Australian National University

Research School of Biology Summer Scholars ePortfolio

2018/2019

ANU COLLEGE OF SCIENCE

WHAT HAPPENED TO THE ROCKY MOUNTAIN LOCUST?

Peiyu Yuan, ANU
Supervisor: Sasha Mikheyev

We may be in the middle of an insect mass extinction. The Rocky Mountain locust (*Melanplus spretus*), was once dominant in the North American ecosystem. It was found in western Canada and the United States until the end of the 19th century. It occurred in incredible masses. One sighting in 1875, is believed to be the biggest aggregation of insects on record with 12 trillion individuals. However, in 1902, the last alighting of this species was recorded in southern Canada. Question have been raised about both the causes for its extinction and the status of the locust.

My project involved genomic comparison of related extant locust species, to those generated from museum data. In particular, I aimed to identify bioinformatic artefacts inherent to museum samples, which may provide direct comparisons between the two data sets.

Due to the low coverage of the museum samples of the extinct species and draft genome for the present-day samples, I filtered and optimize the quality of the samples by vcftools. However, under the visualization under Geneious Prime, there is many potential misalignments present in the museum samples due to the short fragment. Then I genotyped the museum sanguinipes and spretus individuals using samtools based on the bam filed offered at all the markers in the reference. To further improve the quality of samples, I calculated the informativenss of each single nucleotide polymorphisms and selected the most informative loci at various standard which was used for the phylogenies tree generation. The draft phylogenies tree was generated by the package “SNPRelate” which illustrated that the museum samples are grouped together, indicating that the artifacts are driving the data.

Apart from the finding of the research, I am grateful to Sasha’s suggestion and had an opportunity to attend the three-day SNPs in population- and phylo-genomics workshop in University of Canberra, which impressed me a lot about those excellent researchers in evolution field. I am immensely grateful to my supervisors, Sasha (Alexander) Mikheyev, who provided the patient guidance, encouragement during my research. I have been very lucky to have supervisors who cared so much about my work, and who responded to my questions so promptly, although I am extremely new in this area.

INVESTIGATING EFFECTOR RECOGNITION IN WHEAT PROTOPLASTS.

Salome Wilson, University of Queensland
Supervisor: Benjamin Schwessinger

My project aimed to refine a wheat protoplast assay for use in transient gene expression studies to test effector interactions and recognition.

Effectors are small proteins that are secreted by pathogens to suppress or manipulate plant defence systems. Plants have, in turn, evolved their own mechanisms to detect and respond to specific pathogen effectors. This gene-for-gene interaction is analogous to an arms race, with selection pressures driving the evolution of these mechanisms on both sides. Thanks to advances in whole genome sequencing, candidate effectors can be predicted, however they require functional validation. Many of the approaches used in other systems are not efficient for transforming wheat, therefore refining a robust and high-throughput assay to investigate gene expression in vivo opens up opportunities for screening many different gene functions.

As part of this project I have begun to generate plasmid constructs to test the expression of pathogenesis-related genes, which contribute to the plant’s defence responses when a pathogen is detected. I will use the these as reporters for gene-for-gene interactions in protoplasts.

A highlight of the program was the opportunity to attend the Stromlo Plant Pathology conference, giving me the opportunity to engage with a wide range of exciting research in the field of molecular plant pathogen interactions. Senior staff and fellow students in the Research School of Biology provided encouragement and support, helping me to step outside of my academic comfort zone and generously sharing their advice and experience with me.

DISCOVERING NOVEL ANTIMALARIAL DRUG TARGETS

Sathya Perera, University of Tasmania
Supervisors: Christina Spry and Kevin Saliba

Malaria is an important global concern, affecting over 200 million and killing around half a million people annually. This disease is caused by parasites of the genus *Plasmodium*, with *Plasmodium falciparum* causing the most severe infections. With the prevalence of drug resistant *P.falciparum* infections rising rapidly, antimalarial drugs with novel targets are essential.

