Ecp2 effector homologs from fungal pathogens can be recognised by non-host Nicotiana tabacum plants by eliciting a hypersensitive response. In this study the Cz-Ecp2 gene was cloned with an apoplast-targeting signal peptide to mediate extracellular transport in tobacco. Agroinfiltration of tobacco leaves enabled the transient expression of the Cz-Ecp2 gene. The results demonstrated that the Cz-Ecp2 effector causes a strong hypersensitive response in tobacco, visible as necrosis 12 days post infection. The Dothistroma septosporum Ecp2 and Phytophthora infestans INF1 effectors were used as positive controls which caused necrosis in tobacco. The protein sequences of Cz-Ecp2 and Ds-Ecp2 have 58% identity and caused similar hypersensitive response levels in tobacco. The Cz-Ecp2, Ds-Ecp2, and Pi-INF1 effectors caused significant levels of necrosis compared to Agrobacterium and vector control treatments which did not induce hypersensitive response ($p < 0.005$). This study indicates that the C. zeina Ecp2 effector can be regarded as a possible virulence factor of the fungus.

65

TOWARDS HAPLOTYPE AND STRUCTURAL VARIANT BASED GENETIC DISSECTION OF COMPLEX TRAITS IN EUCALYPTUS HYBRIDS

Ms Julia Candotti[1], Ms Anneri Lötter[1], Dr Marja M. Mostert-O'Neill[1], Ms S. Melissa Reynolds[1], Dr Nanette Christie[1], Dr Raphael Ployet[1], Dr Gert van den Berg[2], Dr Sharmane Naidoo[2], Dr Nicoletta Jones[2], Prof Sanushka Naidoo[1], Prof Eshchar Mizrachi[1], Dr Tuan A. Duong[1], Prof Alexander A. Myburg[1]
[1]Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa, [2]Sappi Forests, Research, Planning and Nurseries, Shaw Research Centre, Howick, South Africa

In order to understand the genetic mechanisms underlying complex traits, a more detailed understanding of genomic variation is required. Gene variants (haplotypes) and structural variants (SVs) are relatively unexplored sources of genomic diversity in outbred organisms. In Eucalyptus plantation trees, genome-wide single nucleotide polymorphism (SNP) analysis is a standard approach used for trait dissection. However, there is limited understanding of haplotype and SV diversity within Eucalyptus. Therefore, we aim to characterize haplotype and SV diversity in Eucalyptus grandis and E. urophylla, and to determine how this variation is associated with growth and wood properties in their hybrid progeny. Towards this, we have developed a haplotype mining panel using 11,999 oligonucleotide probe sets targeting coding and/or non-coding regions of 6,293 genes. To identify SVs, we produced approximately 100X short-read (Illumina) coverage of the parental genomes and 200X long-read (Oxford Nanopore) coverage of three F1 progeny. Using a trio-binning approach to separate long-reads originating from the pollen and seed parent of each F1 hybrid, we assembled the haplogenomes inherited from three E. grandis pollen parents and two E. urophylla seed parents (average contig N50 of 18 Mb) and then identified SVs within and between the parental species. Next, we will use the haplotypes and SVs to perform genetic dissection of growth and wood properties in a large, multi-parent, F1 hybrid mapping population. This is an important step in Eucalyptus genetics research and will assist in breeding programmes of this globally important forest genus.

34

DATA DRIVEN MOLECULAR BREEDING OF TROPICAL PINES

Dr Nanette Christie[1], Prof Alexzander A (Zander) Myburg[1]
[1]*University of Pretoria, Pretoria, South Africa*

Plant breeding has become a data science as a result of advanced phenotyping platforms and the application of next-generation DNA sequencing and genotyping technologies. However a lack of genome-wide genotyping capability in pine tree species, partly due to the enormous size of their genomes, has hampered progress with the implementation of genomic technologies in pine breeding programmes. We performed genome and gene targeted SNP discovery towards the development of a genome-wide 50,000 marker, multi-species genotyping array for tropical pines (Pitro50K) using the Axion 384-format SNP array platform (ThermoFisher). This cost-effective resource is serving as a basis for a new trans-disciplinary framework supporting genome-assisted breeding of tropical pines in South Africa. Thousands of pine trees from breeding populations of SA forestry companies, as well as natural populations in Mexico and Central America are being genotyped towards a genome diversity atlas and the development of methods for pine species and hybrid identification. We will perform QTL mapping in clonal F1 Pinus patula x Pinus tecunumanii families, as well as develop and implement genomic selection approaches in P. patula x P. tecunumanii F1 hybrid breeding programmes. Furthermore, we are initiating a landscape genomics study of P. patula in collaboration with York Timbers, integrating information on genotype (G), environment (E) and growth, form and processing phenotypes (P) and will ultimately implement artificial intelligence methodology to integrate the various data types for genomic selection and deployment of superior, climate resilient genotypes supported by new molecular breeding informatics resources.

43

MOLECULAR DOCKING OF ZINC DATABASE NATURAL COMPOUNDS TO SARS-COV-2 CORONAVIRUS PROTEINS TO IDENTIFY NOVEL INHIBITORS WITH ANTIVIRAL ACTIVITY

Dr Ruben Cloete[1], Prof Megan Shaw[1]
[1]*University Of The Western Cape, Cape Town, South Africa*

COVID-19 infections are on the rise and warrants the identification of novel drugs to treat and prevent further loss of life. In this study, we investigated two SARS-CoV-2 coronavirus proteins, RNA-dependent RNA polymerase (nsp12; PDBID: 6m71) which is responsible for viral RNA replication and transcription, and guanine N7-methyl transferase (nsp14; PDBID: 5c8t), which plays a key role in mRNA capping and also possesses 3' to 5' exonuclease proofreading activity. Each protein was exploited in molecular docking studies and interaction analyses. Firstly, two known viral RNA polymerase inhibitors (Remdesivir and Favipiravir) were docked to nsp12 and one known inhibitor (Sinefungin and substrate S-adenosylmethionine) to nsp14 as comparative controls. Secondly, a subset of the Specs natural compounds consisting of 1339 compounds from the ZINC database were docked to each protein target to identify diverse compounds with stronger affinity for each protein, higher number of interactions and possessing cell permeable properties as compared to the known compounds. Molecular docking identified 677 compounds with higher docking scores compared to the known antiviral drugs. Interaction analysis of the top 20 compounds indicated a total of seven compounds with a higher number of interactions for the two target proteins compared to the known antivirals and having Total Polar Surface Area scores of less than 140. This suggest that the seven compounds have a stronger affinity for the protein targets and are cell permeable. Future work will involve purchasing these seven compounds and testing their activity against the novel SARS-COV2 coronavirus in a whole cell viral assay.

38

CHARACTERISATION AND EXPRESSION OF ORF2 IN GENOME SEGMENT 10 (SEG-10) OF AHSV

Miss Carmen Coertzen[1], Prof Vida van Staden[1], Prof Christiaan Potgieter[2], Prof Eshchar Mizrachi[1]
[1]*University of Pretoria, Pretoria, South Africa,* [2]*Deltamune, Die Hoewes, South Africa*

Viruses are often under strong selective pressure to enhance the coding capacity of their genomes, allowing expression of multiple proteins from a single mRNA. A second additional open reading frame, ORF2, was identified in African horse sickness virus (AHSV) Seg-10 which encodes the non-structural protein NS3. The ORF2 conservation, size, nature, expression, function and potential subcellular localisation of the putative protein product AHSV is not known and is investigated in this project. A bioinformatic analysis of all available AHSV Seg-10 sequence indicated that ORF2 was maintained in over 400 AHSV Seg-10 sequences, and six size variants were identified ranging from 183 to 252 nucleotides. ORF2 showed high conservation in the central region. AHSV Seg-10 ORF2 shows strong positive selection. The protein encoded by AHSV ORF2, or an ORF2-eGFP fusion protein, was recombinantly expressed in insect and mammalian cells. Localisation of ORF2-eGFP was specific within the nucleus and cytoplasm of insect cells. ORF2 transiently expressed in BSR-T7 cells initially localised in the cytoplasm and later moved to the nucleus, and ultimately caused cells to shrivel showing a cytotoxic effect. Antiserum raised against ORF2 was used to investigate the presence and subcellular localisation of the protein during AHSV infection. Using this antibody, ORF2 could be detected in AHSV infected cells, potentially identifying a novel AHSV non-structural protein. This is the first detection of a putative protein product of Seg-10 ORF2 during normal AHSV infection, and paves the way for further functional analysis of this protein in the AHSV replication cycle.

69

DISCERNING THE GLOBAL PHYLOGEOGRAPHIC DISTRIBUTION OF PHYLLOSTICTA CITRICARPA BY MEANS OF WHOLE GENOME SEQUENCING

Dr Beatrix Coetzee[1], Dr Elma Carstens[2], Dr Megan Dewdney[3], Dr Paul Fourie[2,4], Dr Jeffrey Rollins[5], Prof Aletta Bester-van der Merwe[1]

[1]Department of Genetics, Stellenbosch University, Stellenbosch, South Africa, [2]Citrus Research International, Nelspruit, South Africa, [3]Citrus Research and Education Centre, University of Florida, Lake Alfred, United States of America, [4]Department of Plant Pathology, Stellenbosch University, Stellenbosch, South Africa, [5]Department of Plant Pathology, University of Florida, Gainsville, United States of America

Phyllosticta citricarpa is a fungal pathogen and causative agent of citrus black spot. As a regulated pest in some countries, the presence of the pathogen limits fruit exports, and is therefore of agricultural and economic importance. This is the first study to use high throughput sequencing (HTS) to infer the global phylogeographic distribution of this pathogen. We assembled whole genomes of 71 isolates from eight countries, and used pairwise read mapping to detect and count inter-isolate variants. The HTS data were mined for SSRs and in silico genotypes were generated for all isolates with 1,987 SSR markers. We also identified 32, 560 SNPs relative to the reference genome. The pairwise variant counts, SSR genotypes and SNP datasets presented the same phylogeographic patterns. The Chinese population is the most diverse, and is genetically the furthest removed from all other populations, and therefore it is assumed to be the origin of the pathogen. There is a clear pathway of dispersal from the origin of the pathogen in Asia, to Australia and from there to Southern Africa and the South American countries. Southern Africa is also most likely the source of P. citricarpa that are now found in North America. This study represents the largest whole genome sequencing survey of P. citricarpa to date, and provide a more comprehensive picture of the population genetic diversity and connectivity of P. citricarpa from different geographic origins. This information provides a better understanding of the epidemiology of the citrus black spot pathogen and its dissemination pathways.

13

ELUCIDATING THE GENOMIC CAUSES OF SKELETAL DYSPLASIAS IN SOUTH AFRICA

Miss Kimberly Coetzer[1], Professor Shahida Moosa[1,2]
[1]*Stellenbosch University, Cape Town, South Africa,* [2]*Tygerberg Hospital, Cape Town, South Africa*

Skeletal dysplasias (SDs) are a broad group of heritable disorders encompassing over 450 individual diseases. To date, 437 SD genes have been identified, in pathways related to bone and cartilage formation, mineralization and homeostasis. Deep phenotyping of 51 molecularly undiagnosed patients with rare SDs was undertaken. Consented samples were sent for sequencing either as a skeletal gene panel test or whole-exome sequencing (WES). Variant calling was performed using bcbio. Variants were filtered depending on family structure prior to annotation. Known SD genes were extracted and in silico predictive algorithms were used to produce a shortlist of likely disease-causing variants. These variants were then further analysed and interpreted using ACMG/AMP guidelines and clinical presentation. Segregation was confirmed with Sanger sequencing. The overall diagnostic yield was 90.2%. The biggest group of patients (n=39) had osteogenesis imperfecta (OI). 22 OI patients came from three large unrelated families of mixed South African ancestry and shared the same pathogenic variant in COL1A2 (c.1892G>T, p.Gly631Val). Further studies are required to determine if it is a founder mutation. In line with previous studies, COL1A1/2 underlie the majority of OI in our population as 92.3% of patients had a variant in either gene. Two novel and five known variants were found in known SD genes. The five patients with “negative” exome are enrolled in further gene discovery analyses. This project provided undiagnosed SD patients with a confirmed molecular diagnosis, thereby directly impacting their treatment and management while providing insights into the basis of SDs in our understudied population.

55

PERFORMANCE OF ARTIFICIAL INTELLIGENCE ALGORITHMS ON OMICS DATA

Miss Ashley Ehlers[1,2], Professor Gerard Tromp[1,2,3,4,5]

[1]Centre for Bioinformatics and Computational Biology, Stellenbosch University, Stellenbosch, South Africa, [2]Bioinformatics Unit, South African Tuberculosis Bioinformatics Initiative, Stellenbosch University, Cape Town, South Africa, [3]Division of Molecular Biology and Human Genetics, Department of Biomedical Sciences, Stellenbosch University, Cape Town, South Africa, [4]DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research, Stellenbosch University, Cape Town, South Africa, [5]South African Medical Research Council Centre for Tuberculosis Research, Stellenbosch University, Cape Town, South Africa

Mapping and sequencing of the human genome in 1990 soon led to new technologies that could obtain many molecular measurements within a tissue or cell. These technologies have allowed researchers to study the underlying biology at a resolution that has never been possible. The measuring of such biological molecules in a high-throughput way is called “omics”. Furthermore, advanced bioinformatics and computational biology are enabling improved performance of data-based predictions through the application of machine learning algorithms, a form of artificial intelligence which is placed to transform the twenty-first century. ML plays an important role in biomedical research, especially in the analysis of big ‘omics data. Due to the intrinsic high cost of acquiring large South African omics data sets many of the data sets available in the domain are small. The sample sizes (n) pose an analytical problem since the data sets are derived from few independent observations. Thus, these data sets present many parameters (p) leading to a p>>n problem. Such data easily lead to overfitting (biased models) and therefore would require identifying characteristics of data-based prediction performance to deal with the p>>n problem. This study aims to associate omics-based molecular measurements, including data sets with sample sizes that range from tens (small) to thousands (large), with a clinical outcome of interest by identifying the key trends among several different ML algorithms, evaluating their performance metrics, and identifying their usability for human disease prediction.

40

INTERACTION OF PARABURKHOLDERIA SPECIES WITH DIFFERENT AGRICULTURAL AND ACACIA LEGUMES IN SOUTH AFRICA

Mr Lindokuhle Emmanuel[1]

[1]L.E Masimula, , City of Tshwane, South Africa, [2]C.W Beukes, City of Tshwane, South Africa, [3]S.N Venter, City of Tshwane, South Africa, [4]E.T Steenkamp, City of Tshwane, South Africa, [5]E.K Muema, City of Tshwane, South Africa

Rhizobia are Gram-negative soil bacteria, which provide nitrogen to leguminous plants through symbiotic nitrogen fixation, in return receiving protection and carbohydrates from photosynthesis. They belong to two classes, the Alpha and Betaproteobacteria. The majority described rhizobia belong to the Alphaproteobacteria whereas only three rhizobial genera belong to the Betaproteobacteria, including species of Paraburkholderia (formerly Burkholderia). The aim of this study was to investigate the symbiotic interaction between Paraburkholderia strains and a range of agricultural legumes, as well as a native-invasive, and a non-native Acacia species in a controlled environment in South Africa. Eight Paraburkholderia strains (31.1, CI3, WC7.3b, WSM4176 and MM6662R1, originally isolated from legumes belonging to various tribes samples in South Africa) and (JPY169, JPY251 and CCGE1002, originally isolated from mimosoid legumes in Latin America) were investigated. The legumes included were (common bean, cowpea, pigeon pea, siratro, lablab, chickpea, soybean, groundnuts, lentils, lucerne, as well as Vachellia karroo and Acacia mearnsii). All strains nodulated common bean, cowpea, pigeon pea, siratro, lablab, and V. karroo under control environmental conditions. This meant that the considered Paraburkholderia species have the potential for development as inoculants for some agricultural crops, although this would need to be tested under field conditions. The species also cross-nodulated the African invasive V. karroo, a first study to show the Latin American Paraburkholderia species ability to nodulate an African legume. These results were supported by the phylogenetic results, which confirmed the identities of the recovered strains tested.

58

PHARMACOGENETICS OF CYP2A6, CYP2B6 AND UGT2B7: RELEVANCE TO HIV TREATMENT IN AFRICAN POPULATIONS

Mr Graeme Ford[1,2], Prof Fourie Joubert[1,2], Mrs. Antionette Niehaus[1,2], Prof Michael Pepper[2]
[1]Center for Bioinformatics and Computational Biology, Faculty of Natural and Agricultural Sciences, University of Pretoria, Pretoria, South Africa, [2]Institute for Cellular and Molecular Sciences, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa

Objectives: This study focuses on identifying variation in selected CYP genes related to treatment response in patients with HIV by investigating variant characteristics and effects in African cohorts.
Design: Cytochrome P450 (CYP) 2A6, 2B6, and Uridine 5'-diphospho-glucuronosyltransferase (UGT) 2B7 allele frequencies were studied using public-domain datasets obtained from the 1000 Genomes Phase 3 project, the African Genome Variation Project (AGVP) and the South African Human Genome Programme (SAHGP).
Methods: Variant annotation was performed using self-identified ethnicity to conduct allele frequency analysis in a population-stratification sensitive manner. The NCBI DB-SNP database was used to identify documented variants and standard frequencies, and the E! Ensembl Variant Effect Predictor tool was used to perform the prediction of possible deleterious variants.
Results: A total of 4 468 variants were identified across 3 676 individuals following pre-filtering. Eight variants were identified at an allelic frequency (1% or more) which also predicted deleterious consequences. Of these, four were found to occur at a clinically significant frequency (4% or more), with only one corresponding to the CYP2A6*2 haplotype.
Conclusions: This study describes allele frequencies in several African sub-populations and identifies previously documented and potentially novel variants of clinical relevance. Despite the mixed sequence coverage, the variants identified pose notable avenues of future inquiry. This study illustrates the need to perform genetic pre-screening of patients as a standard of practice in HIV care, particularly on the African continent where HIV is highly prevalent.

