December 2020

Volume 2, issue 4

SASBi

Newsletter

Inside the Issue

South African Society for Bioinformatics

UNIVERSITEIT iYUNIVESITHI STELLENBOSCH UNIVERSITY 100 1918·2018
saam vorentoe · masiye phambili · forward together

PRESIDENT'S LETTER

Greetings!

I have now served as your president for 2 years, and Ruben Cloete, Oleg Reva, Alisa Postma, Phelelani Mpangase, Cedric Werely, Mahtaab Hayat, Tulio de Oliveira, Yuman Moosa and Werner Smidt have served as the other Council Members. Due to the COVID-19 pandemic, the Council's term has been extended till September 2021, when we hope to have our joint SAGS/SASBi BIO2021 Conference and elect a new

(Continued on page 2)

Prof Gerard Tromp

Photograph: Stefan Els

SASBi Student Council: Virtual Symposium

In August 2021, over the course of three days 163 attendees joined the SASBi Student Council (SC) online for our first ever virtual Student Symposium. On day one we had a total of 58 attendees, of which 88% had never attended a SASBi-SC Student Symposium before, highlighting the advantages of hosting the Student Symposium online.

Oral, project overview and group presentations presented by the young researchers at this symposium covered transdisciplinary research incorporating Bioinformatics. A big congratulations to our prize winners:

Kaitlyn Flynn - Best Oral Presentation

Minette Havenga - Second Best Oral Presentation

Nicole Dames and Ncité Da Camara - Joint Third Best Oral Presentation

Lauren Martin - Best Project Overview Presentation

Morne du Plessis - Second Best Project Overview Presentation

The presentation recordings are available on YouTube at the links below:

Day 1 - SASBi SC Virtual Symposium Day 1

Day 2 - SASBi SC Virtual Symposium Day 2

Day 3 - SASBi SC Virtual Symposium Day 3

Thank you to each and every attendee for contributing to making the SASBi-SC Online Student Symposium a great success!

We hope to see everyone in person for the 2021 SASBi-SC/SAGS Student Symposium on Sunday the 26th of September in Stellenbosch, South Africa.

(Presidents Letter from page 1)

president and new council members.

During 2020, SASBi Council 'convened' virtually eight times to plan and discuss SASBi activities. We have published four SASBi newsletters during 2020. All issues are available on our website at http://sasbi.weebly.com/.

We have continued to work hard to revitalize SASBi. For example, we have established SASBi as a group on the "Google Groups" platform to facilitate communication. To post on the group you can send an e-mail to *sasocbioinformatics@googlegroups.com*. Aside from announcements, the group can be used for discussions of bioinformatic topics. Note that the group is moderated to avoid abuse and reduce spam.

Another new SASBi initiative is to organize workshops on various bioinformatics topics. We have formed a curriculum working group tasked in designing virtual workshops on RNA-sequencing analysis and using R, which you, our members, voted as the topics for workshops you would like to attend. We hope to be able to run the first workshop in 2021.

In October 2020, we launched our virtual seminar series. You can read more about this on page 4 of this newsletter.

We are also in the process of having SASBi become a member of the South African Council for Natural Scientific Professions (SACNASP), which would allow our conferences and workshops to be used towards continuing education.

To improve our visibility as a Society, we have established SASBi Silver and Gold Awards. The Silver Award is meant for a more junior bioinformatician with 5-20 years of experience after a PhD degree and who has made major contributions to a minimum of two peer-reviewed publications and advanced the discipline of bioinformatics by teaching, advising, or influencing public advocacy. The Gold Award is meant for a more senior bioinformatician with over 20 years of experience after a PhD degree and who has made major contributions to several recent publications and advanced the discipline of bioinformatics by teaching, mentoring, advising, or influencing public advocacy. Both Silver and Golder awardees are also expected to have been active in our Society. More information about the award criteria is available on page 3. The call for nominations will be released soon and the first awardees to these awards will be announced during the BIO2021 Conference.