Plasmodium species are transmitted by female Anopheles mosquitoes and have a complex life cycle. In humans, the parasite

first needs to invade liver cells to mature into the merozoite stage which can infect red cells. Inside red cells parasites reproduce asexually, rupturing the cell in approximately 48 hours to infect other red cells and continue their cycle. The red cell stage is responsible for the clinical symptoms of malaria.

My project focused on investigating the mechanism of action of a group of novel antiplasmodial compounds identified through high-throughput screening. I worked with the blood stage parasites of a specific strain of *P. falciparum*. I maintained routine cultures of them and assessed their growth in response to varying concentrations of the antiplasmodial compounds. Parasite growth was measured in two different ways: 1) using SYBR Green-I, a dye which fluoresces upon binding to double-stranded DNA, to detect parasite DNA, and (2) using the absorbance-based 'Malstat assay', in which *Plasmodium* lactate dehydrogenase is detected by monitoring its activity in reducing APAD (3-acetyl pyridine dinucleotide), which in turn is coupled to the conversion of tetrazolium (yellow) to formazan (blue). I also performed 'parasite rescue assays', where I attempted to identify antagonisers of the inhibitory effect of the antiplasmodial compounds as a means to elucidate the corresponding targets. Additionally, I attempted to establish a fluorescence-based assay for measuring the production of reactive oxidative species (ROS) in parasite-infected red blood cells using the dye CellROX.

My work has: (1) provided confirmation, by independent means, of the inhibitory activity of previously identified antiplasmodial compounds;
2) revealed compounds which "rescue" or antagonise the activity of some antiplasmodial compounds, with some antagonisers being unique to certain antiplasmodial compounds, consistent with the compounds studied having different mechanisms of action.
(3) shown that a commonly used positive control in ROS measurements kills malaria parasites at concentrations often used in ROS assays, thus necessitating the use of lower concentrations in future assays.

The SRS program at Saliba Lab has been a fascinating experience! I have enjoyed every moment of it, with the main highlights being: the Saliba lab team, workshops, guest presentations and festivities. I am grateful to my supervisors and ANU RSB for being so welcoming and this indelible opportunity to gain an insight to drug discovery research with an immensely supportive team.

ALTERNATIVE MATING STRATEGIES AND SPERM COMPETITION: A META-ANALYSIS

Michael Skirrow, Massey University, Auckland, New Zealand
Supervisor: Michael Jennions

Michael Skirrow and Professor Michael Jennions

Objectives:
The main objective of this project was to address two key research questions using a meta-analytic framework:
1) Does investment in sperm competition differ significantly between males that employ the main versus alternative mating strategies when each strategy has a fixed morphology?
2) Does investment in sperm competition differ with the proportion of individuals in the population employing an alternative mating strategy which has a fixed morphology?

Importance:
This project will clarify the effect of different mating strategies on sperm competition parameters across a number of taxonomic groups. It will provide a quantitative test of sperm competition theory, and provide an excellent starting point for further examination of how alternative mating strategies affect investment into traits that are advantageous under sperm competition. The study can also be extended to include the examination of how alternative behavioural tactics affect investment in sperm competitiveness in future research.

Research Approach:
To conduct this research, I used a meta-analytic approach common to ecological reviews. First, I predefined search terms* and inclusion criteria to allow for the objective screening of scientific publications during searches of the literature database Web of Science. I then uploaded the titles and abstracts of the publications obtained during literature searches to Rayyan, to form an online database. I then used the predefined criteria that I developed to include or exclude publications as appropriate. The included publications will later be examined in depth to extract data and to calculate effect sizes relating to the two main research questions.

Findings:
My initial literature searches returned a total of 5891 publication 'hits' from 19 sets of search terms*. From these publication hits, I identified 2604 unique papers. After screening, I identified 218 publications that were suitable for inclusion under my criteria. In a second literature search, I did a reverse search of highly cited publications related to the study questions. This returned a further 1704 publication hits, containing 479 duplicates. The 1225 unique papers identified remain to be screened as part of my ongoing research collaboration with the Jennions Group.