64

POPULATION DYNAMICS OF TWO ENDEMIC CATSHARK SPECIES INFERRED FROM MOLECULAR AND TAG-RECAPTURE DATA

Miss Emma Hubble[1], Doctor Juliana Klein, Professor Aletta Bester-van der Merwe
[1]*Department of Genetics - Stellenbosch University, Stellenbosch, South Africa*

The catshark family, Scyliorhinidae, forms part of the rich endemic elasmobranch biodiversity of Southern Africa. As they are exposed to many threats, information regarding their movements and genetic diversity is vital to aid in the implementation of applicable conservation strategies. Currently, there is a lack of scientific information on distribution patterns and the extent of population structure for endemic elasmobranchs, such as Poroderma africanum and Poroderma africanum. It is therefore not possible to evaluate how vulnerable these species are to over-exploitation from local fisheries. This study aims to address these knowledge gaps by assessing genetic diversity, population structure and movement patterns of these two co-distributed Poroderma catshark species through integration of molecular and tag-recapture data. The application of genus-specific microsatellite markers for population genetic assessment of these catsharks provided insight into the genetic diversity and population connectivity of Poroderma catsharks from various geographical locations along the South African coastline. The use of a long-term tag-recapture dataset from the Oceanographic Research Institute's Cooperative Fish Tagging Program confirmed high site fidelity in both species, as the majority (72.52%) of recaptures occurred in the same location where they were tagged with a very small number travelling > 50km from the original tagging location. While the two Poroderma catsharks are majorly sympatric along the South African coastline, P. pantherinum was seen to have a wider and more-eastward geographical range than P. africanum. Several unique instances of P. pantherinum individuals tagged and recaptured in Namibia as well as further north than Durban were observed.

18

FUNCTIONAL GENOMIC PROCESSES OF EARLY DOMESTICATION AND ARTIFICIAL SELECTION IN BLACK SOLDIER FLY COLONIES

Mr Kelvin Hull[1], Mr Matthew Greenwood[1], Ms Melissa Lloyd[2], Professor Aletta Bester-van der Merwe[1], Dr Clint Rhode[1]

[1]Stellenbosch University, Stellenbosch, South Africa, [2]Insect Technology Group Holdings UK Ltd, Guildford, United Kingdoms

Insects are among the most diverse and abundant groups of animals, and are increasingly being utilised in emerging industries. The black soldier fly (BSF), Hermetia illucens, shows promise as a biotechnological vector for the bioremediation of organic waste into environmentally friendly nutrient sources for use in aqua- and agriculture practices. However, little is known regarding the functional genetic adaptations that occur over generations of mass-rearing. We assessed transcriptomic differences associated with early stages of domestication and artificial selection in BSF colonies. Differential gene expression was evaluated regarding (i) two selection strategies [no artificial selection (NS); artificial selection for greater larval mass (SEL)], and (ii) generational time within the captive environment (F2 vs F3). RNA-Seq was conducted using 5th instar BSF larvae (n = 36), representing equal proportions of NS (F2 = 9; F3 = 9) and SEL (F2 = 9; F3 = 9) lines. Reads were aligned to a publicly available BSF genome, and the resulting count data was used to identify differentially expressed genes (DEGs). Differential response to artificial selection was evident, with 245 DEGs observed between selection strategies (FDR-corrected P-value < 0.05). Genes related to lipid (CRYL1, FGGY), carbohydrate (ENOSF1, IDH1) and protein (CG7200, NEP4) metabolism were up regulated in SEL larvae, which were significantly (P < 0.001) greater in mass than NS larvae. Temporal differences revealed 376 DEGs, with the F3 generation overexpressing developmental (PNPO) and defence response (CYP6A2, CYP12A2) genes, pointing to an elevated developmental rate as a consequence of generations of mass-rearing.

6

IN-SILICO BASED APPROACH TO DESIGN A T-CELL MULTI-EPITOPE BASED VACCINE FOR LEISHMANIASIS

Miss Tamara Ishwarlall[1], Dr Abiodun Fatoba[1], Miss Leah Maharaj[1], Dr Mrs Victoria Adeleke[1], Dr Moses Okpeku[1], Dr Adebayo Adeniyi[2,3], Dr Matthew Adeleke[1]

[1]University of KwaZulu Natal, Durban, South Africa, [2]University of the Free State, Bloemfontein, South Africa, [3]Federal University, Oye-Ekiti, Nigeria

Leishmaniasis is a vector borne disease associated with lesion formation on the skin, membranes, or organs. It is the second most life-threatening parasite disease following malaria. At present, there is no vaccine available for all the manifestations of leishmaniasis. The study was aimed at developing a multi-epitope vaccine using in-silico tools. Various immunoinformatic techniques were performed to construct a multi-epitope vaccine capable of stimulating cellular immune responses. CD8+ and CD4+ T-cell epitopes were screened from several Leishmania genomes. Antigenicity, allergenicity, binding affinity to major histocompatibility complex I and II, and cytokine inducing tests were performed. Docking was performed with the Toll-Like Receptor 4. Various physicochemical and structural analyses were carried out. A final list of 18 CD8+ and 12 CD4+ overlapping T-cell epitopes were identified. The final epitopes are antigenic, non-allergenic, and capable of binding to major histocompatibility complex I and II, respectively. The CD8+ and CD4+ epitopes were bound to the Human Leukocyte Antigens -A*02:06 and -DRB1*01:01, respectively, with a high binding affinity. The population coverage of the alleles was higher in endemic areas (95.72% - 68.55%) as compared to non-endemic areas (26.93%). Physicochemical analysis revealed that the construct was thermostable, basic, and hydrophobic. The complex contained nine hydrogen bonds and two salt bridges. The anti-leishmanial vaccine produced in this study consists of appropriate properties to induce protection against leishmaniasis. This in-silico design should serve as a rational step and lead to a laboratory experimental design of an effective vaccine against leishmaniasis.

44

IDENTIFICATION OF IMMUNOGENIC T-CELL AND B-CELL EPITOPES FROM MYCOBACTERIAL MEMBRANE PROTEIN LARGE AND RESISTANCE-NODULATION-CELL DIVISION TRANSPORTER PROTEINS OF MYCOBACTERIUM ULCERANS

Miss Tamara Ishwarlall[1], Miss Leah Maharaj[1], Dr Mrs Victoria Adeleke[1], Dr Moses Okpeku[1], Dr Adebayo Adeniyi[2,3], Dr Matthew Adeleke[1]

[1]University of KwaZulu Natal, Durban, South Africa, [2]University of the Free State, Bloemfontein, South Africa, [3]Federal University, Oye-Ekiti, Nigeria

Buruli ulcer is a neglected tropical disease caused by Mycobacterium ulcerans. It is associated with skin lesion formation, that may lead to disability. At present, there are no fully effective vaccines against M. ulcerans infection. This study aimed at constructing two multi-epitope based vaccines using in-silico tools. The bacterial genome was screened to identify CD8+, CD4+ T-cells and B-cell epitopes from mycobacterial membrane protein large and resistance-nodulation-cell division transporter proteins. Antigenicity, allergenicity, toxicity, binding affinity to major histocompatibility complex I and II; cytokine inducing testing, and T-cell epitope docking were performed. The population coverage of the CD8+ and CD4+ T-cell associated human-leukocyte alleles was determined. Two vaccine constructs were created using the LprG and RpfE adjuvants, respectively. The constructs underwent refinement, physicochemical and structural analysis, and molecular dynamics simulation. Following screening, 9 CD8+ and 7 CD4+ conserved T-cell epitopes and 1 B-cell epitope were identified. The T-cell epitopes are capable of binding to their respective major histocompatibility complexes. The CD4+ T-cell epitopes are positive inducers of interferon-gamma and interleukin-4. The population coverage for the T-cell epitopes ranged from 15.33-97.80% for endemic areas. The final epitopes and vaccine constructs are antigenic, non-allergenic, non-toxic and thermostable. Molecular dynamics simulations indicated that once the constructs were bound to the respective adjuvant, they entered a more stable state. The vaccines computationally constructed in this study display appropriate properties to elicit a protective immune response against M. ulcerans infection. These constructs will serve as the initial step towards the design of in-silico multi-epitope-based vaccines.

17

DRAFT GENOME ASSEMBLY AND ANNOTATION OF ARGYROSOMUS JAPONICUS IN SOUTH AFRICA, PROVIDES INSIGHT INTO THE EVOLUTIONARY CHARACTERISTICS OF THIS SPECIES

Ms Tassin Jackson[1], Dr Clint Rhode[1]
[1]*Department of Genetics, Stellenbosch University, Stellenbosch, South Africa*

The biology and life history traits of Argyrosomus japonicus (Dusky kob) has been relatively well studied in South Africa, Australia, and more recently in China. However, the life history characteristics and morphology of A. japonicus within these regions differ significantly, which may be an indication of separate genetic sub-populations or even subspecies. In this study, we report the draft genome assembly and annotation of the South African dusky kob and conduct a comparative genomic analysis with the Chinese Dusky kob genome sequence and other available fish genomes to provide a better understanding of the evolutionary ecology and history of the species. A draft genome of 708,288,114 bp (N50 = 59,954; BUSCO = completeness of 97.8%) was assembled from the combined sequencing data of two individuals (male and female). This data was generated utilising a dual sequencing strategy, producing both long and short reads to attain both optimal coverage and depth. Gene prediction using BRAKER identified a total of 50,858 protein-coding genes, of which 20,670 unigenes were assigned to 113 subcategories of Gene Ontology (GO), with biological process making up the majority (84%). The comparative genomic analyses of A. japonicus and five closely related species revealed an average similarity of 78%, indicating high levels of conservation and low levels of divergence among the species. However, despite the slow evolutionary rate of A. japonicus, analyses detected the presence of positive selection. Selection which would allow for region specific adaptions, thus providing insight into the dynamics involved in the evolution of the species.

23

COMPARING THE RESISTANT AND SUSCEPTIBLE DEFENSE RESPONSE OF SUGARCANE CHALLENGED WITH ELDANA SACCHARINA

Robyn Jacob[1], Dr Malcolm G Keeping[1], Dr R Stuart Rutherford[1]
[1]*SASRI, Mount Edgecombe, South Africa*

Eldana saccharina Walker (Lepidoptera: Pyralidae) has plagued the sugarcane industry for over 50 years, with losses caused by the pest estimated at around R1 billion per annum. Commercial sugarcane cultivars vary in their resistance to E. saccharina, which is determined in variety selection trials and subsequently monitored in post-release variety evaluation trials. In this study, RNA sequencing (RNA-seq) was used to detect early and late defense mechanisms induced by E. saccharina challenge in two sugarcane cultivars, N11 (susceptible) and N33 (resistant). A total of 2697 differentially expressed genes (DEGs) were identified in the resistant response (94.4% upregulated and 5.6% downregulated), with 1442 DEGs in the susceptible response (81.3% upregulated and 18.7% downregulated) (log2|ratio|>2; FDR corrected P value ≤ 0.01). These figures highlight the increased transcriptional response required for successful defense. Both early responses were characterised by an enrichment of DEGs involved in 'hormone biosynthetic process' indicating the role of changes in hormone levels in driving defense. The early resistant response was enriched in upregulated DEGs annotated in defense related biological processes ('response to wounding' and 'cyanogenic glycoside biosynthetic process'), whilst these processes were delayed and more enriched amongst the DEGs identified in the late susceptible response. The expression profiles of a few E. saccharina induced genes were chosen for validation using qRT-PCR confirming upregulation as a result of the challenge. The dataset provides a rich source of candidate genes for further study and currently a promising dirigent domain-containing gene induced in the resistant response only, is under investigation.

76

IDENTIFICATION OF GENETIC VARIANTS IN HUMAN-MOUSE ORTHOLOGOUS GENES IN PATIENTS DIAGNOSED WITH NON-SYNDROMIC HEARING IMPAIRMENT

Ms Kili James[1], Dr Oluwafemi Oluwole[1], Professor Ambroise Wonkam[1]
[1]*University Of Cape Town, Cape Town, South Africa*

Despite, hearing impairment (HI) being a global cause of disability, there is little that is known about the causative mechanisms of the disease particularly in African patients. Recently, 38 human-mouse orthologous HI genes were investigated in the African population. A homozygous pathogenic variant was found in the MCPH1 gene in a single Cameroonian non-syndromic HI (NSHI) patient. The variants found in previous research (MCPH1 c.2311C>G and SPNS2 c.867C>A) were genotyped using Sanger sequencing and high-resolution melt in a larger and well-characterised cohort (n = 90) of sporadic and familial NSHI cases from Cameroon and South Africa HI in order to investigate the roles of these targeted variants. These variants were not found in this study. However, two missense pathogenic variants in the MCPH1 gene: c.2222G>A p.(Arg741Gln) rs779231385 and novel c.2234A>C p.(His745Pro) were found in two different Cameroonian patients and were absent in ethnically matched unaffected controls. In conclusion, the identification of another homozygous pathogenic variant in the MCPHI gene in a Cameroonian patient supports that the gene may contain mutation hotspots and therefore play a causative role in NSHI in Cameroonians. Further analysis using cell-based assays should be considered to understand the role of this gene in African patients diagnosed with NSHI.

36

A SYSTEMATIC INTEGRATION OF EMPIRICAL AND COMPUTATIONAL STUDIES TO BIOPHYSICALLY DESCRIBE KLEBSIELLA PNEUMONIAE NICOTINATE-NUCLEOTIDE ADENYLYLTRANSFERASE

Ms Olamide Jeje[1], Dr Ikechukwu Achilonu[1]
[1]*University of the Witwatersrand, Johannesburg, South Africa*

The continuous threat of drug-resistant Klebsiella pneumoniae justifies identifying novel targets and developing effective antibacterial agents. A potential target is nicotinate nucleotide adenylyltransferase (NNAT), an indispensable enzyme in the biosynthesis of pyridine dinucleotides. Given the vital role of NAD+ in controlling key cellular processes required for bacteria survival, NNAT represents an attractive target for the design of novel broad-spectrum antibiotics. NNAT catalyses the adenylation of nicotinamide/nicotinate mononucleotide (NMN/NaMN), using ATP to form nicotinamide/nicotinate adenine dinucleotide (NAD+/NaAD). Its activity is widely conserved, and the respective enzymes exhibit various structural and catalytic properties that reflect their species and specificity. To successful design inhibitors with therapeutic potential, there is a need to understand the biophysical structure of the enzyme. In this study, we used computational modelling (homology modelling, molecular docking, and molecular dynamic simulation) to biophysically described NNAT from K. pneumonia (KpNNAT). The results obtained showed that KpNNAT exist mainly as a monomer with a predominately α-helical secondary structure content and a binding site that is partially hydrophobic. Its substrates ATP, NMN and NAD+, all share the same binding pocket with similar affinity and exhibit an energetically favourable binding. Overall, ATP binding affects KpNNAT dynamics, and the dynamics of ATP binding depend on the presence of Mg2+. The information obtained from this study would serve as a basis for further evaluation towards designing structure-based inhibitors with therapeutic potential.

52

EVALUATION OF GENE FLOW FROM COMMERCIAL SUGARCANE HYBRIDS TO COMPATIBLE WILD RELATIVES

Dr Shailesh Joshi[1,2], Ms Nokulunga Maima[1,2], Dr` Dyfed Lloyd Evans[1,2], Dr Meenu Ghai[2]
[1]South African Sugarcane Research Institute, Durban, South Africa, [2]School of Life Sciences, College of Agriculture, Engineering and Science, University of KwaZulu-Natal, Durban, South Africa

South African sugarcane industry is in the process of developing a genetically modified sugarcane. The evaluation of potential natural hybridization between a GM crop and related plant species forms a part of pre-commercialization Environmental Risk Assessment. Miscanthidium capense, M. junceum and Narenga porphyrocoma are phylogenetically closely related to sugarcane, occur in the sugarcane cultivation area. To ascertain the likelihood of gene flow between sugarcane and its antecedent species, replicated field trials were established in the Pongola region of KZN; in addition to testing them in an artificial photoperiod environment at Durban. Agronomic evaluation for various growth characters were recorded at monthly intervals for both plant and first ratoon crops. Flowers were collected fortnightly from the field trial, whereas, from photoperiod treatments, plants were inspected daily and were then used for crossing. Phylogenetic analyses amongst the species were conducted based on whole chloroplast and ITS barcode sequences to determine the phylogenetic distance between them. Miscanthidium species were found to be highly fertile both in natural and artificial conditions, whereas differential fertility was observed within sugarcane varieties. Non-synchronization between Miscanthidium and sugarcane was observed, resulting in no natural hybridization. Human mediated crosses at photoperiod facilities did not result in seed set. This is likley due to genetic distance and chromosomal incompatibility between these species, which is also confirmed via phylogenetic analyses. Further studies need to include the specific sugarcane genotype/s used for transformation in GM studies, to check their flowering pattern and fertility, as this varies with both genotype and environmental conditions.

