

SAGC 2022 Workshop Series: Workshop 1. Metagenomics

Flinders at Victoria Square, 182 Victoria Square, Adelaide, South Australia

July 4th-6th, 2022



Workshop Schedule

Note: Schedule subject to change depending on progress of the workshop!

Day 1: Monday July 4th

Time	Topic	Presenter
0900-0915	Welcome and Introductions. PATRIC account	RE
0900-1000	Introduction to computing	BNP
1000-1030	QC/QA hands-on	BNP
1030-1100	Morning coffee	
1100-1130	Introduction to metagenomics	MD
1130-1230	Kraken2 hands-on	MD
1230-1330	Lunch	
1330-1400	Functional analysis	SG/MD
1400-1430	Super-Focus hands-on	SG/MD
1430-1500	Sequence Assembly and Evaluation with QUAST	VM/BNP
1500-1530	Afternoon tea	
1530-1600	Metagenomic binning	VM
1600-1700	Binning hands-on	VM

Day 2: Tuesday July 5th

Time	Topic	Presenter
0900-0930	Recap of Day 1	EDi
0930-1000	Metagenome-assembled genome annotation - PATRIC	BNP/SG
1000-1030	KEGG mapper	BNP
1030-1100	Morning coffee	
1100-1200	AMR and WW	EDo
1200-1230	AMR finding	MR

1230-1330	Lunch	
1330-1400	Data visualization in R and Python	MR
1400-1500	Hands-on data visualisation	MR/VM
1500-1530	Afternoon tea	
1530-1645	Hands-on data visualisation	MR
1645-1700	Wrap up and summary	RE

Day 3: Wednesday July 6th

Time	Topic	Presenter
0900-0930	Recap of Day 1 and 2	EDi
0930-1000	Introduction to workflows (Snakemake)	NWH
1000-1030	Snakemake hands-on	NWH
1030-1100	Morning coffee	
1100-1200	Snakemake hands-on	NWH
1230-1330	Lunch	
1345-1530	TreeClimb Adelaide Waiver: https://roller.app/treeclimbadelaide/waivers/3119#/	
1400	Arrival	
1420	Session 1	
1440	Session 2	
1500	Session 3	
1530	Drinks at the Goodwood Park Hotel	SAGC

WELCOME TO THE WORKSHOP

The organisers of the 2022 SAGC Metagenomics Workshop:

- South Australian Genomics Centre (SAGC) (<https://sa-genomics.com.au>)
- Flinders Accelerator for Microbiome Exploration (<https://fame.flinders.edu.au>)
- The Flinders University of South Australia (<https://www.flinders.edu.au>)
- Robert Edwards and Elizabeth Dinsdale, Matthew Flinders Fellows, Flinders University of South Australia

extend a warm welcome to speakers and participants to Flinders, CBD and to the workshop, and trust that you will have an enjoyable and interesting experience.

We look forward to seeing you at the Workshop!

ABOUT THE WORKSHOP

This workshop is designed as a hands-on introduction to metagenomics sequence analysis. The workshop is designed with end-users in mind – people that want to leverage the power of next generation sequencing but are unsure exactly where to start. In the workshop we will walk through the steps of DNA sequence analysis, from quality control and quality assurance, through DNA assembly, to gene and protein identification and annotation. We will also analyse metagenomics DNA sequences, with hands-on experience analysing random community metagenomics. Students in the course will learn the latest cutting-edge techniques for sequence analysis including extracting complete genomes from metagenomes and learning how to assess those genomes.

The workshop is taught through a combination of lectures highlighting and discussing the key topics, followed by hands-on practical experience analysing DNA. No prior experience with programming is required. The course will include an introduction to using the terminal in Linux.

At the end of the workshop students should be able to:

1. Identify the appropriate software for the analysis of their metagenomics sequencing project.
2. Starting from a simple DNA sequence file, be able to analyse the sequences and generate meaningful conclusions from the DNA sequences
3. Analyse random community metagenomic sequences to identify which organisms are present in the sample and what they are doing.

Professor Robert Edwards (course leader) has been at the forefront of microbial genomics and metagenomics for more than 20 years. He has developed software to analyse and annotate bacterial genomes and metagenomes. Using those tools, he discovered a novel virus that is present in the microbiomes of about half of the people in the world.

Professor Elizabeth Dinsdale has been studying microbial genomes and metagenomes associated with tropical and temperate marine environments for over 20 years. She is an expert in microbial isolation and culturing, DNA sequencing, sequence analysis, and connecting genotype and phenotype.