In 2019, SASBi and SAGS committed to working more closely together and have had reciprocal representation on each other's councils to promote collaboration and communication. One of the most important collaborations is planning for the next conference, which was

(Presidents Letter from page 2)

scheduled for September 2020, but had to be postponed till September 2021 (see page 5) due to the COVID-19 pandemic. We are still hoping that we can go ahead with the face-to-face conference on 26 – 29 September 2021, but we are monitoring the situation closely and will keep you posted.

We would welcome your input! We would especially appreciate nominations for Council and SASBi officers. If you have suggestions, please contact any of the Council or Student Council members, as appropriate.

Gerard Tromp

SASBi Awards

Council acted on a proposal to create two awards that are to be given in recognition of effort to promote bioinformatics in South Africa. As mentioned in the President's letter, the two awards, Gold and Silver, are to recognize senior (> 20 years) and junior (5 — 20 years post PhD) bioinformaticians for service to the discipline and to the Society.

Council passed a resolution that:
Candidates for prizes shall be nominated as follows:

- A SASBi member shall nominate a candidate for one of the two categories
- Nominations shall be accompanied by a motivation document that shall address the award criteria.
- The motivation shall:
    * emphasize the contribution of the candidate to the discipline and the Society
    * emphasize a few key publications by the candidate with rationale as to their influence
- Nominations shall include an extensive curriculum vitae of the candidate (not a résumé)

Nominations shall be made no later than two months before the biennial conference and general meeting.

Nominations shall be directed to the Council Secretary (currently Ruben Cloete, ruben@sanbi.ac.za), who shall forward the collected nominations to the duly constituted committee for evaluation.

Awards shall be presented at the biennial conference.

Simonsberg, 2020-12-25, Gerard Tromp

SASBi Virtual Seminar Series Launched

On October 20, 2020, SASBi launched its Virtual Seminar Series. The first speaker of the series was Prof. David Tabb of Stellenbosch University. He presented *"PCA-Powered Quality Control for Large-Scale Proteomics Experiments."* Dr. Tabb had previously served as part of the USA National Cancer Institute Clinical Proteomics Technology Assessment for Cancer (CPTAC) effort, publishing several papers relating to the reproducibility of proteomics technologies.

For proteomics, the pairing of quality metrics with PCA has already made several applications possible. First, PCA is useful for recognizing outliers and batch effects. If one or two experiments fall far from the others on a PCA plot (using a distance metric such as a Euclidean distance), it suggests that their metrics were not in sync with the main mass of the data. If many weeks' worth of data are visualized via PCA, it may be that experiments group only with others that were conducted at similar

In this talk, Dr. Tabb emphasized the need for dimensionality reduction algorithms such as Principal Components Analysis (PCA) to reduce large sets of quality metrics derived from proteomic experiments to relatively few components. The components are ordered by the amount of measurement variability that each explains among experiments. PCA functions as a type of "projection pursuit" for unsupervised learning; experiments are grouped by their similarity in component values, and this can lead to discovery of groupings that were previously unknown among the experiments.

number of times. If week one experiments and week two experiments are located separately in PCA components, a substantial batch effect may complicate comparing experiments between weeks. While Dr. Tabb spoke about these challenges in the context of proteomics, he noted that batch effects and other types of technical variability are quite common among different biotechnologies. All users of systems biology should be cognizant of technical variation in complex molecular biology measurements.

Stay tuned for the next SASBi virtual seminar planned for February 2021.

Save the Date – Bio2021 CONFERENCE

26-29 September 2021

Venue: Spier Wine Estate

Conference website: https://uctcmc.eventsair.com/south-african-genetics-society-the-southern-african-society-for-bioinformatics

Key dates:

- **Early Bird Registration**: 2 March – 23 July, 2021
- **Registration:** 24 July – 3 September, 2021
- **Abstract Submission**: 1 February – 21 May, 2021

Co-chairs: Clint Rhode (SAGS) & Gerard Tromp (SASBi)

South African Society for Bioinformatics

SASBi

SAGS

South African Genetics Society

Student Profiles @ Stellenbosch

Lovemore Nyasha Sigwadhi

PhD Candidate: Bioinformatics, Tygerberg Campus, Stellenbosch University

What is your current project?