57

GERMLINE SEQUENCE VARIANTS CONTRIBUTING TO CANCER SUSCEPTIBILITY IN SOUTH AFRICAN BREAST CANCER PATIENTS OF AFRICAN ANCESTRY

Mr Dewald Eygelaar[1], Prof Elizabeth J van Rensburg[1], **Prof Fourie Joubert[1]**
[1]*University Of Pretoria, Pretoria, South Africa*

Since the discovery of the breast cancer susceptibility genes, BRCA1 and BRCA2, various other genes conferring an increased risk for breast cancer have been identified. Studies to evaluate sequence variants in cancer predisposition genes among women of African ancestry are limited and mostly focused on BRCA1 and BRCA2. To characterize germline sequence variants in cancer susceptibility genes, we analysed a cohort of 165 South African women of self-identified African ancestry diagnosed with breast cancer, who were unselected for family history of cancer. With the exception of four cases, all others were previously investigated for BRCA1 and BRCA2 deleterious variants, and were negative for pathogenic variants. We utilized the Illumina TruSight cancer panel for targeted sequencing of 94 cancer susceptibility genes. A total of 3.6% of patients carried a pathogenic/likely pathogenic variant in a known breast cancer susceptibility gene: 1.2% in BRCA1, 0.6% in each of BRCA2, ATM, CHEK2 and PALB, none of whom had any family history of breast cancer. The mean age of patients who carried deleterious variant in BRCA1/BRCA2 was 39 years and 8 months compared to 47 years and 3 months among women who carried a deleterious variant in other breast cancer susceptibility genes.

21

THE GUT MICROBIOTA IN FOETAL ALCOHOL SPECTRUM DISORDERS

Ms Natasha Kitchin[1], Dr Jacqueline S. Womersley[1], Mrs Andrea Engelbrecht[1], Mrs Anna-Susan S. Marais[1], Mrs Marlene M. de Vries[1], Prof Philip A. May[2], Prof Soraya Seedat[1], Prof Sian M. J. Hemmings[1]

[1]Department of Psychiatry, Stellenbosch University, Cape Town, South Africa, [2]Department of Nutrition, Gillings School of Global Public Health, Nutrition Research Institute, University of North Carolina, Chapel Hill, USA

The prevalence of Foetal Alcohol Spectrum Disorders (FASD) in the Western Cape is 31%, significantly higher than the global prevalence of 0.77%. Neurodevelopment is dependent on microbial composition and the corresponding microbial metabolic outputs, therefore alcohol-induced microbial alterations may alter infant gut microbiota functioning, increasing the risk of FASD development. This study therefore aimed to compare the gut microbial composition of infants diagnosed with and without FASD. 16S rRNA sequencing was performed on microbial DNA extracted from stool samples collected from 211 infants. The dada2 pipeline, PhyloSeq and vegan were used to process the data, calculate diversity measures and compute the statistical analyses of microbial composition. The infant gut microbiota was dominated by Prevotella, Bacteroides, Faecalibacterium, Bifidobacterium and Eshcerichia/Shigella. Bifidobacteria, which degrade the sugars in breast milk and have beneficial probiotic effects, was found to be higher ($p = 0.017$) in infants diagnosed with FASD. A lower abundance of Bifidobacteria has been observed in children with Autism Spectrum Disorder (ASD), making this finding unexpected. Prevotella was higher ($p = 0.003$) in infants diagnosed with FASD, a finding that mirrors findings in children diagnosed with ASD in other low- and middle-income countries. Prevotella readily breaks down mucin - a structural component of mucus which protects the colon. Increased abundance of Prevotella may compromise the intestinal barrier, allowing bacteria and their metabolic outputs to enter the bloodstream and influence neurodevelopment. These findings are promising for microbe-based therapeutic interventions to reduce the extent of neurocognitive deficits and the debilitating symptoms associated with FASD.

41

TISSUE- AND SEX-SPECIFIC ODOUR CODING FOR ECOLOGICAL NICHE ADAPTATION IN THE FOREST PEST, SIREX NOCTILIO

Mrs Leandri Klynsmith[1], Mrs Alisa Postma[1], Dr Tuan Duong[1], Dr Jeremy Allison[2], Prof Bernard Slippers[1]
[1]Department of Biochemistry, Genetics and Microbiology (BGM), Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria,, South Africa, [2]Natural Resources Canada, Canadian Forest Service, Great Lakes Forestry Centre, Sault Ste. Marie,, Canada

The evolutionary and ecological success of insects is partly ascribed to their sophisticated and uniquely adapted chemosensory systems. Insect chemosensation genes form some of the largest multigene families and are important in reverse genetics approaches to study the relationship between ecology, evolution and genomics of these organisms. The woodwasp, Sirex noctilio, poses a significant threat to pine plantations in various parts of the world. Genomics, transcriptomics, phylogenetics and differential expression analyses were used to investigate the chemosensory gene families of this economically important pest. Here, 129 chemosensation genes, including 45 ORs, 12 GRs, 48 IRs, 1 SNMPS, 14 OBPs and 9 CSPs, were identified in the genome of S. noctilio and phylogenetically characterised. Most of these genes exhibited basal phylogenetic clustering within the order Hymenoptera. RNA-sequencing of the olfactory and non-olfactory tissues of adult and pupa S. noctilio woodwasps allowed for expression profiling of the chemosensory genes identified. Tissue- and sex-specific expression patterns were observed in these data. Receptors potentially involved in recognition of ecologically relevant chemical stimuli, such as pheromones and host-tree volatiles, were identified. The phylogenetically basal position of this woodwasp within the Hymenoptera allows for a deeper understanding of the evolution of the multigene families associated with chemosensation in this ecologically and economically important order. This study broadens our understanding of the molecular processes involved in perireceptive events of S. noctilio chemosensation. Tissue- and sex-specific genes are promising targets for future deorphanisation studies which could aid the identification of novel attractants or repellents for biorational pest management.

83

IDENTIFICATION AND STRUCTURAL ANALYSIS OF CASSAVA MOSAIC DISEASE RESISTANCE (R) GENES AND ASSOCIATED QUANTITATIVE TRAIT LOCI

Mx Keelan Krinsky[1], Dr Bulelani Sizani[1], Prof Chrisse Rey[1]
[1]*University Of The Witswatersrand, Johannesburg, South Africa*

• The cassava crop constitutes a vital component of food security in sub-Saharan Africa. Cassava Mosaic Disease, which is widespread throughout this region, drastically decreases crop yield. Breeding for resistance remains the only viable control strategy, however, long cropping cycles impede progress. Marker-assisted breeding can fast-track this process allowing for rapid development of improved germplasm. Here, we aimed to analyse several molecular markers common in cassava breeding and structurally characterize one ligand-binding LRR protein putatively involved in cassava's immune response.

• 24 germplasm including susceptible, tolerant, and resistant phenotypes, were screened with six SSR/SCAR markers associated with three identified cassava resistance loci (CMD1,2 and 3). Marker bands were sequenced with both in vitro and in silico methods and analysed for the presence of SNPs. Genomic DNA of the LRR protein was sequenced in 14 germplasm and, using homology modeling in Modeller 10.1, the protein structure of native and mutate protein was predicted.

• No clear connection between marker presence and phenotypic response was apparent, with high levels of marker polymorphism, and multiple genomic matches. A protein polymorphism exclusive to susceptible germplasm, leading to alterations in predicted protein structure of the LRR protein, was identified.

• These results indicate that marker presence/absence alone may not guarantee the presence/absence of resistance loci, as markers may contain primer binding region SNPs preventing amplification, or may appear in gene orthologs/paralogues with no connection to resistance loci. Additionally, SNP/SAP mutations within our LRR protein may compromise ligand binding resulting in attenuated effector recognition.

22

SEQPREDICTNN: PREDICTING A SEQUENCE FROM A PROTEIN CONFORMATION USING A DEEP FEED-FORWARD NEURAL NETWORK.

Mr Francois Adriaan Lategan[1], Prof Hugh -G. Patterton[1]
[1]*Stellenbosch University, Cape Town, South Africa*

High-throughput pipelines in protein X-ray crystallography and a deeper understanding of the protein chemistry-structure relationship have placed us on the cusp of designing novel proteins with properties not found in nature. Such "designer" proteins may have significant medical and commercial value. Recent advances in the accuracy of protein structure prediction suggests that machine learning may be a valuable tool for protein design. Here we report a proof-of-concept study to establish the feasibility of predicting the sequence of a protein based solely on its spatial conformation. As a first step we considered a simple feed-forward neural network (SeqPredictNN) to predict residues based on features of the surrounding backbone conformation in a protein. The model accurately predicted 26.4% of the residues in a test dataset with an equal frequency of occurrence of each residue, and 27.6% of residues in a test set with a natural distribution of residues. The trained neural network predicted glycine residues with high precision and recall but performed poorly on some residues. The model often predicted residues with similar physico-chemical properties as the true residue. A model trained on imbalanced data – with a non-uniform distribution of amino acids – fails to identify uncommon amino acid residues. Advanced network architectures may produce more accurate amino acid predictions, but better metrics are required to evaluate model performance. The ability of the simplest neural networks to learn the relationship between a protein structure and its sequence gives a promising outlook for the future use of machine learning in protein design.

15

IDENTIFYING GENETIC VARIATION IN CANDIDATE GENES FOR HEAT TOLERANCE AND ASSOCIATION WITH GROWTH PERFORMANCE IN DUSKY KOB (ARGYROSOMUS JAPONICUS)

Miss Stephanie Jane Linder[1], Miss Tassin Kim Jackson[1], Dr Clint Rhode[1]
[1]*Stellenbosch University, Stellenbosch, South Africa*

Dusky kob (Argyrosomus japonicus), is a high value commodity in South Africa due to both its palatability and size. However, this species is particularly vulnerable due to overfishing, as it has been targeted for decades by commercial, recreational and subsistence fisheries, resulting in the steady decline of the natural populations. As such, a shift towards aquaculture has been initiated to meet the current market demands. However, conventional breeding methods are suboptimal for dusky kob, and the application of marker assisted selection is required to accelerate the rate of genetic improvement of commercially important traits such as growth performance in this species. Although, knowledge pertaining to the molecular mechanisms underpinning the expression of this complex trait is lacking. Notably, numerous studies have found associations between growth and heat stress. This is of particular interest in fish, given the important role temperature plays in their life cycle, and the steady rise in global temperatures. Therefore, this pilot study aimed to investigate whether candidate genes for heat tolerance are associated with the growth performance of dusky kob, using a case-control analysis. Two novel non-synonymous single nucleotide polymorphisms within the heat shock protein 90 beta gene, were identified to be significantly associated with growth in dusky kob. Upon validation in future studies, these markers can be utilised in marker assisted selection, coupled with a breeding programme, to allow for the effective and sustainable utilisation of this species.

54

PHASED GENOME ASSEMBLY AND HAPLOGENOME COMPARISON IN AN F1 HYBRID OF EUCALYPTUS UROPHYLLA AND E. GRANDIS

Miss Anneri Lotter[1], Miss Julia Candotti[1], Dr Tuan A. Duong[1], Dr Eshchar Mizrachi[1], Prof Jill L. Wegrzyn[2], Prof AA Myburg[1]

[1]Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University Of Pretoria, Pretoria, South Africa, [2]Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, USA

Molecular breeding strategies that incorporate structural and haplotype variations can provide greater accuracy for selection and improvement of complex traits than biallelic DNA markers. Identification of such variants has been enabled by de novo long-read genome assemblies that can be phased via long-reads spanning haplotypes and connecting SVs. The aims of this study were to assemble, annotate and compare the parental haplogenomes of an F1 hybrid of Eucalyptus urophylla, a tropical species and E. grandis, a temperate species. Using a trio-binning approach, 67 Gb of Nanopore sequencing data from the F1 hybrid was binned into two parental haplotype read groups and assembled independently. The resulting haplogenome assemblies were 544.5 Mb and 566.7 Mb for the E. urophylla and E. grandis haplogenomes, respectively (>98% BUSCO completeness). By using high-density linkage maps, more than 88% of the assembled contigs were anchored to 11 chromosomes. A genome-wide comparison between the two haplogenomes revealed that 257 Mb of the chromosomal assemblies was syntenic and there were 48,729 SVs which ranged from 100 bp to 4.01 Mb in size. Besides the SVs, a total of 8 million SNPs were identified underlying a high heterozygosity estimate of 3.46%. A total of 37,942 and 39,849 genes were annotated for the E. urophylla and E. grandis haplogenome assemblies, respectively (>94.6% BUSCO completeness). This study is the first genome-wide view of genetic differentiation between haplogenomes of an F1 hybrid of E. urophylla and E. grandis that may contribute to hybrid performance.

42

WASTEWATER-BASED EPIDEMIOLOGY AS A VIABLE MINIMUM DATA STRATEGY FOR ANTIMICROBIAL RESISTANCE SURVEILLANCE IN A PERI-URBAN SETTING

Dr Carla Louw[1], Dr Elanna Bester[1], Dr Edward Archer[1], Dr Francois Kamper[1], Prof Gideon Wolfaardt[1,2], Prof Tobias Louw[1]

[1]Stellenbosch University, Stellenbosch, South Africa, [2]Ryerson University, , Canada

The increase in antimicrobial resistance (AMR) seen on a global scale has been exacerbated by the inappropriate use of antibiotics in the healthcare and agricultural industries. AMR surveillance has been implemented world-wide to address the urgency of the problem. Urban-wastewater profiling and wastewater-based epidemiology are viable methods of monitoring population health. The detection of antimicrobial resistant genes, antibiotic residues and microbes in wastewater samples suggests the potential use thereof as a surveillance mechanism for AMR infections. Urban-wastewater profiling for public health surveillance has proven to be effective, however, monitoring wastewater treatment works alone in peri-urban settings may not be sufficient. Informal settlements with run-off not connected to municipal infrastructure remains ubiquitous across the world. In this study Stellenbosch, South Africa is used as a case study representing a region consisting of combined agricultural, industrial, urban, suburban, and informal zones, as is commonly found in low-to-middle-income countries. We developed a viable, generalised minimum data strategy for AMR surveillance in a peri-urban setting using machine learning. We use generalized linear modelling with recursive feature elimination to identify the top predictive features for AMR surveillance from an array of measurable features and the optimal sampling sites from different types of sampling locations. AMR incidence of some microbes on the WHO's critical priority watchlist, such as Acinetobacter baumannii and Staphylococcus aureus, can be predicted with 5-fold cross validated R2 scores of 0.9 or higher. This study demonstrates the combined use of wastewater-based epidemiology and machine learning as a powerful tool for AMR surveillance.

32

T-CELLS AND B-CELLS EPITOPE-BASED VACCINE DESIGNED AGAINST TOXOPLASMA GONDII INFECTION IN CHICKENS USING IMMUNOINFORMATIC APPROACH

Miss Thabile Madlala[1], Dr Matthew Adeleke[1], Dr Moses Okpeku[1], Dr Selaelo Ivy Tshilwane[2]
[1]University Of Kwazulu Natal, Durban, South Africa, [2]University of Pretoria, Onderstepoort, South Africa

Poultry meat serves as the main easily accessible protein source for most families worldwide, making detection of infectious diseases (toxoplasmosis) in chickens a serious concern. Toxoplasmosis is a zoonotic disease caused by Toxoplasma (T.) gondii, which not only contribute to substantial reproductive and economic losses, but also pose critical public health concern since transmission of this parasite from animals to humans is easily facilitated by consumption of poorly cooked meat containing parasite cysts or contact with contaminated water. Currently, available treatments for this disease often results in parasite resistance, drug residues in food and are inefficient in eliminating parasite tissue cysts. This poses a great challenge to poultry industry globally due to high cost of treatment, prevention, and control. Development of novel vaccines against T. gondii through implementation of immunoinformatics has become imperative in the advancement of vaccine design. Hence this study aimed to develop a potentially cost-effective multi-epitope based vaccine against Toxoplasma gondii by exploring potential antigenic epitope candidates identified through immunoinformatic techniques. In this study, the in-silico approach successfully identified 2 CD8+, 7 CD4+ T-cell and 1 B-cell epitopes from Apical membrane antigen 1 (AMA1). These vaccine candidates were used to construct a vaccine that could effectively induce an immune response and prevent host reinfection. Identification and design of epitope-based vaccines against T. gondii through the in-silico approach shows potential in drastically reducing vaccine production costs and ensuring food security. This strategy aims to prevent potential food/meat contamination by chemical residues often used in currently available vaccines.