Invited Keynote Speaker

Erica Donner

Professor Erica Donner is an environmental scientist with expertise in chemical and microbiological risk assessment and management. She specialises in systems-based contaminants analysis, focussing on the water cycle, circular economy, and food production systems. Erica is intent on understanding the environmental dimensions of antimicrobial resistance (AMR) in these contexts and how we can best harness this knowledge to deliver positive benefits for societies. Her research, which is grounded in the concepts of One Health and One Water, reveals the close interconnectedness of human, animal, and environmental health.

To address the complex challenge of AMR Erica engages in industry-partnered research to support informed discussion and the co-development of sustainable solutions. In recent years, she has led an SA government-funded international collaboration project on AMR in wastewater treatment plants and downstream environments. She was also a core participant in European COST Action ES1403: 'New and Emerging Challenges and Opportunities in Wastewater Reuse' (NEREUS; www.nereus-cost.eu), a four-year collaborative network program bringing together scientists from around the world to share insights, conduct joint research, and mentor early career researchers.

Erica is a member of the Australian Strategic and Technical Advisory Group on Antimicrobial Resistance (ASTAG) and is a Steering Committee member for the Victorian Antimicrobial Resistance Strategy and the South Australian AMR Action Plan. She is also co-lead of the Water, Soil, and Food Security Theme in the NHMRC-funded national HEAL network (www.heal2021.com.au/heal-network). This newly established democratic research network, which has strong Aboriginal and Torres Strait Islander leadership, is co-creating regional hubs and communities of practice dedicated to advancing One Health solutions across Australia.

Building on her interest and expertise in AMR, Erica has worked with industry collaborators in the water, organic waste, and agricultural sectors to develop the new Cooperative Research Centre (CRC) for Solving Antimicrobial resistance in Agribusiness, Food and Environments (CRC SAAFE), a cross-sector CRC designed to help translate Australia's National AMR Strategy into industry best practice. The CRC consortium provides the ideal forum to co-create knowledge with industry experts and researchers through a fundamentally transdisciplinary approach.

Instructors

Sen Wang. South Australian Genomics Centre (SAGC) Manager

Sen received his joint PhD degree from Shanghai Medical school and University of Basel and has more than ten years' experience using NGS technologies with an in-depth understanding of its applications in the research and clinical fields. Before this role, Sen was running the NATA accredited Genomics facility back at Monash Health Translational Precinct over the past 5 years. Working together with the genomics specialists and the Bioinformatics platform at the SAGC, Sen is endeavoured to deliver the state-of-the-art, end-to-end genomics services to the research and clinical communities.

Mike Doane

Michael Doane recently started as a Research Associate at Flinders University, working in the FAME group with Elizabeth Dinsdale and Rob Edwards. He is originally from the United States, where he did his Ph.D. work at San Diego State University. His Ph.D. work focused on using ecological principles to define processes that shape the microbiome of sharks. He then went on to work as a research associate at the Sydney Institute of Marine Science, interrogating the IMOS microbial dataset to identify driving principles shaping marine water column microbial community patterns. As a research associate at Flinders University,

Michael will continue to investigate the role of microbial communities with host organisms, principally elasmobranchs such as sharks, and explore means of expanding microbiome research into applied settings, such as the aquaculture industry.

Vijini Mallawaarachchi

Vijini Mallawaarachchi is a Research Associate in Bioinformatics at the Flinders Accelerator for Microbiome Exploration (FAME). She completed a Ph.D. in Computer Science at the Australian National University in 2022, where she focused on developing models and algorithms for improved binning of metagenomic data. She has developed several metagenomic binning tools including GraphBin and MetaCoAG which have leveraged assembly graphs to accurately bin bacterial metagenomes. Currently, her research at FAME focusses on developing methods to bin viral metagenomes which can facilitate the discovery of novel bacteriophages from metagenomic data.

Michael Roach

Michael Roach is a Bioinformatician at the Flinders Accelerator for Microbiome Exploration. His PhD was in drug discovery and protein structural bioinformatics, and he transitioned to genomics when starting his first post-doc at the Australian Wine Research Institute. During this time, he pioneered grapevine clonal marker discovery with his work on the Chardonnay genome and built the tool that is the current standard for assembly clean-up of heterozygous genomes. More recently, Mike has led the development of the Hecatomb pipeline for viral metagenomics and is leading the effort to build a pan-virome across the various human microbiomes.

Susie Grigson

Susie is currently a PhD student at the Flinders Accelerator for Microbiome Exploration (FAME) supervised by Prof. Rob Edwards. She has undertaken Bachelors degrees in Molecular Biology and Mathematics and in 2021 completed Honours in Microbial Ecology where she used metagenomics to study migrating wastewater communities. Her current PhD research aims to develop probabilistic genome annotation models which facilitate annotation of viral proteins with unknown function.