I am designing a Bayesian clustering algorithm for mass cytometry data.

Please fill us in on your career up to this point?

I completed a BSc degree in Mathematics and Statistics at University of Zimbabwe in 2013. I then completed BSc Honours degree in Statistics (2014) and an MSc Master's degree in Operations Research and Statistics (2016) at the National University of Science and Technology. In 2018, I enrolled at Stellenbosch University for a MSc degree in Biostatistics which I completed in 2020. My background is in mathematics and statistics with a MSc degree focusing on assessment of performance measures using simulation in epidemiology and biostatistics. I proceeded to work as a Statistical consultant in the tertiary sector, part-time lecturer at a training college, then in data analyst in the medical industry. As a consultant, I provided statistical support to researcher in the medical field from study design, sample size calculation, data management, statistical analysis plan, data analysis and results interpretation. During my MSc degree I joined the Bioinformatics Group as an intern and got more interested in statistical application in immunological data. In 2020, I joined Prof Gian van der Spuy's team for my PhD.

If you could give advice to students who are new in the field, what would you say?

The best index to a person's character is how he treats people who cannot do him any good, and how he treats people who cannot fight back "Abigail Van Buren". Keep on working hard, the journey may be tough but you are not alone so don't lose hope. ***Rejoice always, pray without ceasing, give thanks in all circumstances; for this is God's will for you in Christ Jesus, 1 Thessalonians 5: 16-18***

"R is my favourite. I love the way it handles data wrangling and statistical analysis for complex methods."

What are your hobbies/activities you do in your free time?

Playing snooker, siteseeing, playing with kids, and meeting with new friends.

If you were an animal – what would you be and why?

A honeybee. It is dedicated and passioned about fulfilling its goal. It is the only insect that produces food eaten by man. A honeybee can fly for up to six miles, and as fast as 15 miles per hour, hence it would have to fly around 90,000 miles -three times around the globe – to make one pound of honey. It takes one ounce of honey to fuel a bee's flight around the world.

Do you have a favourite computer language?

R is my favourite. I love the way it handles data wrangling and statistical analysis for complex methods.

Benilo Slabbert

MSc Candidate: Bioinformatics, Center for Bioinformatics and Computational Biology, Stellenbosch University

What is your current project?

I'm currently investigating mechanisms whereby epigenetics might increase the efficiency of transcription, specifically in the context of facilitated diffusion. Facilitated diffusion is the observation that DNA binding proteins search for their cognate binding site not only by 3D diffusion, but also by sliding along the DNA. It is known that facili-

(Continued on page 7)

tated diffusion increases search efficiency, but the potential role of epigenetics has not been considered thoroughly. To test the above mentioned effects, I am building a model for simple explicit DNA binding protein dynamics.

Please fill us in on your career up to this point:

I completed my BSc in Molecular Biology and Biotechnology in 2018 majoring in Genetics and Biochemistry. The following year I received my BSc Hons in Biochemistry, during which I was privileged to attend a Behavioral Genetics Summer School hosted at VU Amsterdam. My Honours project involved using Molecular Dynamics Simulations to probe the effect of acetylated histone tails on nucleosome stability. At present I am concluding the first year of my MSc in Bioinformatics with the CBCB at Stellenbosch University.

If you could give advice to students who are new in the field, what would you say?

You're in the right place at the right time. There are so many opportunities for new, valuable research in Bioinformatics and Computational Biology. Figure out what excites you in the field and pursue it with diligence.

"There are so many opportunities for new, valuable research in Bioinformatics and Computational Biology. Figure out what excites you in the field and pursue it with diligence"

What are your hobbies/activities you do in your free time?

I enjoy playing football, reading contemporary philosophy and listening to science podcasts.

If you could change one thing about yourself, what would it be?

I'd really like to be a morning person, but alas.

Do you have a favourite computer language?

Python. Simple and powerful, although I would not die on this hill.

Your idol? (Can be someone in your life, a celebrity or a fictional character)

I'm not usually one for idols, but someone I truly admire is the theoretical physicist Sean M. Carroll at CalTech. Interested in almost everything, thoughtful on most topics, kind and honest even in his disagreements.