74

EXPLORING THE ROLE OF MICROBIAL CARS AS POTENTIAL BIOCATALYSTS, AND ITS POTENTIAL CONTRIBUTION TO SOUTH AFRICA'S BIOECONOMY

Miss Masethabela Maphatsoe[1], Prof. Karl Rumbold[1]
[1]*University Of The Witwatersrand, Johannesburg, South Africa*

With the ever-increasing demand to produce "natural" aromatic compounds for industrial applications, microbes and enzymes are amongst the most sustainable and advanced biotechnological methods used in markets. Aldehydes, including vanillin and benzaldehyde, are aromatic compounds widely used as flavor agents and/or precursors for pharmaceuticals. They are prepared using conventional methods (plant material), synthesized using chemicals (reducing agents - LiAlH4) and are generated using biocatalysis (whole cells systems and isolated or purified enzymes). First two methods are implicated with economic and environmental pressure. Approaches offering "greener" solutions and that are cost-efficient such as the use biocatalysts for organic synthesis are encouraged. This study aims characterize and immobilized microbial CAR/s from fungal source to elucidate possible roles as potential biocatalysts. Carboxylic acid reductases (CAR) enzymes reduce of carboxylic acid to their corresponding aldehydes using a single step, and under mild conditions. Herein, we determined biochemical properties of a fungal CAR from Pycnoporus cinnabarinus in vitro. The purified CAR was also immobilized onto a nickel-charged resin (Ni-Sepharose) and the immobilization efficiency was calculated and potential biocatalyst was tested for reusability. Outstandingly, the immobilized CAR was recyclable up to 6 times without loss of activity. Furthermore, during each cycle, the potential biocatalyst maintained a relative activity of ~ 85 % bioconversion of benzoic acid to benzaldehyde. Undeniably, with the rich diversity of flora in South Africa (SA), there is a great opportunity to identify novel microbial enzymes which could be biotechnologically relevant and contribute to SA's bioeconomy.

29

SPECIES-LEVEL PROFILING OF THE MATERNAL VAGINAL BACTERIOME WITH APPLICATION TO FETAL ALCOHOL SPECTRUM DISORDERS

Miss Lauren Martin[1,3], Miss Natasha Kitchin[2,3], Mrs Anna-Susan Marais[2], Mrs Marlene de Vries[2], Prof Philip May[4], Prof Soraya Seedat[2,3], Prof Sian Hemmings[2,3]

[1]Division of Molecular Biology and Human Genetics, Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa, [2]Department of Psychiatry Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa, [3]South African Medical Research Council / Stellenbosch University Genomics of Brain Disorders Research Unit, Cape Town, South Africa, [4]Department of Nutrition, Gillings School of Global Public Health, University of North Carolina, , United States of America

Affecting approximately 16-31% of children in the Western Cape of South Africa, Fetal Alcohol Spectrum Disorders (FASD) describes varying severities of physical and cognitive deficits associated with prenatal alcohol exposure. Exposure to vaginal microbes during delivery results in the acquisition of intestinal bacteria which, via the microbiome-gut-brain axis, play a significant role in neurodevelopment. Alcohol-associated vaginal microbial alterations may thus increase FASD risk in infants. While species-level classifications provide greater insight into the bacterial composition, our previously obtained hypervariable sequencing data lacks the resolution to accurately identify bacterial species.

This study, therefore, aims to improve the species-level resolution of the vaginal bacteriome of 28 women who bore infants with definitive FASD diagnoses through long- and short-read sequencing of the full-length 16S rRNA amplicon using the PacBio Sequel IIe and the Illumina iSeq100, respectively.

Following library preparation and sequencing, we intend to assemble 150bp libraries to form full-length 16S rRNA amplicons. For all data types, taxonomy will be assigned using the SILVA reference database and microbiome-related bioinformatic, composition and diversity analyses will be performed using R packages, dada2, vegan and PhyloSeq. Thereafter, species-level classifications obtained from V1-V2 as well as long- and short-read sequencing data of the entire gene will be compared, and the extent to which the full-length amplicon improves species-level identification will be determined.

Species-level profiling of the vaginal bacteriome will provide greater insight into the underlying bacterial composition and dynamics that may contribute to FASD development. Additionally, the feasibility of a novel sequencing technique will be evaluated.

11

RETINOBLASTOMA BINDING PROTEIN 6 (RBBP6) EXPRESSION ELICITS THE SENSITIZATION OF CERVICAL CANCER CELLS TO CISPLATIN TREATMENT

Dr Pontsho Moela[1], Ms Harshini Mehta[1], Dr Melvin Ambele[1]
[1]*University Of Pretoria, Pretoria, South Africa*

RBBP6 is considered a potential cancer biomarker due to its association with cell proliferation and the fact that it is overexpressed at cervical cancer sites where there is marked apoptosis and elevated p53. More information is emerging regarding the role of RBBP6 in cancer treatment, specifically, its potential to sensitize cancer cells to radiation and certain chemotherapeutic agents via BCL-2 gene regulation. Cisplatin is an FDA-approved chemotherapeutic agent that still presents with acquired resistance in certain cervical cancer cases through p53 repression and BCL-2 upregulation. The study aims to investigate the relationship between cisplatin and RBBP6 expression in cervical cancer cells. RBBP6 was silenced in cancer and normal cells using the RNAi technology, followed by measurement of p53 and BCL-2 at mRNA level using qPCR. Cells co-treated with cisplatin and siRBBP6 were analyzed for apoptosis induction and real-time growth monitoring using flow cytometry and the xCELLigence system, respectively. Cancer cells in the co-treatment group showed a reduction in apoptosis compared to cisplatin-only group and real-time growth monitoring revealed a reduced growth rate in RBBP6-knockdown cells treated with cisplatin. Although wild-type p53 remained unchanged in the co-treatment group of cancer cells, BCL-2 was completely repressed. Findings from this study suggest that RBBP6 expression promotes sensitivity of HeLa cells to cisplatin through BCL-2 downregulation. Knockdown of RBBP6 limits apoptosis induction and delays cell growth inhibition in response to cisplatin. These data have the potential to help improve cisplatin efficacy through personalized administration based on the expression profile of RBBP6 among individual patients.

25

DEVELOPMENT OF COMPREHENSIVE SINGLE-CELL RNA SEQUENCING PIPELINE: PRE-PROCESSING, QUALITY CONTROL AND IDENTIFICATION OF OUTLIERS

Miss Tiego Mohlaba[1,2,3,6], Prof. Gerard Tromp[1,2,3,4,5,6]

[1]Stellenbosch University, Stellenbosch , South Africa, [2]Centre for Bioinformatics and Computational Biology, Faculty of Science, Stellenbosch, South Africa , [3]SAMRC-SHIP South African Tuberculosis Bioinformatics Initiative (SATBBI), Tygerberg, South Africa, [4]DST/NRF Centre of Exellence for Biomedical Tuberculosis Research , Tygerberg, South Africa, [5]SAMRC Centre for Molecular and Cellular Biology , Tygerberg, South Africa, [6]Division of Molecular Biology and Human Genetics, Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Tygerberg, South Africa

Data and results that are findable, accessible, interoperable, and reproducible (FAIR) are necessary for independent data verification and advancing science. This is even more important for complex data with high dimensionality. Single-cell RNA sequencing (scRNA-seq) is such a type of data. ScRNA-seq enables characterisation of individual cells in tissues providing substantial insight into the variability within cell types and the complexity of tissues. As with any other method technical noise, complications with library construction, variable cDNA capture, sequencing depth, batch effects, and bias affect the utility and interpretation of results. Amplification of the limited quantity of RNA material found within each cell compounds the variability. The diversity of scRNA-seq protocols make it challenging to create a universally applicable workflow. Most existing pipelines cater only to droplet-based data, excluding microwell and microfluidics data, and assume basic quality control has been applied. Robust and reproducible quality control and standardisation increase the utility and interpretability of data in downstream analyses and reduce the probability of false discovery of novel cell types or gene expression in cell sub-types. We are developing a robust scRNA-seq pipeline focused on improving initial pre-processing and quality control. After careful evaluation of existing workflow management systems (WMS), we selected Nextflow as the most versatile WMS for our purpose. We will test the robustness and reproducibility of the pipeline. To further ensure robustness and reproducibility we will containerise the pipeline.

61

EXPRESSION PROFILING OF INTERFERON-STIMULATED GENES IN PERIPHERAL BLOOD MONONUCLEAR CELLS FROM HEALTHY AND SARS-COV-2 INFECTED INDIVIDUALS

Ms Darisia Moonsamy[1], Dr Nikki Gentle[1]
[1]*University Of The Witwatersrand, Johannesburg, South Africa*

The world is currently in the midst of the coronavirus disease 2019 (COVID-19) pandemic. The majority of individuals experience mild symptoms, however a small percentage develop severe COVID-19. There is increasing evidence highlighting a potential link between severe COVID-19 and an imbalanced type-I interferon (IFN) response. The molecular mechanisms underlying this potential link are not well understood. Therefore, we used a single-cell approach to study the changes in gene expression of type-I IFN signaling molecules between healthy immune cells and those from individuals experiencing both mild and severe symptoms. To obtain high-quality gene expression data, single-cell RNA sequencing data (scRNA-seq) for peripheral blood mononuclear cells (PBMCs) were preprocessed using BUStools v0.39.2 and Seurat v3.2.0. Once normalised, the datasets were integrated in order to enable cell type-specific comparisons. To identify specific immune cells, a graph-based clustering algorithm was used. The clusters were then annotated using known cell-type markers and differential gene expression analysis was performed in a cell type-specific manner. Our analysis revealed potential dysregulation of multiple interferon-stimulated genes (ISGs) including interferon regulatory factor 3 (IRF3), an ISG involved in the transcription of type-I IFN. IRF3 was upregulated in both T and B cells in severe cases. Additionally, different ISG signatures were seen in monocytes compared to T and B cells. Thus, these observations highlight a role for ISGs in the development of severe COVID-19. In conclusion, the findings from this study warrant further research into these upregulated ISGs and their involvement in severe COVID-19.

45

NRPS-DEPENDENT SIDEROPHORE SYNTHETASE GENE CLUSTERS AND CHARACTERISTICS OF NRPS SIDEROPHORE SYNTHETASE GENES IN ARMILLARIA AND OTHER SPECIES IN THE PHYSALACRIACEAE

Mrs. Deborah L. Narh Mensah[1], Prof. Brenda D. Wingfield[1], Prof. Martin P.A. Coetzee[1]
[1]*Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Hatfield, South Africa*

Armillaria spp. are terrestrial ubiquitous basidiomycetes in the family Physalacriaceae, most of which are facultative necrotrophs of plants. In fungi, secondary metabolites are often pathogenicity or virulence factors. Genes involved in biosynthesis of these metabolites are usually contained in secondary metabolite gene clusters, such as nonribosomal polypeptide synthetase (NRPS) clusters. NRPSs contain domains, are either mono- or multi-modular, and produce peptides such as siderophores. Siderophores are high affinity ferric iron chelating compounds required for iron uptake under aerobic conditions. NRPS-dependent siderophore synthetase encoding clusters of Armillaria spp. and selected species from the Physalacriaceae were investigated using a comparative genomics approach. Siderophore biosynthesis by Armillaria spp. was also evaluated using CAS and split-CAS assays. The genomes studied contained one NRPS-dependent siderophore synthetase cluster each. All NRPS clusters were multimodular with the domain architecture (ATC)3(TC)2. NRPS clusters of the Armillaria spp. showed a high degree of microsynteny. Genes encoding L-ornithine-N5-monooxygenase were not identified in the NRPS clusters and atypical Stachelhaus codes were predicted for the A3 domains. Thus, we postulate that the siderophore biosynthesised by the identified NRPS clusters will be hydroxamates based on homology with characterised NRPSs, domain architecture, and predicted substrates of A-domains of the NRPS genes. However, the siderophores will differ from earlier characterised siderophores. Bioassays with strains of five Armillaria spp. revealed production of mainly hydroxamate siderophores and some catecholate siderophores. This is the first report on siderophore biosynthesis by Armillaria spp. Results from this and future studies will elucidate the molecular biology of fungal pathogenicity.

9

SCALABLE PRODUCTION OF H22(SCFV)-ETA' TARGETING CD64 IN ACUTE MYELOID LEUKEMIA (AML)

Mr Sikozile Ncembu[1], Dr Siew Tai[1], Professor Sikozile Ncembu[1], Professor Susan Harrison[1]
[1]*University of Cape Town, Cape Town, South Africa*

Cancer immunotherapy is a promising innovative and effective treatment for many forms of cancer. Amongst hematologic malignancies, acute myeloid leukemia (AML) remains an unmet medical need, as it is primarily treated by chemotherapy which is characterised by severe side effects. H22(scFv)-ETA' is an immunotherapeutic recombinant protein that has been successfully shown in vitro to be highly potent in selectively destroying CD64-positive dysfunctional myeloid tumour cells in AML. However, this study has only managed to produce 22(scFv)-ETA' in shake flasks, a scale that cannot supply sufficient quantity to carry out preclinical and/or clinal studies. Therefore, the current phase of this study focuses on optimizing the productivity of H22(scFv)-ETA' and conducting a scale-up production from shake flask to a 5 L stirred-tank reactor (STR). H22(scFv)-ETA' is recombinantly expressed in E. coli Bl21 (DE3) under osmotic stress and purified by metal ion affinity chromatography and size exclusion chromatography. Various scale-up criteria are evaluated to achieve effective batch and fed-batch fermentation processes. The therapeutic efficacy of H22(scFv)-ETA' is evaluated by several biological assays, including binding assays by flow cytometry and cytotoxicity by annexin V bioassay. Development of a successful scale-up production of H22(scFv)-ETA' is crucial, as it enables insights of a process to be established at a pilot scale and ultimately commercial scale in the context of biopharmaceutical manufacturing.

91

CHARACTERIZATION OF MYCOBACTERIUM BOVIS PERSISTERS IN SOUTH AFRICAN WILDLIFE

Miss Pamela Ncube[1], Dr Bahar Bagheri[1], Prof Michele Miller[1], Prof Samantha Sampson[1]
[1]*DST/NRF Centre of Excellence for Biomedical Tuberculosis Research, South African Medical Research Council Centre for Tuberculosis Research, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa*

In this study, we aimed to characterize Mycobacterium bovis persister formation upon in vitro acid stress which mimics the macrophage phagolysosome microenvironment. A total of 23 samples from naturally infected M. bovis wildlife species were successfully decontaminated, purified, and genotyped to the strain level. Twenty-two of the 23 isolates were successfully transformed with the Fluorescent Dilution reporter plasmid pTiGc. Three M. bovis strains were selected persister assay experiments, upon in vitro acid stress believed to enrich for persisters, as reflected by a sub-population of viable but non- or slowly replicating bacteria (VBNR). Laboratory strains, Severely Attenuated Mutant of M. tb::pTiGc and BCG::pTiGc, had the highest VBNR mean percentage of cells of 12.2 ± 1.5% and 7.2 ± 0.6% (± SD), respectively on day 4. The VBNR mean percentages of clinical M. bovis PN18067_1::pTiGc was 1.3 ± 0.1%, while PNMP20_1::pTiGc and PN18062_1::pTiGc had a similar very low mean percentage of 0.2 ± 0.0%. These data suggest that upon acid stress: (i) laboratory strains seem to have a higher propensity to form VBNR populations than three clinical isolates examined, (ii) M. bovis may demonstrate VBNR populations following acid stress, although these are very small, (iii) VBNR formation may vary depending on strain genotype. Notably, the VBNR populations detected under the conditions employed in this study were very small (0.5 ± 0.6 %) as expected, as well as supported by previous studies conducted on bacterial persisters and leaves this an open question to be further investigated in the future.

24

MOLECULAR MODELLING, DOCKING AND SIMULATION STUDIES TO IDENTIFY NOVEL INHIBITORS AGAINST THE MYCOBACTERIUM TUBERCULOSIS DRUG TARGET, RV2194 (QCRC)

Ms Lunathi Mihlali Ndlondlo[1]

[1]*South African National Bioinformatics Institute, Bellville, Cape Town, South Africa*

Mycobacterium tuberculosis, manifesting as tuberculosis (TB) in hosts, continues to be responsible for the most deaths caused by a single infectious agent in human history. Even with treatment regimes, such as the directly observed treatment strategy in place, the disease has continued to spread as well as mutate into strains that are resistant to presently available drugs. Due to the need for additional drug targets with a new mode of action, a search for essential genes in the M. tuberculosis metabolic pathway was carried out by Cloete et al., 2016. Amongst the essential genes, Rv2194 was identified and as an enzymatic form, it is named the probable ubiquinol cytochrome C reductase (QcrC). Rv2194 was found to be of importance in the electron transporting for M. tuberculosis survival, lacked a human homologue and had no experimental structure resolved. In this study, we aim at exploiting Rv2194 through structure-based computational methods to identify compounds that have potential to specifically target Rv2194 (QcrC) via a new mode of action. This is achieved through modelling along with preparing the unresolved structure, consequently followed by docking and molecular dynamics simulation to filter promising compounds that can bind to Rv2194 and inhibit the enzyme function, and as proof of concept we test these compounds
using in vitro assays for activity against M. tuberculosis.