Bhavya Nalagampalli Papudeshi

Bhavya is currently a Ph.D. student at The Flinders Accelerator for Microbiome Exploration (FAME) an interdisciplinary research group at Flinders University. My research currently focuses on developing methods to assemble and annotate bacterial and viral genomes isolated from different environments. Through these methods we improve taxonomic and functional annotations and our understanding of the role of microbes within microbial communities.

Nathan Watson-Haigh

Nathan has 20 years of bioinformatics experience, with much of this time spent outside of the traditional academic research setting. He has a true multidisciplinary skill-set, possessing a broad understanding of many areas of bioinformatics coupled with a deep knowledge of genome informatics. He first started down the slippery slope of bioinformatics when undertaking his PhD in phylogenetics and molecular sequence analysis. His undergraduate training in pharmacology has always stood him in good stead, allowing him to communicate effectively with biologists. A significant amount of Nathan's time has been spent working with biologists to understand a problem space, devise genomic experiments, undertake analyses and to present findings to the researchers. An exemplar of this type of work was the development of a genomics resource

called DAWN (Diversity Among Wheat geNomes), together with the reproducible analytical workflow for processing many terabytes of raw sequence data using Snakemake. The driving force behind DAWN, was the need to provide wheat researchers with easy access to a plethora of public genomic/transcriptomic data sets in order to facilitate the identification of genomic loci responsible for phenotypic traits of interest.

Rob Edwards

After receiving his Ph. D. from the University of Sussex, in England studying nitrogen regulation in bacteria, Dr. Edwards moved to the United States to continue his studies. He worked as a Post-Doctoral Researcher at the University of Pennsylvania, Philadelphia, understanding how a leading cause of traveller's diarrhea (*E. coli*) causes disease. Dr. Edwards then moved to the University of Illinois, Urbana Champaign to study another food-borne pathogen, *Salmonella*. These studies merged the nascent area of genomics with traditional microbial genetics to investigate how a particular type of *Salmonella* became the leading cause of food-borne illness in the United States.

From 2000 to 2004, Dr. Edwards was an Assistant Professor at the University of Tennessee Health Sciences Center in Memphis, TN. Here, Dr. Edwards continued his studies on pathogenic bacteria, notably *Salmonella* and the bioterrorism weapon *Francisella*. Dr. Edwards received FBI clearance to work on these bacteria and was invited to the NIH to comment on the use of Select Agents at basic research laboratories.

In 2004, Dr. Edwards moved to the non-profit Fellowship for Interpretation of Genomes to work at the interface of biologists and computer scientists and worked with their team at Argonne National Laboratory. He remains an active software developer for Argonne and the Fellowship, developing open-source software including PERL and Python software for biological analysis and parallel computing that are used by scientists worldwide. Using breakout DNA sequencing technologies, Dr. Edwards' studies have continually pushed the forefront of both sequencing technology and bioinformatics. His work has been published in leading journals including multiple papers in both *Nature* and *Science*.

Dr. Edwards returned to academia in 2007, taking a research and teaching position in the Departments of Computer Science and Biology at San Diego State University where he rose through the ranks to become a Full Professor. He continued to work at the interface of biology and computing. The National Institutes of Health, the National Science Foundation, the Department of Education, the Department of Defense, the USGS, and private donors funded Dr. Edwards' research at SDSU, and his work led to breakthroughs in our understanding of how viruses interact with their hosts, and how viruses from around the world carry important genetic information. Dr. Edwards has continued to push current sequencing and bioinformatics technologies, in 2013 took a next-generation sequencing machine to the remote Southern Line Islands to explore metagenomics of coral reefs in real-time. In 2014 Dr. Edwards' team identified a virus that is present in the intestines of approximately half the people in the world, and in 2019 Dr. Edwards demonstrated the global spread of the virus in a paper that includes collaborators from every continent who collected and sequenced samples. In 2017, Dr. Edwards was elected to the American Academy of Microbiology in recognition of his contributions to the field of microbiology. In 2020, Dr. Edwards took the position of Matthew Flinders Fellow in Bioinformatics at Flinders University, in Adelaide, South Australia, Australia to start the Flinders Accelerator for Microbiome Exploration, to enhance microbiome and metagenome studies in South Australia.