Jessica Reid

MSc Candidate: Bioinformatics, Center for Bioinformatics and Computational Biology, Stellenbosch University

What is your current project?

Research into the molecular mechanism that causes htz1Δ strains to resist evolutionary improvement of fitness.

Please fill us in on your career up to this point:

I obtained my BSc degree in Molecular Biology and Biotechnology in 2017 and then went on to complete my Honours in Biochemistry the following year at Stellenbosch University. I then took a year break from my studies in 2019. During this time, I was lucky enough to travel and work in China as an English teacher for working professionals. This year abroad really highlighted my passion for biochemistry and so I was thrilled to be able returned to Stellenbosch this year to start my MSc under the guidance of Prof. Patterton.

"Be patient with yourself. Learning is a process and a beautiful one at that. "

If you could give advice to students who are new in the field, what would you say?

Be patient with yourself. Learning is a process and a beautiful one at that. The field you've chosen is so expansive which is exciting but it can simultaneously be incredibly daunting. So be patient, be curious and never shy away from asking questions. You are

(Continued on page 8)

surrounded by so many brilliant minds at this University - so make the most your time with them.

What are your hobbies/activities you do in your free time?

I love a challenge, so it'll make sense when I say triathlon is what I spend a lot of my free time training for. When I'm not waking up early to make the most of the Stellenbosch routes, I really enjoy trying to find new fun and interesting Asian dishes to subject my friends to at the dinner table.

When did you first realize you love science/computers?

I've always been a creative person who enjoys a problem to ponder over. After really focusing on Biochemistry in third year I realized it was the perfect avenue for me to put my creativity to work in a way that was both fun and mentally stimulating.

Please describe what you were like at age 10.

Wow, so different. I was fascinated with creepy crawlies and so would carry my 'bug box' everywhere I went. Now I am probably the last person to call if there's a spider in the room.

Hannah Simba

PhD Candidate: Department of Global Health, Tygerberg Campus, Stellenbosch University

What is your current project?

My study is on the role of genetic and environmental factors associated with Oesophageal Cancer (OC) in the South African population. OC is one of the most aggressive cancers worldwide, with high mortality rates, particularly in developing countries. It is endemic in sub-Saharan Africa, and its etiology is not fully understood. In South Africa it is highly prevalent in the Eastern Cape Province. My study aims to assess and characterize the role of genetic and environmental factors in the development of OC, and investigate the underlying molecular pathobiology using gene expression. I have worked on two systematic reviews so far, focused on OC genetic risk factors reported in relevant Africa literature, doi.org/10.3389/fgene.2019.00642 , and OC environmental risk factors reported in relevant African literature (manuscript under preparation). I am currently working on OC GEO (Gene Expression Omnibus) datasets to identify biological pathways involved in OC development by assessing genes demonstrating altered mRNA expression using weighted gene co-expression network analysis (WGCNA) of. This will be the most comprehensive bioinformatic analysis of existing GEO mRNA expression datasets on OC and we expect it to provide novel insights into the pathobiology of OC. Finally, I will be assessing genetic data from the first African Genome Wide Association Study (GWAS) using a candidate gene approach by a)validating genetic variants identified in the systematic review using the GWAS data, b)assessing polymorphisms in genes from biological pathways that are associated with specific environmental risk factors (tobacco smoking, Smoke exposure and alcohol consumption) in the South African population. All this knowledge generated from my study will help us understand the genetic and environmental basis of OC in Africa, and ultimately play a role in our efforts to modulate risk and susceptibility to OC development and progression.

"Always try and find platforms and spaces to speak about your scientific work, outside the Ivory Tower of Academia. It is important that people know and understand what scientific research is and why it is important, and you may inspire someone in the process."

Please fill us in on your career up to this point:

My academic background includes a BSc in Biomedical Sciences for the University of Kwa -Zulu Natal, Durban, a BSc Honours in Chemical Pathology and a MSc in Environmental Health from the University of Pretoria. I am currently doing a PhD in Public Health. My

career aspirations are to continue in research and academia.