33

STRUCTURE-BASED VIRTUAL SCREENING OF SELECTED MALARIA BOX COMPOUNDS AGAINST FALSTATIN, A MULTI-STAGED PROTEIN IN PLASMODIUM FALCIPARUM

Bolu Oladunjoye[1]
[1]*University Of The Western Cape - Bellville, Bellville, South Africa*

The structure-based virtual screening approach is a cost-effective and time-efficient strategy for discovering new drug candidates. This approach was explored to identify antiplasmodial compounds that could inhibit falstatin, a multi-staged drug target that facilitates the degradation of haemoglobin, erythrocytic invasion of merozoites and rupture of erythrocytes by free merozoites. In this study, 18, 000 antiplasmodial compounds retrieved from the Medicines for Malaria Venture database were docked on to the active site of a homology modelled falstatin. Compounds with high binding energy were ranked based on their ADMET (absorption, distribution, metabolism, excretion, and toxicity) profile and the top-ranked compound was subjected to molecular dynamics. Redocking analysis was done using the MM-GBSA and a pharmacophore hypothesis developed from the top compounds. The research identified TCMDC 131646 as a promising antimalarial candidate that has a high affinity for falstatin and good drug-like properties. Analysis of trajectories obtained from the MD simulation revealed that TCMDC 131646 interacted with and conferred stability on falstatin. The results demonstrate that TCMDC 131646 is predicted to be a drug-like and safe compound that can inhibit falstatin in Plasmodium falciparum. However, in-vitro and in-vivo studies are required to further substantiate the potential of this compound as an innovative antimalarial compound.

12

LONGITUDINAL AND COMPARATIVE ANALYSIS OF TUNISIAN NEWBORNS GUT MICROBIOTA ACCORDING TO DELIVERY MODE

Ms Mariem Hanachi, Ms Olfa Maghrebi, Dr Emna Harigua-Souiai, Pr Lamia Guizani-Tabbane, Pr Alia Benkahla, Dr Souiai Oussema
[1]*Institut Pasteur De Tunis, Tunis, Tunisia*

Microbiota colonization is a dynamic process that impacts the health status during one individual's lifetime. The composition of the gut microbiota of newborns is conditioned by multiple factors, including the delivery mode (DM). Nonetheless, the DM's influence remains uncertain and is still the subject of debate. In this context, the medical indication and the emergency of a cesarean delivery might have led to confounding conclusions regarding the composition and diversity of the neonatal microbiome. Herein, we used high-resolution shotgun sequencing to decipher the composition and dynamics of the gut microbiota of Tunisian newborns. Stool samples were collected from 5 Elective Cesarean section (ECS) and 5 vaginally delivered (VD) newborns at each of the following time points: Day0, Day15, and Day30. The ECS and VD newborns showed the same level of bacterial richness and diversity. In addition, the gut microbiota of both groups showed a transition profile dominated by Proteobacteria, Actinobacteria and Firmicutes. However, ECS showed an underrepresentation of Bacteroides and an enrichment of opportunistic pathogenic species of the ESKAPE group; specially from the second week. Besides revealing the intestinal microbiota of Tunisian newborns, this study provides novel insights in the microbiota perturbations caused by elective CS.

75

UNRAVELLING THE WHOLE GENOME SEQUENCE OF THE SOUTHERN GROUND HORNBILL (BUCORVUS LEADBEATERI)

Miss Jasmin Patel[1], Dr Pieter De Maayer[1], Dr Jean Mollett[1], Dr Angela Botes[1]
[1]*University Of The Witwatersrand, Johannesburg, South Africa*

The Southern Ground Hornbill (Bucorvus leadbeateri) is a large species of bird belonging to the family Bucorvidae and order Bucerotiformes (hornbills), which is restricted to southern Africa. Being both culturally and ecologically important to the savanna and grassland biomes these birds are known as 'thunder' or 'rain' birds whose conservation can benefit other threatened species inhabited in the same territory. Unlike other hornbill species, the Southern Ground Hornbill lead a largely sedentary lifestyle and are carnivorous. Furthermore, B. leadbeateri has few natural enemies that prey on them. Irrespective of this, the Southern Ground Hornbill population is declining due to habitat loss and cultural beliefs. As such, it is listed as vulnerable by the International Union for the Conservation of Nature (IUCN), while in South Africa it is listed as endangered. Recently, B. leadbeateri has been the subject of extensive conservation efforts to increase their population. While numerous studies on the lifestyle and behavioural patterns of these birds have been conducted, knowledge regarding their genetic potential is limited. This study focuses on the whole genome sequencing of the Southern Ground Hornbill. The genome sequence will serve as the genetic backbone for the development of molecular markers that will assist in population genetic studies and concomitantly in conservation efforts. Furthermore, comparative genomics of the B. leadbeateri genome with the available genomes of B. abyssinicus as well as other flighted hornbill species will provide insights into these unique aspects of biology, ecology, behaviour, and evolution of the Southern Ground Hornbill.

60

DIFFERENCES AND COMMONALITIES INDUCED BY PMA AND 1,25(OH)2D3 IN THP-1 CELLS DURING MONOCYTE-TO-MACROPHAGE DIFFERENTIATION

Miss Kelda Perumal[1], Mr Duncan Cosser[1], Dr Vanessa Meyer[1], Dr Nikki Gentle[1]
[1]*University Of The Witwatersrand, Johannesburg, South Africa*

Monocytes can differentiate into macrophages in response to cytokines and/or pathogens. The process of monocyte-to-macrophage differentiation is studied in vitro through the use of promonocytic model cell lines, like the THP-1 cell line, where commonly used differentiation inducing agents include phorbol-12-myristate-13-acetate (PMA) and the active metabolite of vitamin D3, (1,25(OH)2D3). This study used RNA sequencing (RNA-seq) to characterise changes in gene expression induced in THP-1 cells by these two differentiation inducing agents during the process of monocyte-to-macrophage differentiation. Gene expression was quantified in differentiated and undifferentiated THP-1 cells, treated with both PMA and 1,25(OH)2D3 using Salmon (v1.2.1) and differential gene expression analysis between differentiated and undifferentiated cells was performed using DESeq2 (v1.30.1). Finally, gene ontology and pathways analysis was performed using the R package, clusterProfiler (v3.1.4). In this way, changes in gene expression uniquely and commonly induced by both these differentiation inducing agents were identified.

37

NOVEL GENE DISCOVERY IN A XHOSA FAMILY WITH PARKINSON'S DISEASE

Miss Nikita Simone Pillay[1], Prof. Soraya Bardien[2], Prof. Alan Christoffels[3]
[1]SANBI, University Of The Western Cape, Cape Town,, South Africa, [2]Division of Molecular Biology and Human Genetics, Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town,, South Africa, [3]SA Medical Research Council Bioinformatics Unit, SANBI, University Of The Western Cape, Cape Town,, South Africa

Parkinson's disease (PD) is a neurodegenerative disorder with complex genetic aetiology. The limited number of genetic screening PD studies in Sub-Saharan African (SSA) populations has also typically indicated an unknown cause of the disease. Whole-exome sequencing (WES) approaches can be utilized to find novel pathogenic mutations/genes in families portraying Mendelian inheritance. This study aims to identify a potentially novel pathogenic variant/gene in the first Next-Generation Sequencing-based analysis of a Xhosa family with PD. WES was performed on two PD-affected siblings and their two unaffected siblings on a HiSeq 4000 at the Mayo Clinic Core Facility, USA. Raw sequencing data was analyzed through a bioinformatic workflow incorporating BWA-MEM (GRCh38/hg38 reference alignment), GATk HaplotypeCaller (variant calling) and EnsemblVEP (variant annotation). VCFs were scanned for variants in both known (n=24) and putative (n~80) PD genes to eliminate known genetic causes. VCF files were then filtered to include heterozygous (assumed autosomal dominant inheritance), exonic, non-synonymous variants with a PHRED QS > 30, present in population databases with a MAF < 0,01 and a CADD > 15. A total of 68 variants, shared between the affected individuals, were identified. These candidate genes were then subjected to gene/protein expression analysis to determine neuro-specific tissue and pathway expression. Twenty-four variants were prioritized and are currently undergoing Sanger sequencing, prior to further population-specific screening and functional analysis. Identifying novel PD genes may improve clinical diagnoses and treatment options by providing insight into unknown PD molecular mechanisms, detecting rare PD bio-markers and determining novel drug targets.

71

ELUCIDATING THE ROLES OF EGRARF1, EGRARF10 AND EGRMYB14 IN THE REGULATION OF XYLEM FORMATION IN EUCALYPTUS

Ms Ipeleng Randome[1], Dr Raphael Ployet[1], Mrs Adri Veale[1], Mr N Nangammbi[1], Ms Aysha Ndou[1], Prof Jaqueline Grima-Pettenati[2], Prof Alexander Myburg[1], Dr Steven Hussey[1]
[1]University Of Pretoria, Pretoria, South Africa, [2]University of Toulouse, Castanet-Tolosan, France

Wood (secondary xylem) is the most abundant raw material produced by plants on earth. In industry wood is important for the production of pulp, paper, timber and biofuels. Regulation of wood development occurs through a transcriptional regulatory network encompassing at least 3 tiers of transcription factors. The identification of novel transcription factors and upstream regulators of this three-tiered model is an ongoing area of research. We used a network-based lines-of-evidence (LOE) pipeline to identify Eucalyptus transcription factors with possible roles in wood formation. The pipeline involved using five major categories of Eucalyptus RNA-seq datasets (n = 15) to calculate co-expression scores for potential candidates and known secondary cell wall genes. Using this approach, we identified 3 candidates (EgrARF1, EgrARF10 and EgrMYB14) with strong links to wood formation. Two Eucalyptus clones (CG96 and ZG14) were transformed with EgrARF1 and EgrARF10 overexpression constructs respectively, yielding 9 EgrARF10-OX lines and 8 EgrARF1-OX lines. Root length and diameter measurements of the overexpression lines and empty vector controls were taken at 10 weeks post-acclimation. Additionally, EgrARF10 was overexpressed in poplar and a greenhouse growth trial was conducted over a period of three months. Microscopy, pyrolysis GC/MS and RNA seq analysis are currently in progress. For EgrMYB14, Eucalyptus seedlings were successfully transformed with the candidate using the hairy root strategy and the composite plants are currently being acclimated in soil. The results from the study will aid in understanding the regulation of wood formation in tree species such as E. grandis.

72

HYBRID DE NOVO SEQUENCING AND ASSEMBLY OF THE BEAUMONT/ HAES 695 AND SANTA ANNA ACCESSIONS (MACADAMIA SP.)

Miss Mary Ranketse[1,2], Dr C,A Hefer[3], Dr Rian Pierneef[1], Dr Gerda Fourie[2], Prof A,A Myburg[2]
[1]Biotechnology Platform, Agricultural Research Council – Onderstepoort Veterinary Institute, Pretoria, South Africa, [2]Department of Biochemistry, Genetics and Microbiology, Forestry and Agriculture Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa, [3]AgResearch Ltd, Lincoln, New Zealand

Commercial macadamia cultivars are mainly hybrids derived from two species, M. integrifolia and M. tetraphylla. Macadamia tree nuts are the most expensive in the world and are also nutritious and healthy, with high fats and low sugar content. South Africa commercially plants various cultivars imported originally from Australia and Hawaii, and is the world's largest producer of the nuts. Although a successful crop with a high quality nut product, genomic resources and molecular breeding are limited for macadamia. Genome assemblies assist in developing genetic markers that can be used to identify variation amongst accessions. Furthermore, high-quality genomes of elite accessions can be used to understand how genomics may contribute to traits of interest for breeding and crop improvement. The objective of this study was to generate high-quality genome assemblies of two macadamia accessions, Beaumont/HAES 695 (M. integrifolia x M. tetraphylla hybrid planted extensively in South Africa) and Santa Anna (M. tetraphylla representative). This was done through high coverage (>80x) sequencing of short-reads on the Illumina HiSeq, long-reads in the Oxford PromethION and optical mapping on the BioNano Saphyr instruments. Sequence data was combined to successfully assemble the Beaumont and Santa Anna genomes. Genomic variation was identified in the form of SNPs and structural variants present within the genomes. This study will add to the existing resources for macadamia and assist researchers and breeders to better understand macadamia genomics for accelerated and sustainable future breeding initiatives.

14

GENOMIC AND EPIGENETIC COMPARISON OF BACILLUS STRAINS SUITABLE FOR PLANT PROTECTION

Prof Oleg Reva[1]
[1]*University Of Pretoria, Pretoria, South Africa*

Strains of the Bacillus subtilis taxonomic group are widely applied as ecologically safe biopesticides and plant growth promoting rhizobacteria. A comparative study is reported here on the whole genomes SMRT PacBio sequencing of 14 natural Bacillus isolates characterized with distinct phenotypes and plant growth promoting activities, which were assessed by Illumina sequencing of total RNA samples and by profiling epigenetic modifications of the chromosomal DNA. A uniform organization of the genomes of all these isolates was observed whereas the patterns of gene expression and transcriptional regulation under the effect of root exudate stimuli were quite different. It was assumed that epigenetic modification of the chromosomal DNA may be behind the alternative patterns of gene regulation. Methylated nucleotides were identified using the SMRT-Link v.10-2021 software tools and in-house scripts on Python. It was a remarkable discovery that different strains possessed unique DNA restriction-modification systems responsible for methylation of either cytosine or adenine residues at strain-specific motifs. Appearance of methylated nucleotides in promote regions of different genes may explain the observed differences in the gene regulation patterns at similar growth conditions. Although there are many publications on the role of different genes in plant growth promoting activities of rhizobacteria, this is the very first report of alternative patterns of genome methylation resulting from replacements and horizontal exchange of bacterial restriction-modification systems, which can re-program the genome orchestra causing significant changes in phenotype and bioactivities being of interest for biotechnology.

26

SEQUENCE VARIATION IN THE CYTOPLASMIC MALATE DEHYDROGENASE (MDHC) GENE AND ASSOCIATION WITH GROWTH PERFORMANCE FOR SOUTH AFRICAN ABALONE

Ms Gheda Ali Salum[1], Mr Thendo Stanley Tshilate[1], Dr Clint Rhode[1]
[1]*Stellenbosch University, Stellenbosch , South Africa*

South African abalones (Haliotis midae) are economically relevant aquaculture species, accounting for two-thirds of South African aquaculture investment and more than 90% of the industry's revenue. However, due to climate change, environmental temperatures are rising. Abalones then suffer from heat stress, becoming more susceptible to infection, resulting in a decline in growth rate. Understanding genetic variation in response to changing environmental conditions could lead to genetic improvement and resilience. Therefore, this study aimed to quantify the extent of genetic variation in the heat tolerance-associated gene cytoplasmic malate dehydrogenase in abalone populations as well as assess its association with commercial cultured abalone growth performance. A total of 68 abalone samples were sequenced, with 28 being wild and 40 being cultured. For the wild population, specimens were sampled along a temperature gradient from the east coast, through the south coast, and finally to the west coast. The cultured abalones were four first-generation families from a consortium of five aquaculture facilities. Genetic diversity analyses were used to examine SNP sequence variation. Meanwhile, a case-control model was used to conduct SNP association analysis of growth performance, with size (large or small) as the categorical phenotype. Heterozygous deficiency was found within the four identified SNPs, potentially indicating directional selection. Private alleles were found in the east coast population, which may represent adaptation. Ultimately, no significant association was found between the SNPs and growth performance. This research could lead to more effective genetic management strategies for South African abalones.

35

THE RELATIONSHIP BETWEEN GENETIC VARIATION IN NEUROBEHAVIORAL DEVELOPMENTAL GENES, AND GROWTH TRAITS IN DUSKY KOB (ARGYROSOMUS JAPONICUS)

Mr Joshua Sampson[1], Dr Clint Rhode[1]
[1]*Department of Genetics, Stellenbosch University, Stellenbosch, South Africa*

The ever-increasing demand for food production has exhausted global fisheries, necessitating the transition towards aquaculture to meet these demands and facilitate the conversation of marine species. One such organism is the dusky kob (Argyrosomus japonicus), which is in the early stages of establishment for aquaculture in South Africa. Proper management is essential to the success of any selective breeding scheme, with the implementation of molecular breeding techniques using genetic markers increasing the efficiency (i.e., faster growth) and profitability of aquaculture. The complex nature of developmental traits such as growth rate requires the inclusion of environmental factors when attempting to characterise genetic variation. Individual responses to stress, and their behavioural plasticity, may serve as key selectors for improved fitness and growth. By using an annotated draft genome for the dusky kob, primers were designed for the exons, 5'- and 3'-UTRs for five key genes associated with developmental and neurological pathways in vertebrates. Thirteen putative SNPs were discovered in 3 of the genes, from which four were found to be significantly associated (adjusted for family structure) with increased growth. Three of the four were synonymous, with the other marking a change from Serine to Proline in the fourth exon of BDNF. Linkage analysis found strong LD in 3 of these SNPs, and an association of haplotype -AGT- with larger individuals. This study has produced potentially valuable genetic resources which could be implemented in marker assisted selection (MAS), aiding in the acceleration of effective aquaculture for dusky kob in South Africa.