Committed to teaching, Dr. Edwards received the graduate student award for the outstanding educator at the University of Tennessee, the teacher-scholar award and outstanding faculty award four times at San Diego State University. He was Graduate Advisor to the Biological and Medical Informatics Program at SDSU. Rob travels extensively to share his passion for bioinformatics and has taught bioinformatics classes around the US, and in Australia, China, Chile, Europe, Mexico, and North and South America. Dr. Edwards holds a visiting professor position at the Federal University of Rio de Janeiro.

In addition to science and teaching Dr. Edwards is also an advanced scientific SCUBA diver having led teams to study Coral Reefs all over the world. In his spare time, he is an avid international yachtsman, navigating in long-distance offshore races, including navigating the 2019 TransPac race from Los Angeles to Honolulu finishing 4th out of 89 boats.

Liz Dinsdale

Prof. Liz Dinsdale was a research officer at James Cook University, prior to receiving her Ph.D. (1992-2001), conducting research into the ecology of coral along the entire Great Barrier Reef. Highlights of the research included a latitudinal analysis of coral recruits, adult cover, and fecundity published in *Nature*. In addition, Dinsdale documented the prevalence of coral disease occurring along the GBR.

Dr Dinsdale received her Ph.D. from James Cook University, Australia in 2005. Her Ph.D. integrated ecological and social measures of coral reef health, to identify indicators to measure the effectiveness of conservation strategies. While ecology and social science have different philosophical underpinnings, indicators that have meaning to different members of the community were identified and were able to be cross-referenced with ecological measures. Therefore, a description of the health of a coral reef has community and ecological significance.

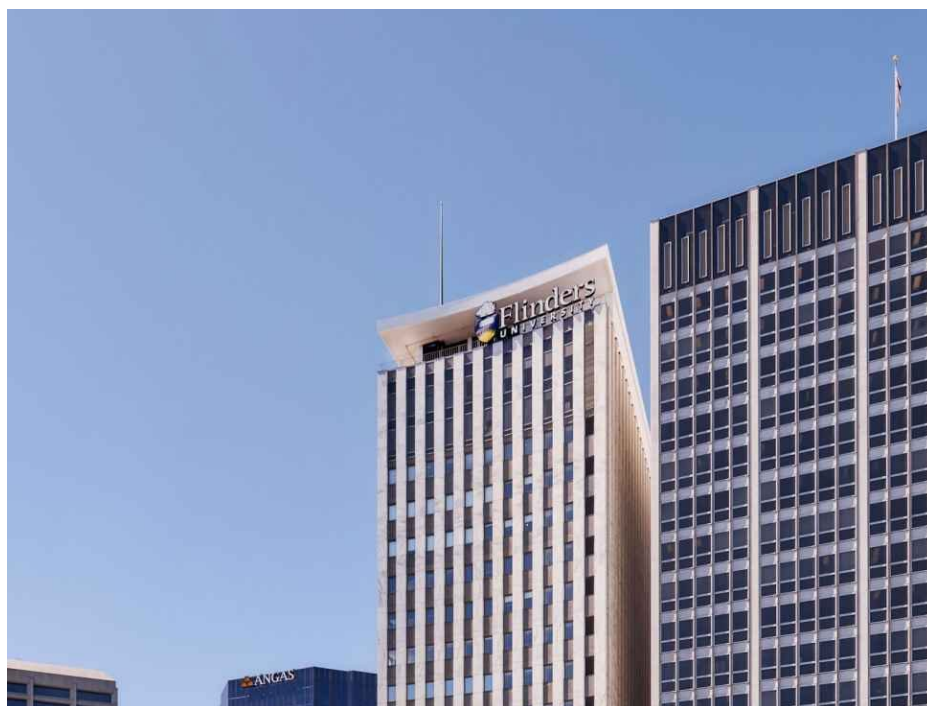
After spending a great deal of time investigating the macro-fauna that lives on coral reefs, Dinsdale joined a cruise to the central Pacific with researchers from San Diego State University and the Scripps Institute of Oceanography and developed techniques to describe the effects of microbes to coral reef ecology. During this cruise some of the earliest coral metagenomes were produced and the analysis identified that on a degraded coral reef not only are the shark, fish, coral and algae perturbed, but so are the microbes. The metagenomes showed that the microbes are becoming more prevalent and virulent with degradation and are a driver of coral reef decline. The Pacific Ocean findings were published in *PloS One* and subsequently combined with research for across three ocean basins, which confirmed the role microbes in the death of coral reefs, published in *Nature Microbiology*.