If you could give advice to students who are new in the field, what would you say?

Don't be afraid to ask questions when you don't know, and always stay updated on new literature.

If you could preach about something to other scientists, what would it be? (For example, "Please add your data to a repository once it's published" or "Please stop using the word "uhm" in presentations".)

Always try and find platforms and spaces to speak about your scientific work, outside the Ivory Tower of Academia. It is important that people know and understand what scientific research is and why it is important, and you may inspire someone in the process.

What are your hobbies/activities you do in your free time?

I play tennis, hike and do yoga.

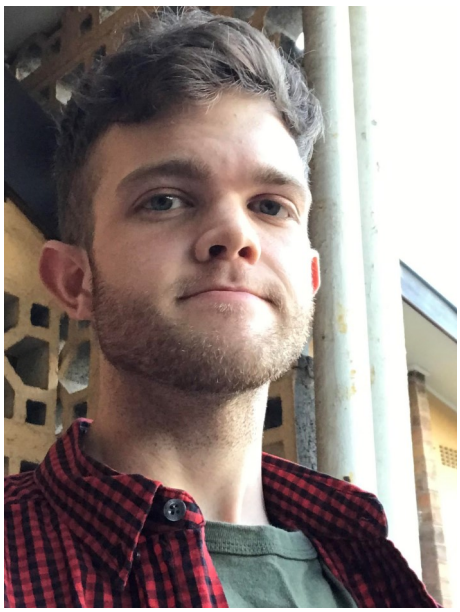
If you were an animal – what would you be and why?

I would be a bird. I have always found flying fascinating, and I would be able to just fly anywhere I want to go and explore. As to which specific bird, I would want to be one of the extant birds with largest wingspan, therefore it would have to be the Wandering Albatross.

When did you first realize you love science/computers?

I was always inquisitive and observant, and when I realized that science could answer most of the questions I had early in primary school, I fell in love with it. Science class was my favorite class.

Aidan Swartz



MSc Candidate: Bioinformatics, Tygerberg Campus, Stellenbosch University

What is your current project?

My current project is meant to assemble the transcriptome of the spotted hyena, while also including a proteomics component.

Fill us in on your career up to this point:

So, I began my studies doing a BSc in microbiology and human genetics at Stellenbosch University, main campus, for about three years, before I moved over to the Tygerberg branch of SU, in order to continue my studies in an Honours degree in Molecular Biology and Human Genetics, for one year. At current, I'm busy with my Master's degree, still in Molecular Biology and Human Genetics, working on a continuation of my project from my Honours degree.

If you could give advice to students who are new in the field, what would you say?

The best advice I'd ever received was actually by a professor in my third year undergraduate degree. She told us that she wanted us to prepare for failure. And, more importantly, to not be afraid of it. You're going to fail a lot in science. Experiments will go wrong. Sometimes it will feel like the world is conspiring against you. In the end, all you can do is push through.

If you were an animal – what would you be and why?

Probably a raven. Highly intelligent birds. They've run tests where the raven had to escape a cage and was capable of using simple tools to get out of the cage, based on a reward system, of course. They're also just really beautiful birds.

Describe what you were like at age 10?

Right. So my memories of the age 10 are… fuzzy. Not because of anything serious, but around that time I spent most of my time with my nose buried in a book. Including during class time. The rest of the time I was finding someplace to read.

If you could preach about something to other scientists, what would it be?

Quality assessment is something that should be carried out more extensively in science. In some fields it is more prevalent than others, but ultimately it is probably one of the most important aspects of your work. Especially for the initial data you receive, your "raw files", this should be extensive. If your foundation is shaky you want to know sooner rather than later, so that problems further down can be completely avoided or planned around.

"The best advice I'd ever received was actually by a professor in my third year undergraduate degree. She told us that she wanted us to prepare for failure. And, more importantly, to not be afraid of it. You're going to fail a lot in science. Experiments will go wrong. Sometimes it will feel like the world is conspiring against you. In the end, all you can do is push through."

Editorial Team: ***This issue brought to you by H Kuivaniemi & G Tromp, Stellenbosch University***