56

KRAS, EGFR, ALK AND ROS1 MUTATION CONFIRMATION IN NON-SMALL CELL LUNG CARCINOMA CASES

Dr Natasha Sanabria[1], Ms Ntebogeng Kgokong[1], Dr Naseema Vorajee[1]
[1]*NIOH / NHLS, Johannesburg, South Africa*

Prognosis of lung cancers with conventional cytotoxic chemotherapy treatment is poor. However, certain sub-types of non-small cell lung carcinoma (NSCLC), determine the patient's response to Tyrosine kinase inhibitor (TKI)-based treatments. Since generic drugs will soon be on the market, state hospitals will be able to allow patients to receive precise treatment, of a specific therapy, which is targeted to their own genetic profile. The most useful biomarkers for predicting the efficacy of this targeted therapy are somatic genome alterations known as "driver mutations". Mutations warrant investigation and must be monitored to provide specific care and understand the carcinogenesis. A total of 3901 histology reports were accessed from 1st January 2008 to 30th June 2014. Of these, 111 tissue biopsies of patients from the Charlotte Maxeke Johannesburg Academic Hospital (CMJAH) and the Helen Joseph Hospital (HJH), were available for mutational analysis. Immunohistochemistry (IHC) and quantitative polymerase chain reaction (qPCR) assays specific for Kirsten rat sarcoma viral oncogene homolog (KRAS), Epidermal growth factor receptor (EGFR), Anaplastic lymphoma kinase (ALK) translocation/EMLF fusions and c-ros oncogene 1 (ROS-1) mutations were used to identify and confirm NSCLC cases. A match could be confirmed with qPCR if the IHC samples were identified as positive for EGFR mutation when detected with the EGFR-111 antibody. This implies a population specific response, where this antibody might be better suited for the SA-Black population that made up most of the cohort. Different bioinformatics approaches were used to investigate the types of mutations, their locations and antibody detection.

50

THE ENCEPHALARTOS NATALENSIS-CYANOBACTERIAL CORALLOID ROOT PARTNERSHIP FOR NITROGEN ACQUISITION

Miss Cassandra Bianca Schoeman[1], Dr Danielle Roodt[1], Dr Desre Pinard[1], Prof. Almuth Hammerbacher[1], Prof. Thulani Makhalanyane[1], Prof. Eshchar Mizrachi[1]
[1]*University of Pretoria, Pretoria, South Africa*

Mutualistic partnerships between plants and microbes arose during land plant terrestrialisation around ~450 million years ago, and have led to the widespread existence of symbiosis among plants, driven by the need for plants to access scarce nutrients. Plant-cyanobacterial symbioses are amongst the major evolutionary innovations linked to the acquisition of nitrogen via partnerships with microorganisms. Cyanobacterial symbiosis evolved independently across unrelated lineages of land plants, unlike those involving symbioses with arbuscular mycorrhizae, rhizobia, Frankia diazotrophic bacteria and flowering plants. However, information regarding the precise genes and processes linked to cyanobacterial partnerships are dated and lacking. We hypothesise that genes and pathways have been coopted multiple times for terrestrial symbioses. We investigated gene expression in symbiotic and control tissues from four cyanobacterial partnership plant lineages, including a comprehensive tissue-specific transcriptomic dataset from cycad (Encephalartos natalensis). We analysed genes preferentially expressed in symbiotic tissue from Anthoceros punctatus, Azolla filiculoides, E. natalensis and Gunnera perpensa, along with known genes involved in other symbiotic partnerships. These data were combined with light and transmission electron microscopy on selected cyanobioses to confirm an active nitrogen-fixing symbiosis. The results show high conservation of common symbiotic pathway genes in the cycad-cyanobacterial partnership, despite their absence or loss in other cyanobiosis plant lineages. This result suggests that cycads may have subfunctionalised the same genes and pathways for its cyanobacterial partnership from its existing arbuscular mycorrhizal symbiosis, apart from those involved in carbon-nitrogen source biosynthesis.

30

GENOMIC CHARACTERISATION OF ACINETOBACTER BAUMANNII ASSOCIATED WITH NEONATAL SEPSIS AND STILLBIRTHS IN A SOUTH AFRICAN POPULATION

Mr Bonginkosi Shabangu[1], Dr Courtney Paige Olwagen[2], Professor Shabir Ahmed Madhi[2]
[1]University of the Witwatersrand - Vaccine And Infectious Diseases Analytics Research Unit, Faculty of Health Sciences, Johannesburg, South Africa, [2]South African Medical Research Council, Division of Research Capacity Development, Cape Town, South Africa

Multi-drug resistant hospital-acquired organism including Acinetobacter baumannii (A. baumannii) overwhelmingly contributes to neonatal deaths, particularly in low-middle income countries (LMIC). Concerns continue to grow that without significant interventions, hospital-acquired infections will soon be untreatable. Strategies for preventing A. baumannii associated neonatal deaths and stillbirths are thus needed to optimise the health of the mother-baby dyad. While the antigenic structure of A. baumannii organisms is diverse and complex, studies investigating virulence factors and clonal groups associated with pathogenicity are scarce, especially in LMIC settings where molecular epidemiology is poorly characterised. Thus, the study will undertake whole-genome sequencing (WGS) of A. baumannii isolates associated with neonatal invasive disease (including from fatal cases) and stillbirths. Whole-genome sequencing will be used to genetically characterise A. baumannii neonatal sepsis isolates that have been collected during several studies currently underway at Chris Hani Baragwanath Academic Hospital and Rahima Moosa Mother and Child Hospital in the city of Johannesburg, South Africa. Data generated from this study will allow us to define the population structure of A. baumannii, its antibiotic resistance gene repertoire, the size and content of its pan-genome, and the phylogenetic relationships amongst strains associated with neonatal invasive disease. Thus, providing means for determining the probability and the severity of infection these strains may cause, and possibly further identifying neonates that are at a high risk of demising. Genomic characterisation studies are needed to develop prevention approaches (vaccine, monoclonal antibodies, innovative infection prevention, and control tools) to reduce antimicrobial resistance to sepsis in neonates.

78

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF NLR PROTEINS DIFFERENTIALLY EXPRESSED IN CASSAVA PLANT INFECTED WITH SACMV

Dr Bulelani Sizani, Mr Keelan Krinsky, Prof Chrissie Rey
[1]*University of Witwatersrand, Johannesburg, South Africa*

• Yield of the cassava crop is reduced greatly by cassava mosaic viruses particularly in the sub-Saharan region, which is dependent greatly on cassava for food security. Identification of disease resistance genes in cassava such as NLR encoding proteins is one of the strategies to combat CMD. These genotypes can be selected by farmers and be used in breeding techniques for better improvement of cassava crops against CMD.
• Here, we performed structural-based multiple sequence alignment and homolog modelling of differentially expressed NLR proteins upon SACMV infection during leaf development in susceptible T200 and tolerant TME3 genotypes. Moreover, selected regions on the CC-NLR and TIR-NLR C-terminal proteins associated with SACMV response were mutated using CRISPR/Cas9 to verify their function in protoplasts of susceptible/resistant genotypes.
• Thirty-five genes were differentially regulated in susceptible/tolerant cassava genotypes. However, there is very low genetic variation in six sequenced NLR proteins from 24 susceptible, tolerant or resistant genotypes.
• Our results show that cassava NLR protein share conserved structural function with known NLR proteins providing another source of CMD resistance in cassava.

82

SIMULATION OF TRANSCRIPTION FACTOR DYNAMICS INVESTIGATES THE EFFECT OF IN VIVO CONDITIONS ON COGNATE SITE SEARCH.

Mr Benilo Slabbert[1]
[1]*Stellenbosch University, Cape Town, South Africa*

Transcription factor (TF) DNA binding and its role in transcription regulation is essential to every living creature throughout its life. The timely arrival of a TF at its cognate binding site is crucial for any function dependent on a gene product. Recent evidence suggests that non-specific chromatin binding, the same property which enables a mode of transport called facilitated diffusion (FD), is necessary for a large range of pioneer factors tasked with invading Eukaryotic heterochromatin and initiating silenced regulatory programs. We hypothesize that TFs with non-specific binding capabilities locate their targets more readily than those lacking it, particularly in vivo with conditions including coiled, highly concentrated and nucleosome occupied genomes. Short time scale simulations of individual TFs searching for their cognate sites, with and without FD capabilities, are performed under the cellular conditions listed above. The DNA is modelled physically with conformational accuracy using the Worm-like chain (WLC) model. Non-specific binding strengths, DNA compaction and accessible chromatin are varied to investigate the search dynamics these physiologically relevant constraints impose. The results from the experiments described above are used to assess the possibility of epigenetic features acting as an index for pioneer TFs to initiate regulatory programs. It is concluded that conditions common to silenced chromatin including FD search dynamics by TFs, DNA compaction and occlusion by nucleosomes support that the novel functional role of an epigenetic index may not only be feasible, but advantageous to rapid and reliable transcriptional program initiation.

79

PHYLOGENETIC ANALYSIS OF CIRCADIAN CLOCK GENES IN CHRYSOPORTHE SPECIES

Mr. Michal Slupski[1,2], Prof. Markus Wilken[1,2], Prof. Nicolaas van der Merwe[1,2]
[1]*Forestry and Agricultural Biotechnology Institute, Pretoria, South Africa,* [2]*University of Pretoria, Pretoria, South Africa*

Circadian rhythms constitute an organism's ability to anticipate and respond to regularly changing environmental stimuli. These endogenous rhythms are controlled by a network of circadian clock proteins. In fungi, this network is well conserved and involves a central transcriptional/translational negative feedback loop. Core clock genes that are important in this loop include frq, wc-1, wc-2, and frh. Recent research showed that circadian rhythms influence the virulence of fungi, which highlights a need for research into the circadian clocks of pathogenic fungi. The fungal pathogens of the genus Chrysoporthe are good candidates for such studies, due to their importance in the infection of economically important plants in Southern Africa, such as Eucalyptus. We aimed to identify and characterise the evolution of circadian clock genes within Chrysoporthe species and the order Diaporthales in which they occur. The identities of the genes were confirmed via PCR and Sanger sequencing. The evolution of these genes was investigated via phylogenetic analysis of both the genes and their protein functional domains. These gene trees were incongruent to the known Chrysoporthe species trees, but genetic variation was insufficient to separate the isolates into distinct clades. Additionally, the functional domains were highly conserved within the order Diaporthales. Therefore, we conclude that these genes are highly conserved within the genus Chrysoporthe and in the order Diaporthales. From these results, we have established a foundation for future research into the impact of circadian rhythms on fungi, their hosts, and on experimental design.

51

THE ULTRASTRUCTURE AND ROLE OF PLASTIDS IN CARBON PARTITIONING DURING XYLOGENESIS IN VND7-INDUCIBLE ARABIDOPSIS

Miss Medha Sood[1], Dr Desrè Pinard[1], Dr Victoria Maloney[1], Prof Vida Van Staden[1], Prof Eshchar Mizrachi[1]

[1]*University Of Pretoria, Hatfield, South Africa*

Lignin is a fundamental component of the plant secondary cell wall (SCW) and comprises ?30% of lignocellulosic biomass, which is a renewable feedstock for timber, paper, pulp, biomaterials and biofuel production. Although plastids are the exclusive sites for the synthesis of phenylalanine - a lignin monomer precursor - the role of organelles in carbon partitioning between polysaccharides and lignin during xylogenesis has been largely ignored. Previous research in our laboratory has shown that a developing-xylem specific plastid (the xyloplast) is uniquely regulated to facilitate carbon partitioning during SCW biosynthesis. Plastid ultrastructure and function during their differentiation is linked, but we know little about the morphology of xyloplast differentiation in trees as the SCW in woody tissues of plants hinders detailed examination. The aim of this study was therefore to dissect the ultrastructural changes that accompany induced xyloplast differentiation from chloroplasts. We utilised a VND7-inducible system in Arabidopsis, for which transcriptomic and some metabolomic data is publicly available. In this system, non-vascular cells differentiate into xylem vessel elements upon induction, resulting in SCW growth. We show that vessel element induction results in major ultrastructural changes in chloroplasts, with an intermediate morphology typical of plastid differentiation. This supports the classification of the xyloplast as a distinct plastid type. Integration of morphology and gene expression analysis showed that carbon partitioning linked to plastid terminal differentiation and degradation play a key role in xylogenesis. These findings shed insight on vascular plant development and evolution, as well as advanced biotechnological applications in engineering biomass crops.

27

PARKINSON'S DISEASE, PLINK, AND PUTATIVE DISEASE SUSCEPTIBILITY VARIANTS

Miss Kathryn Step[1], Ms Alvera Vorster[1], Prof Soraya Bardien[1]
[1]*Stellenbosch University, Cape Town, South Africa*

Parkinson's disease (PD) is an incurable, degenerative disease affecting nerve cells. The cardinal features of PD include loss of balance, muscle tremors, rigidity, and bradykinesia. In the present study, we used raw data generated from a collaborative project known as COmprehensive Unbiased Risk Factor Assessment for Genetics and Environment in Parkinson's Disease (Courage-PD) whose goal was to identify PD-associated variants. The NeuroChip array, used to genotype the study participants, contains a total of 306 670 tagging variants and 179 467 custom content variants, including 348 associated with PD. South African cases and controls were genotyped on the array and the aim of this study was to analyse the genotyping data. The South African dataset comprises of 464 cases and 288 controls. The quality control (QC) and statistical procedures were completed using PLINK. The results were visualized using RStudio. QC steps conducted included missingness of single nucleotide polymorphisms (SNPs) and individuals (<1%), sex discrepancy, minor allele frequency (<1%), Hardy-Weinberg equilibrium, heterozygosity, relatedness, and population stratification. Further analysis included controlling for population stratification and statistical tests of association. Our preliminary results showed 3 SNPs of interest (rs7695365, rs17835005 and rs3835037) showing significant association with the disease, and an additional 4 SNPs approaching significance. Population stratification analysis showed clear clusters for the European ancestry and Mixed ancestry participants. Further analysis of the prioritised variants is ongoing. Genetic research involving unique Sub-Saharan African populations, such as the present study, is essential since these populations are severely underrepresented in PD research.

59

HOST GENETIC FACTORS CONTRIBUTING TO SUSCEPTIBILITY TO COVID-19

Ms Chrystal Steyl[1], Prof Marlo Moller[1], Dr Desiree Petersen[1,2], Dr Brigitte Glanzmann[1,3], Prof Craig Kinnear[1,3], Dr Reinhard Hiller[4], Dr Elouise Kroon[1], Dr Caitlin Uren[1], Dr Richard Glashoff[5,6], Dr Helena Cornelissen[6,7], Dr Judith Hornby[8], Dr Zivanai Chapanduka[6,7], Dr Ibtisam Abdullah[6,7], Dr Helena Rabie[9], Dr Deepthi Abraham[9,10], Dr Shane Murray[8], Dr Jeanne Korsman[8], Dr Kate Webb[11], Dr Tongai Maponga[12], Dr Sian Hemmings[13], Ms Nina Vorster[1], Dr Aubrey Shoko[8], Dr Gert van Zyl[6,12], Mr Hendrik La Grange[8], Ms Natrisha Damons[4]

[1]Stellenbosch University, Division of Molecular Biology and Human Genetics, Cape Town, South Africa, [2]South African Medical Research Council (SAMRC), Cape Town, South Africa, [3]SAMRC Genomics Centre, Cape Town, South Africa, [4] Artisan Biomed , Cape Town, South Africa, [5]Stellenbosch University, Division of Medical Microbiology and Immunology, Cape Town, South Africa, [6]National Health Laboratory Service (NHLS), Cape Town, South Africa, [7]Stellenbosch University, Division of Haematological Pathology, Cape Town, South Africa, [8]Centre for Proteomic and Genomics Research (CPGR), Cape Town, South Africa, [9]Stellenbosch University, Department of Paediatrics and Child Health, Cape Town, South Africa, [10]Stellenbosch University, Division of Rheumatology and Immunology, Cape Town, South Africa, [11]University of Cape Town, Department of Paediatrics and Child Health, Cape Town, South Africa, [12]Stellenbosch University, Division of Medical Virology, Cape Town, South Africa, [13]Stellenbosch University, Department of Psychiatry, Cape Town, South Africa

COVID-19, caused by the severe acute respiratory syndrome corona virus 2 (SARS-CoV-2), is a disease that has been declared a global pandemic. To date, most research has been dedicated to elucidating the disease pathology and pathogenesis, however knowledge gaps such as the inter-individual variability in the host immune response remain. Most individuals are asymptomatic or develop mild disease, however, some do develop severe and critical disease. What is of particular interest are individuals who develop these extreme phenotypes despite being young without any underlying comorbidities. Further, a novel rare and severe clinical syndrome in previously healthy asymptomatic children related to SARS-CoV-2 has been described, namely Multisystem Inflammatory Syndrome in Children (MIS-C). A dysregulation of the immune response in these extreme phenotypes is clear and host genetic factors have been proposed as a contributor which may leave these individuals more susceptible to COVID-19. In the current study, whole genome sequencing (WGS) and data analysis of individuals who have been admitted with severe/critical COVID-19 disease or MIS-C at Tygerberg Hospital and Red Cross Children's Hospital will be performed. The generated WGS data will be submitted to an in-house prioritisation tool, TAPER, which will identify rare genetic variants associated with immune system pathways. This will provide the first direct association between MIS-C and genetic susceptibility in Africa while also providing a better understanding of the genetic basis of COVID-19 pathogenesis in humans. It may also lead to additional research involving novel prophylactic treatments and revised treatment strategies.