The involvement in microbial research was a turning point in Dinsdale's career as she integrated her ecological approaches to the new metagenomic methods. Dinsdale in 2008 used metagenomic data to describe the functional profiles of nine biomes published in *Nature*. Dinsdale was employed at San Diego State University in 2009 and developed a new course in 2010 that enabled undergraduate students to use one of the earliest models of next generation sequencer, the Roche 454 Flex. The undergraduate class, conducted for 10 years, has enabled students to use Ion Torrent, Illumina and Nanopore technologies. The students are directly participating in research and learning skills that make them highly sought after in the biotechnology industry. Dinsdale currently runs a DNA sequencing and annotation workshop for Faculty and young researchers in USA, and bioinformatics workshops in Mexico, Australia, Brazil, and Georgia. Dinsdale has received numerous National Science Foundation grants to continue the development of undergraduate education, with a focus of integrating bioinformatics into the life sciences undergraduate curriculum.

Dinsdale and her students are investigating the effects of microbes and viruses in kelp forests, coral reefs, oxygen minimum zones, and on the skin of sharks. Kelp forests are an important habitat forming species in temperate environments and we identified that the microbial community is affected by anthropogenic stressors and these genetically distinct microbes affect the growth of the kelp recruits, which may explain the decline in kelp over the last 50 years. The analysis of coral reef and microbes continue to identify new relationships, including the measurable effect of benthic macro-organisms on the microbiome of the surrounding water. The Dinsdale lab has been involved in two global analysis that question some basic principal of microbial ecology, the first identified that viral lifestyles switch with microbial density published in *Nature* and the second identified that microbial taxonomy and function showed different biogeographical patterns.

Elasmobranchs, including sharks and rays are a focus of the current research, as their skin structure, which cover in dermal denticles, provides an unusual surface for microbial growth. The dermal denticles improve their swimming ability of the shark and reduces the microbial biofilms, however, the microbiome that establishes is highly selected and displays phyllosymbiosis, published in the journal *Microbiome*. Recently, Dinsdale has moved to Flinders University in South Australia to continue microbiome research and investigating the mechanism that develop and maintain microbiomes on sharks and other marine organisms. When not in the lab, Dinsdale is an avid rower, competing in singles to eights in local, national and international regattas!

The Venue



Flinders at Victoria Square, 182 Victoria Square, Adelaide, South Australia

<https://www.flinders.edu.au/campus/victoria-square>

Located in the heart of Adelaide CBD, Flinders at Victoria Square is an accessible, alternative and flexible space for students and staff. We teach a selection of courses and topics and provide general information to students who want to study at Flinders.

The SAGC Metagenomics Workshop short course will be located in level 1 room 2, of the Flinders at Victoria Square building.

Transport

Public transport is recommended, and Flinders Victoria Square is conveniently located close to the tram and public bus lines. Adelaide Metro looks after the public transport including bus, train and trams for Adelaide city and the surrounding suburbs. The Adelaide Metro website has detailed information about the full range of public transport services and timetables including the Terrace to Terrace, Terrace to Entertainment Centre, Terrace to Botanic Gardens and Botanic Gardens to Entertainment Centre free city tram.

Please note that there is no staff or student parking available in the building. A range of commercial car parks offer plenty of off-street spaces:

- Secure Parking (220m)
- Wilson Parking (250m)
- UPark Topham Mall (450m)
- UPark Wyatt Street (450m)
- CarePark (500m)
- Upark Grote Street (550m)
- Myer Centre Car Park (750m)

Manual

This workshop is supported by the online manual available at <https://bioinf.cc/workshop2022/>

Contact Details

Prof Rob Edwards, Flinders University

Phone: (08) 8201 3417

Email: robert.edwards@flinders.edu.au

Concierge at Flinders, Victoria Square

Phone: (08) 7421 9406 - use this number to contact Security also.

Email: concierge.services.tonsley@flinders.edu.au

Location: Ground floor Flinders at Victoria Square Hours: 8.30 am to 4.30 pm Monday to Friday

Prerequisites

Before coming to the workshop we ask that you please take a moment to install some common software. If you have this preinstalled it will smooth the workshop process, but we will also aim to help you with the installation during the workshop.

PATRIC BRC Account: Please create an account at PATRIC BRC (<https://patricbrc.org/>).

PyCharm: <https://www.jetbrains.com/pycharm/> This is a all-in-one software editor for Python and we use it every day. PyCharm has a free license for educational users (<https://www.jetbrains.com/education/>)

RStudio: <https://www.rstudio.com/> is the go-to environment for writing R code.

Windows Users

If you are a Windows user, you will need access to a terminal emulator. We recommend MobaXterm (<https://mobaxterm.mobatek.net/download-home-edition.html>) which is free and very feature rich. (MacOS users already have something similar installed, and we'll show you how to access that).