49

MODELLING THE POPULATION DYNAMICS OF CRISPR-CAS9 GENE DRIVE SYSTEMS IN SIREX NOCTILIO

Mr Hannes Strydom[1], Prof Rachid Ouifki[2], Prof Michael Chapwanya[2], Prof Brett Hurley[1], Prof Bernard Slippers[1]

[1]Department of Biochemistry, Genetics and Microbiology, Faculty of Natural and Agricultural Sciences, Tree Protection Cooperative Program, Forest and Agricultural Biotechnology Institute, University of Pretoria, Pretoria,, South Africa, [2]Department of Mathematics and Applied Mathematics, Faculty of Natural and Agricultural Sciences, University of Pretoria, Pretoria,, South Africa

Sirex noctilio is an invasive pest of pine that has caused significant economic damage in South Africa and many other countries. Current management tools are not always efficient and consequently there is a need for more targeted control measures. An emerging tool for pest management is the use of gene drive systems. In this study we investigate the use of CRISPR-Cas9 gene drive systems in the management of S. noctilio in South Africa. As a first step, we developed a model for the population dynamics of S. noctilio, using historical data and incorporating the influence of two main biological control agents of the pest. We then modeled the influence of two different CRISPR-Cas9 models on the population dynamics of S. noctilio, namely a Toxin-Type CRISPR (TTC) model and Complementary Sex Determination CRISPR (CSDC) model. Each model is also used to simulate different introduction strategies to estimate the effectiveness of the gene drive system. Results suggest that both CRISPR-Cas9 gene drive systems would be effective at controlling the population growth of S. noctilio, but that the TTC model develops complete resistance after a few generations. The CSDC model develops no resistance by causing non-viability in individuals that contain a resistant allele. Overall, the models suggest that CRISPR-Cas9 gene drive systems would serve as a viable form of control, but that further research should be done to identify possible target genes and optimize introduction strategies.

87

THE IN PLANTA EXPRESSION OF AUSTROPUCCINIA PSIDII IN RESISTANT AND SUSCEPTIBLE EUCALYPTUS GRANDIS

Miss Shae Swanepoel[1], Dr Caryn Oates[1], Dr Louise Shuey[2], Dr Geoff Pegg[2], Prof Sanushka Naidoo[1]
[1]Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria, South Africa, [2]Department of Agriculture and Fisheries, Queensland Government, Queensland, Australia

Austropuccinia psidii, commonly known as myrtle rust, is an obligate, biotrophic rust pathogen that causes rust disease on a broad host range of Myrtaceae species. Eucalyptus grandis, a widely cultivated hardwood Myrtaceae species, is susceptible to A. psidii infection, with this pathogen threatening both their natural range and various forest plantations across the world. This study aimed to investigate the A. psidii transcriptomic responses in resistant and susceptible E. grandis at four time points. RNA-seq reads were mapped to the A. psidii reference genome to quantify expressed genes at 12-hours post inoculation (hpi), 1-, 2- and 5-days post inoculation (dpi). A total of eight hundred and ninety expressed genes were found, of which forty-three were candidate effector proteins. These included a rust transferred protein (RTP1), expressed in susceptible hosts at 5-dpi and a hydrolase protein expressed in both resistant and susceptible hosts over time. Functional categorisation of expressed genes revealed processes enriched in susceptible hosts, including malate metabolic and malate dehydrogenase activity, implicating oxalic acid in disease susceptibility. These results highlight putative virulence or pathogenicity mechanisms employed by A. psidii to cause disease and provides the first insight into the molecular responses of A. psidii in E. grandis over time.

88

THE IN PLANTA EXPRESSION OF AUSTROPUCCINIA PSIDII IN RESISTANT AND SUSCEPTIBLE EUCALYPTUS GRANDIS

Miss Shae Swanepoel[1], Dr Caryn Oates[1], Dr Louise Shuey[2], Dr Geoff Pegg[2], Prof Sanushka Naidoo[1]
[1]Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa, [2]Department of Agriculture and Fisheries, Queensland Government, Queensland, Australia

Austropuccinia psidii, commonly known as myrtle rust, is an obligate, biotrophic rust pathogen that causes rust disease on a broad host range of Myrtaceae species. Eucalyptus grandis, a widely cultivated hardwood Myrtaceae species, is susceptible to A. psidii infection, with this pathogen threatening both their natural range and various forest plantations across the world. This study aimed to investigate the A. psidii transcriptomic responses in resistant and susceptible E. grandis at four time points. RNA-seq reads were mapped to the A. psidii reference genome to quantify expressed genes at 12-hours post inoculation (hpi), 1-, 2- and 5-days post inoculation (dpi). A total of eight hundred and ninety expressed genes were found, of which forty-three were candidate effector proteins. These included a rust transferred protein (RTP1), expressed in susceptible hosts at 5-dpi and a hydrolase protein expressed in both resistant and susceptible hosts over time. Functional categorisation of expressed genes revealed processes enriched in susceptible hosts, including malate metabolic and malate dehydrogenase activity, implicating oxalic acid in disease susceptibility. These results highlight putative virulence or pathogenicity mechanisms employed by A. psidii to cause disease and provides the first insight into the molecular responses of A. psidii in E. grandis over time.

68

LOCAL ANCESTRY ADJUSTED ALLELIC ASSOCIATION ANALYSIS ROBUSTLY CAPTURES TUBERCULOSIS SUSCEPTIBILITY LOCI

Miss Yolandi Swart[1], Dr Caitlin Uren[1,2], Prof Paul van Helden[1], Professor Eileen Hoal[1], Professor Marlo Möller[1,2]

[1]DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research, South African Medical Research Council Centre for Tuberculosis Research, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa, [2]Centre for Bioinformatics and Computational Biology, Stellenbosch University,, Cape Town , South Africa

Pulmonary tuberculosis (TB), caused by Mycobacterium tuberculosis, is a complex disease. The risk of developing active TB is in part determined by host genetic factors. Most genetic studies investigating TB susceptibility fail to replicate association signals, particularly across diverse populations. South African populations arose due to multi-wave genetic admixture from the indigenous KhoeSan, Bantu-speaking Africans, Europeans, Southeast Asian and East Asian populations. This has led to complex genetic admixture with heterogeneous patterns of linkage disequilibrium and associated traits. As a result, precise estimation of both global and local ancestry is required to prevent both false-positive and false-negative associations. Here, 820 individuals from South Africa were genotyped on the SNP-dense Illumina Multi-Ethnic Genotyping Array (~1.7M SNPs) followed by local and global ancestry inference using RFMix. Local ancestry adjusted allelic association (LAAA) models were utilized owing to the extensive genetic heterogeneity present in this population. Hence, an interaction term, comprising the identification of the minor allele that corresponds to the ancestry present at the specific locus under investigation, was included as a covariate. One SNP (rs28647531) located on chromosome 4q22 was significantly associated with TB susceptibility and displayed a SNP minor allelic effect (G allele, frequency =0.204) whilst correcting for local ancestry for Bantu-speaking African ancestry (p-value=5.518e-07; OR = 3.065; SE=0.224). Although no other variants passed the significant threshold, clear differences were observed between the lead variants identified for each ancestry. Furthermore, the LAAA model robustly captured the source of association signals in multi-way admixed individuals from South Africa.

10

IMPACT OF LIFESTYLE ON CYTOCHROME P450 MONOOXYGENASE EVOLUTION IS EVIDENT IN BACTERIA

Prof Khajamohiddin Syed[1]
[1]*University of Zululand, Kwadlangezwa, South Africa*

Adaptation is key for the survival of an organism. Organisms adapt to different ecological niches by changing their gene pool and thus changing their physiology to make them suitable for survival in the new environment. P450s are heme-thiolate proteins found in all domains of life and are known for their catalytic versatility and stereo- and regio-specific activity. While the impact of lifestyle or ecological niches on P450 evolution was reported in many eukaryotes, this remains to be addressed in bacteria. This study is aimed to address this research gap. Genome-wide data mining, annotation, and phylogenetic analysis of P450s in different bacterial groups such as Gammaproteobacteria, Firmicutes, Actinobacteria (Streptomyces and Mycobacterium), Cyanobacteria and Alphaproteobacteria revealed lifestyle or ecological niche of an organism certainly impacted P450 profiles in bacterial species where pathogenic lifestyle influenced P450 content to such an extent that species belonging this category lost P450s. On the contrary, the saprophytic lifestyle resulted in the presence of more and highly diverse P450s in bacterial species. It has been observed that species populated specific P450s are helpful in their adaptation to particular ecological niches or valuable in their physiology. This was evident in species known to produce secondary metabolites that help them survive competitive environments or evade the host immune system. Overall, study results strongly support the earlier hypothesis put forward by my laboratory that the impact of lifestyle shapes P450 content in an organism.

77

INFERRING R2R3-MYB FAMILY TRANSCRIPTION FACTOR GENE TARGETS IN EUCALYPTUS GRANDIS USING DNA AFFINITY PURIFICATION SEQUENCING AND MACHINE LEARNING

Mr Lazarus Takawira[1], Prof. Steven Hussey[1], Dr Fabien Mounet[2], Ms Jade Thulloch[1], Dr. Raphael Ployet[1], Prof. Eshchar Mizrachi[1], Prof. A.A Myburg[1]
[1]University Of Pretoria, Pretoria, South Africa, [2]Université de Toulouse III, Toulouse, France

Understanding the transcriptional regulation of secondary cell wall (SCW) development during wood formation in plants may enable strategic re-engineering of woody biomass traits. However, the transcriptional control of wood formation in non-model species such as Eucalyptus is poorly understood. We aimed to reconstruct a transcriptional regulatory network involving Eucalyptus R2R3-MYB transcription factors (TFs) EgrMYB1, EgrMYB2, EgrMYB135, and EgrMYB122, which are thought to regulate SCW biogenesis. We implemented DNA Affinity Purification-sequencing to identify genome-wide in vitro transcription factor binding sites (TFBSs). DNA bound in vitro by HALO-TF fusion proteins, with technical replication, was sequenced (Illumina NovaSeq6000), mapped to the E. grandis genome, and peaks representing TFBSs called. Library preparation and the number of PCR cycles had a significant impact on total and deduplicated mapped reads yield. On average, over 1 million deduplicated mapped reads were generated per binding assay and up to 9,942 peaks (TFBSs) could be reproduced across two independent laboratories for EgrMYB2. Next, we developed a machine learning classifier trained on published Arabidopsis DAP-seq data of 14 TFs to improve gene target assignment, incorporating 11 non-redundant features extracted from co-expression datasets, DAP-seq peak data, conserved non-coding sequences, and DNase I hypersensitive sites. A random forest classifier performed significantly better than the nearest-gene baseline model (precision = 0.751 vs 0.101; recall = 0.711 vs 0.243) and retained reasonable performance across independent datasets (AUC-ROC ≈0.72). The predicted R2R3-MYB gene targets showed enrichment for SCW related biological processes especially lignin biosynthesis, suggesting the successful reconstruction of a MYB-related SCW transcription factor network.

66

ROLE OF VARIATIONS IN MODERN DRUG RESEARCH AND DEVELOPMENT – WAY FORWARD VIA DYNAMIC RESIDUE NETWORKS

Prof Ozlem Tastan Bishop[1]

[1]*Research Unit in Bioinformatics (RUBi), Department of Biochemistry and Microbiology, Rhodes University, Makhanda, South Africa*

The genome sequences of humans and pathogens, and the variations attached to these data, can offer fundamental insights into disease prevention and treatment. Decoding the effects of missense mutations on protein structure and function is crucial to the understanding of the underlying causes of inherited diseases; drug toxicity in particular populations; and mechanisms of drug resistance, among others. While an enormous amount of genome data has been generated, the transition to post genomic analysis has been slow, thus widening the gap between data generation and translational utilization. For example, the human genome consists of about 20,000 protein coding genes and over 900 million variants. Systematic knowledge of the impact of genomic alterations in human is critical for the development of effective medicines. However, it is simply not feasible to study each and every one of these variants in detail. To date the effect of variations at the protein level is poorly studied in computational drug discovery research. This is mainly due to experimental difficulties in analyzing data at the protein structural level. We previously proposed a post-hoc analysis approach of molecular dynamics simulations using dynamic residue network analysis to consider the dynamic nature of functional proteins and protein-drug complexes and to probe the impact of mutations and their allosteric effects. This talk focuses on examples from pharmacogenomics to drug resistance in infectious diseases; proposing computational approaches for common ways of dealing with different health problems.

63

TRANSCRIPTIONAL RE-PROGRAMMING DURING RECOVERY FROM SHORT TERM DROUGHT STRESS IN EUCALYPTUS GRANDIS

Mr Demissew Teshome[1], Prof Godfrey Zharare, Dr Raphael Ployet, Prof Sanushka Naidoo
[1]*University Of Pretoria, Pretoria, South Africa*

The importance of drought as a constraint to agriculture and forestry is increasing with climate change. Genetic improvement of plants' resilience is one of the mitigation strategies to curb this threat under current and future climate. Although recovery from drought stress has been considered as an indicator of drought tolerance in annual crops, this has not been well-explored in forest trees. Thus, we aimed to investigate the physiological and transcriptional changes during drought stress and re-watering in E. grandis. We set up a greenhouse pot experiment where we imposed drought on two-year-old seedlings and re-watered the recovery group after 17 days of drought. Our measurement showed that, while stomatal conductance was reduced by drought stress, it fully recovered at five days of re-watering. Gene ontology enrichment and co-expression network analyses on our RNA-seq data revealed that known stress responses such as phytohormone and reactive oxygen species signaling were up regulated while metabolism and growth were down regulated due to drought. During recovery, we observed a reversal of many of the changes under drought stress indicating an ability of E. grandis to recover quickly. We also observed unique changes during recovery such as down regulation of biotic stress responses. Taken together, this study gives an overview of the molecular changes underlying the mechanisms of response to the repeated cycles of drought stress and recovery that trees face throughout their long-life cycle. This provides a useful reference to the identification of signalling pathways and target genes for future tree improvement.

16

GENOME COMPARISON IN ABALONE SPECIES TO PROVIDES INSIGHTS INTO GROWTH RELATED GENES

Mr Thendo Stanley Tshilate[1], Dr Edson Ishengoma[2], Dr Clint Rhode[1]
[1]Department of Genetics, Stellenbosch University, Stellenbosch, South Africa, [2]Mkwawa University College of Education, University of Dar es Salaam, Iringa, Tanzania

The South African abalone (Haliotis midae) is a valuable molluscan species, with growth rate as the major trait associated with maximum profit and productivity in aquaculture. Growth-related genes have likely undergone structural and functional changes throughout abalone development, leading to significnat size discreptancies between species. In order to draw inferences regarding growth-related genes, we concentrated on finding orthologous groupings across Haliotids.The assembled draft H. midae genome had a total length of 1.45 Gb, a N50 of 31320, and a GC level of 40%. Gene annotation yielded 52280 protein-coding genes and these were used for a comparative genomic analysis with four other abalone species. Gene-based phylogenetic analysis showed that H. rubra and H. laevigata are the most closely related to H. midae. From a total of 19968 orthologous gene-clusters identified among five abalone species, 30.8% (6156) of orthologous genes were shared by the five species (H. midae, H. laevigata, H. rubra, H. discus hannai, and H. rufescens) and 1352 were single-copy genes. Single copy genes were further anlysed for signatures of selection and a considerable number of molecular regulatory proteins involved in growth development were found to be under positive selection accross species. Our study assists in understanding genes related to various biological systems underscoring the evolution and developemet of growth in abalone, with future applications for genetic improvement of commercial stocks.

31

THE NORTHERN CAPE TUBERCULOSIS CASE-CONTROL CONSORTIUM (NCTBC3)

Dr Caitlin Uren[1,2], Mr Justin Myrick[3], Dr Austin Reynolds[4], Mr Oshiomah Oyageshio[5], Dr Cedric Werely[1], Dr Sihaam Boolay[1], Ms Jamie Saayman[1], Prof Brenna M Henn[3,5], Prof Marlo Möller[1,2]

[1]DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research, South African Medical Research Council Centre for Tuberculosis Research, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa, [2]Centre for Bioinformatics and Computational Biology, Stellenbosch University, Stellenbosch, South Africa, [3]Department of Anthropology and UC Davis Genome Center, University of California, Davis, Unites States of America, [4]Department of Anthropology, Baylor University, Waco, United States of America, [5]Center for Population Biology, University of California, Davis, United States of America

Tuberculosis (TB) remains a global health problem. In South Africa, it is the leading cause of death due to a single infectious agent. Genetic variants in a variety of genes have been linked to genetic susceptibility to TB. In addition, the role of ancestry has been confirmed. The Northern Cape province has one of the highest TB incidence rates in the country (943/100 000 individuals). In 2016, based off of prior research recruitment in the area (started in 2009), we began recruiting TB cases and controls in the ZF Mgcawu district. To date, we have collected ~ 2000 DNA samples with matched phenotypic data ranging from TB status and a detailed medical history to demographic details. Here we discuss the initial results from the project which include general cohort characteristics (TB and HIV incidence, gender and age characteristics, global and local ancestry estimations) and genotype-phenotype associations. In addition, we report back on community engagement initiatives and the involvement in other ethical, legal, and social aspects of genomic research in the field. Other than furthering the knowledge on genetic susceptibility to TB, we hope to have a positive impact on the local community. Furthermore, we hope to provide a TB case-control cohort to support future research endeavours.

28

CRITICAL ANALYSES OF THE LITERATURE: IN ISOVALERYL ACIDEMIA, IS ISOVALERYLGLYCINE A PRODUCT FORMED BY GLYCINE-N ACYLTRANSFERASE?

Prof Rencia van der Sluis[1]

[1]*Focus Area for Human Metabolomics, North- -West University, Potchefstroom, South Africa*

Isovaleryl acidemia (IVA) is a rare defect of leucine catabolism caused by mutations in the isovaleryl-CoA dehydrogenase (IVD) gene. This results in the accumulation of isovaleryl-CoA and isovaleric acid, which is toxic to mitochondria. There exists a general assumption in the literature that glycine N-acyltransferase (GLYAT) plays a role in alleviating the symptoms experienced by IVA patients through the formation of isovalerylglycine, a less toxic metabolite. GLYAT forms part of the phase II glycine conjugation pathway in the liver, and detoxifies endogenous and exogenous acyl-CoA. However, after thorough scrutiny of the literature, very few experimental studies support GLYAT as the enzyme that conjugates isovaleryl-CoA to glycine. Two paralogous of GLYAT are located on chromosome 11, namely GLYATL1 and GLYATL2. GLYATL1 is expressed in the liver and GLYATL2 in the gall bladder. Based on functional studies, GLYATL1 is responsible for glutamine conjugation reactions while the biological role of GLYATL2 is unknown. Therefore, GLYATL1 might also be a candidate for the conjugation of isovaleryl-CoA to glycine. It is important to verify whether the assumption that GLYAT can conjugate isovaleryl-CoA to glycine is indeed correct as glycine supplementation is used as a treatment for IVA patients. Unexplained interindividual variation in responsiveness to this treatment has been observed even in a SA cohort, all homozygous for the same IVD mutation. A thorough understanding of the role of these enzymes in detoxification can further inform dietary decisions for IVA patients as the consumption of benzoate (a preservative) might outcompete isovaleryl-CoA as a substrate for GLYAT.

84

FUSARIUM SOIL SURVEY FROM THE KALAHARI

Ms Shannon Bennet[1], Dr Brett Summerell[2], Dr Adriaana Jacobs[3], Prof Eduard Venter[1]
[1]Department of Botany and Plant Biotechnology, University of Johannesburg, Auckland Park, South Africa, [2]Australian Institute of Botanical Science, Royal Botanic Garden and Domain Trust, Sydney, Australia, [3]Plant Health and Protection, Agricultural Research Council, Pretoria, South Africa

South Africa is a mega-diverse country with several efforts to protect our diversity. However, this is focused mainly on the macrofauna and flora and not on the soil that supports them. Soil microbiome studies focus mainly on the bacterial content with poor documentation of soilborne fungal distribution. Better understanding of soil health would include all supporting parameters, as soilborne fungal phytopathogens cause great commercial agricultural losses. Recent studies have indicated that the distribution of toxin producing Fusarium species occur naturally and throughout the grassland biome of South Africa near cattle and grain farming. This highlights a potential problem with the ever-increasing urban encroachment into the grassland biome. Occurrence of fusaria within these soils may lead to future plant disease and crop losses, negatively affecting the health of small farming communities. To further our understanding of the distribution of fusaria in natural ecosystems, the project focuses on soil sampled from the savanna biome at a transition area into the Kalahari. From the collected Fusarium species, DNA was extracted and the translation elongation factor 1-alpha gene was amplified and sequenced. A phylogenetic analysis was performed on these isolates and the phylogenetic groupings were further investigated to group them into the different species complexes that exist within the genus Fusarium. Future extensive sampling throughout this biome from ecosystems with limited anthropological disturbance will benefit understanding of Fusarium distribution in South Africa.

85

PHYLOGENETIC SPECIES DIVERSITY IN FUSARIUM SPP. FROM THE GOLDEN GATE HIGHLANDS NATIONAL PARK, SOUTH AFRICA

Ms Carla Steyn[1], Dr Brett Summerell[2], Prof Eduard Venter[1], Dr Adriaana Jacobs[3]
[1]Department of Botany and Plant Biotechnology, University Of Johannesburg, Johannesburg, South Africa, [2]Australian Institute of Botanical Science, Royal Botanic Garden and Domain Trust, Sydney, Australia, [3]Plant Health and Protection, Agricultural Research Council, Pretoria, South Africa

The genus Fusarium hosts members capable of infecting a wide variety of economically important crops and can cause disease in both animals and humans. Although this group has been extensively studied, the diversity found in areas with less anthropogenic disturbance is largely unknown, especially in South Africa. This study aimed to explore the diversity found in the historically farmed nature reserve, the Golden Gate Highlands National Park, which falls within the grassland biome of South Africa. From the total 257 isolates obtained, six Fusarium species complexes could be identified as based on the Translation Elongation factor 1-α (TEF 1-α) gene region. The sequences were compared with those of two online databases, the FUSARIUM-ID database and the Mycobank database. Phylogenetic relationships and morphological characteristics were used to delineate species within the genus Fusarium. Most of the isolates (77 %) belonged to the highly diverse and widely pathogenic Fusarium oxysporum species complex (FOSC). This study contributes to the known species diversity of the FOSC group as well as other species complexes from the genus Fusarium found in semi-disturbed soils of the grassland biome in South Africa.

86

DIVERSITY WITHIN THE FUSARIUM SAMBUCINUM SPECIES COMPLEX WITHIN SOUTH AFRICA

Mr Marvin Khoza[1], Dr Rian Pierneef[2], Dr Brett Summerell[3], Dr Adriaana Jacobs[4], Prof Eduard Venter[1]
[1]Department of Botany and Plant Biotechnology, University Of Johannesburg, Johannesburg, South Africa, [2]Biotechnology Platform, Agricultural Research Council, Pretoria, South Africa, [3]Australian Institute of Botanical Science, Royal Botanic Garden and Domain Trust, Sydney, Australia, [4]Plant Health and Protection, Agricultural Research Council, Pretoria, South Africa

The genus Fusarium harbours some of the most economically restrictive phytopathogens. A Fusarium sp. residing in the F. sambucinum species complex (FSamSC) has been identified during the South African Fusarium grassland biome survey. This species is closely related to F. brachygibbosum that is a developing phytopathogen with increasing reports of infections in fruits, nuts and vegetables. This Fusarium sp. was demarcated using a phylogenetic analysis of a subset of the FSamSC. The analysis resulted in a high number of species within the grassland isolates that have nBLAST similarities to F. brachygibbosum. The phylogenetic analysis based on the tef-1α gene region indicated that there are over 30 undescribed species located in the analysis. Sequencing the genome of this species will allow comparison to related sequenced fusaria genomes within the FSamSC. This will allow assessment of its potential to develop as a phytopathogen.

81

MULTI-OMICS APPROACHES TO STUDYING HOST-PATHOGEN INTERACTIONS BETWEEN THE FOREST PEST, SIREX NOCTILIO AND ITS BIOCONTROL AGENT, DELADENUS SIRICIDICOLA

Ms Alisa Postma[1], Ms Sumari Venter[1], Ms Zorada Swart[1], Ms Sophie Makua[1], Prof Bernard Slippers[1]
[1]*FABI, University Of Pretoria, Pretoria, South Africa*

Sirex noctilio is an economically important invasive insect pest of pine trees. The nematode, Deladenus siricidicola, is the main biocontrol agent used against this woodwasp in the Southern Hemisphere. The efficacy of the nematode to sterilise the woodwasp is variable, and loss of virulence due to continued laboratory rearing might contribute to this. In this study, genome and transcriptome data of S. noctilio and D. siricidicola were used to investigate the key components underlying the interactions between these organisms on a molecular level. Genes putatively involved in nematode virulence towards insects were identified from the D. siricidicola genome and expression of these genes were studied in order to understand the transcriptomic differences between wild and laboratory reared strains. Furthermore, immune-related genes were identified in the genome of S. noctilio and the transcriptomes of larvae inoculated with D. siricidicola were sequenced and analysed in order to study the immune response of the insect when challenged with nematode infection. These data will enable a better understanding of the effect of the continuous culturing of D. siricidicola and its efficacy in sterilising S. noctilio. This study lays the foundation for understanding the intricate relationship of this parasite and host at a molecular level. These results could ultimately aid in the selection of more virulent biocontrol agents.

39

PHYLOGENETIC RELATEDNESS OF GEOGRAPHICALLY DISTINCT CHICKPEA-ASSOCIATED MESORHIZOBIUM SPP.

Mr Edwin Wanjofu[1,2], Professor Stephanus N Venter[1,2], Dr. Esther Muema[1,2], Professor Emma Steenkamp[1,2], Doctor Chrizelle Beukes[1,2]

[1]Forestry And Agricultural Biotechnology Institute, Pretoria, South Africa, [2]University of Pretoria, Pretoria, South Africa

Chickpea (Cicer arietinum) is non-native to South Africa; hence its cultivation depends on the incorporation of Mesorhizobium inoculants. This host associates with specific Mesorhizobium species affiliated with M. ciceri, M. mediterraneum and M. muleiense. The aim of this study was to infer the phylogenetic relationship of chickpea associated Mesorhizobium strains from different geographical locations across the world. Twelve strains were collected from various culture collections and collaborators. After resuscitation of the cultures and DNA extraction, sequencing of the housekeeping genes atpD, glnII, recA, dnaK and rpoB and symbiotic genes nodA, nodC and nifH were completed. Datasets included all Mesorhizobium type strains as obtained from GenBank(NCBI), which were manually aligned based upon the inferred amino acid sequence using BioEdit. Single gene phylogenies were constructed using MegaX with 1000 pseudoreplicates. Upon congruence of the single genes phylogenies, a multilocus sequence analysis (MLSA) was performed on the concatenated data constructed using FASconCAT-G and subjected to maximum-likelihood phylogenetic analysis using RAxML incorporating 1000 pseudoreplicates. The results revealed four distinct clusters. Seven strains were nested within the M. ciceri lineage. One isolate grouped with M. mediterraneum, while four strains formed two distinct putatively novel lineages. In the symbiotic phylogenies, seven strains clustered with M. ciceri, three with M. mediterraneum and two with M. muleiense and M. wenxiniae. These results revealed additional Mesorhizobium species different from those previously affiliated with chickpea, can interact with this host, however their symbiotic genes are affiliated with M. ciceri, M. mediterraneum, M. wenxiniae or M. muleiense.

48

PANGENOME OF AFRICAN POPULATIONS OF THE MAIZE FUNGAL PATHOGEN CERCOSPORA ZEINA

Ms Tanya Welgomoed[1,2], Dr Tuan A. Duong[1,2], Prof Irene Barnes[1,2], Prof Eva H. Stukenbrock[4,5], Prof Dave K. Berger[1,3]

[1]Forestry & Agricultural Biotechnology Institute, University of Pretoria, Pretoria, South Africa, [2]Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa, [3]Department of Plant and Soil Sciences, University of Pretoria, Pretoria, South Africa, [4]Environmental Genomics, Christian-Albrechts University of Kiel, Kiel, Germany, [5]Max Planck Institute for Evolutionary Biology, Plön, Germany

One of the more serious factors hampering global maize production is grey leaf spot disease. Cercospora zeina is one of the causative pathogens. While some progress has been made in understanding the resistance mechanisms of the maize host, the pathogenicity mechanisms of C. zeina has not been studied to the same extent. Understanding the evolution and virulence mechanisms of this pathogen is therefore vital to prevent and predict breakdowns of host resistance. To aid this, we conducted whole-genome Illumina sequencing of a diverse set of 30 isolates representing African populations of C. zeina. GATK was used to call SNPs in the haploid genomes. DAPC and ADMIXTURE analysis of the SNPs showed population structure with a geographical subdivision between the Eastern African isolates and the Central and Southern African isolates. The C. zeina pan-genome was constructed by de novo assembling the sequencing data using SPADES, annotating the assemblies with BRAKER, and clustering using ORTHOFINDER to group the predicted genes into 10 630 shared core genes and 836 non-shared accessory genes. Based on predicted gene functions we found an enrichment of accessory genes involved in transcriptional regulation and protein signalling and we speculate that these genes have a role in local adapation of C. zeina. Generally, the pan-genome structure of a species relates to its ecology and the accessory genes of fungal pathogens often contain important pathogenicity genes. The variations captured by the SNP and pangenome analyses therefore provide targets for further investigation into the evolution and virulence of C. zeina.

73

LONG READ DE NOVO ASSEMBLY OF THE ACACIA MEARNSII GENOME

Mr Mondli Xaba[1,5], Dr Julian Chan[2], Mr Craig Norris[3], Prof Fourie Joubert[1], Prof Jill Wegzryn[4], Dr Tuan Duong[1,5], Prof Alexander Myburg[1,5]

[1]Department of Biochemistry, Genetics and Microbiology, University Of Pretoria, Pretoria, South Africa, [2]Institute for Commercial Forestry Research , Pietermaritzburg , South Africa, [3]NCT forestry Agricultural Cooperative , Pietermaritzburg , South Africa, [4]Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, United States of America, [5]Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa

Acacia mearnsii (black wattle) is an important plantation forestry species for small growers in South Africa occupying ~7% of the forestry estate. To breed for improved woody biomass production, it is important to understand the genetic make-up of local black wattle breeding populations. A comprehensive understanding of the genome diversity of black wattle is lacking and the availability of DNA markers for molecular breeding is limited to a small number of microsatellite DNA markers. The need for molecular breeding tools in black wattle is further emphasized by an ongoing wattle rust epidemic that can cause major damage to young plantations. To address this need, we have used long read (Oxford Nanopore PromethION) DNA sequencing technology to sequence the genome of a locally selected black wattle genotype. A total of 160 Gbp of nanopore sequencing data, representing 230X coverage of the A. mearnsii genome, was obtained. We have produced a preliminary genome assembly comprising of 469 contigs with an N50 value of 2,2 Mbp, which accounted for 670 Mbp of the estimated 680 Mbp genome size for the species. Analysis of universal-single copy orthologs (BUSCO) genes suggests that the current assembly is 92% complete. We have also initiated in-silico mining for microsatellite DNA sequences suitable for the development of markers for routine DNA fingerprinting and parentage analysis. The genome sequence will also serve as a reference for the development of genome-wide genotyping and molecular breeding resources.

SPONSORS

The South African Genetics Society and the South African Society for Bioinformatics would hereby like to acknowledge and give thanks to our sponsors for their support of the conference:

DIPLOMICS
CONNECTING RESEARCHERS & WORLD CLASS OMICS FACILITIES

NextGen
Molecular Supplies

inqaba biotec™
Africa's Genomics Company

DIAGNOSTECH

Glass Chem

CELTIC
MOLECULAR DIAGNOSTICS

SEPARATIONS
Simply Spectacular

SACNASP
South African Council for Natural Scientific Professions
Integrity in science

Ascendis
MEDICAL

DIAGNOSTECH

Tel: +27 (0)11 469 0097
Email: info@diagnostech.co.za

www.diagnostech.co.za

Committed to being your preferred genomics, pathology and diagnostics partner

Our extensive portfolio includes, but is by no means limited to:

- Immunohistochemistry (IHC)
- Companion Diagnostics (pharmDx)
- Molecular Pathology
- Flow Cytometry
- H&E and Special Stains
- Histology Controls
- Microarray Technology
- Next Generation Sequencing (NGS)
- PCR, RT-PCR and qPCR
- Mutagenesis
- Cloning Vectors and Reagents
- Enzymes
- DNA/RNA Purification
- Protein Expression
- Competent Cells
- Synthetic Biology

Visit our website to download our full brochure and view more on our available products

Agilent Technologies
Authorised distributor

diagenode

HistoCyte Laboratories

LGC

ARCHER®

ACKNOWLEDGEMENTS

The South African Genetics Society and the South African Society for Bioinformatics hereby extend our gratitude to:

- All of our plenary and keynote speakers;
- The joint conference organising committees;
- The joint scientific committee and all of our abstract reviewers and panel of assessors;
- CMC - Connect.Meet.Collaborate;
- The Stellenbosch Institute for Advanced Study (STAIS) at Stellenbosch University; and
- All of our presenters and other delegates.