UNTANGLING WOLBACHIA GENOMES FROM VARROA MITE METAGENOMES

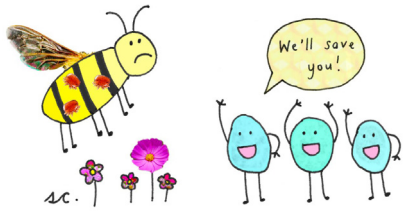
Stephanie Chen, University of Sydney
Supervisor: Sasha Mikheyev

Metagenomics is a branch of bioinformatics which aims to characterise the collective genome of microorganisms of entire ecological communities from a sample. This project capitalised on existing DNA sequence data generated by the Mikheyev Group for the genome assembly of Varroa mites (*Varroa destructor* and *V. jacobsoni*) and involved a three phase approach – Varroa mite metagenome assembly, recovery and comparison of *Wolbachia* genomes, and genome annotation.

Wolbachia are endosymbionts found in a significant proportion of arthropods. These bacteria are well known for inducing parthenogenesis, feminisation, and male-killing and also form mutualistic relationships with hosts such as bedbugs (*Cimex lectularius*) and the nematode family Onchocercidae. Importantly, Wolbachia have the potential to supress vector-borne diseases such as Zika and dengue in humans as well as bee viruses, which have become a significant worldwide threat to apiculture. My summer project investigated the biogeography and evolution of the Varroa mite, an important parasite of the European honey bee (*Apis mellifera*), and focused on its symbiotic relationship with *Wolbachia*.

I used the software Anvi'o to analyse and visualise genomic data and gained experience with high-performance computing (HPC). A diverse range of archaea, viruses, and bacteria, including *Wolbachia*, was detected in the assembled metagenomes of 44 samples collected across 7 countries. *Wolbachia* sequences were present and recovered in 20 % of samples which encompassed the hosts *A. mellifera* and *A. cerana* (Asian honey bee). Comparison of the extracted *Wolbachia* strains to each other and strains published in the literature will contribute to an increased understanding of host, parasite, and endosymbiont biology which is essential for the application of *Wolbachia* in disease control.

The ANU Summer Research Scholarship has been a rewarding research and networking experience that has helped me consider the direction of my further studies.



Wollbachia to the rescue!

THE HEAT IS ON - PROGRAMMING PLANT SURVIVAL

Bridget Fellow, University of Otago, New Zealand
Supervisor: Spencer Whitney

Over the summer my project sought to develop and test differing synthetic genetic networks in plants that can be used to differentially trigger the production of specialized proteins under elevated temperatures. The proteins of choice are those involved in photosynthesis - the biochemical process underpinning survival of the biosphere, and one in which becomes limited by temperature stress. Improving the heat tolerance of crops is a key driver for enhancing agricultural yield, particularly in a world with a booming population and that is increasingly prone to climate change effects.

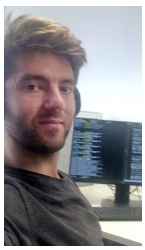
The genetic circuitry developed in my project comprised a range of differing heat sensitive promoters, designer transcription-like effectors (dTALEs) and the marker proteins red fluorescent protein (RFP) and a custom designed thermostable Rubisco activase (RCA). The synthetic, multi-gene, plant transformation plasmids were assembled using Golden Gate cloning and their efficacy tested by transient expression in *Nicotiana benthamiana* plants.

Fluorescent imaging, SDS-PAGE gels and Western blots revealed variability in protein expression among the matrix of expression constructs tested. Most of the heat sensitive promoters were found to stimulate marker protein expression when the plants were exposed for 1 hour to elevated temperature, sometimes producing more protein than in the control plants where marker protein production was regulated by the strong constitutive CaMV 35S promoter. Further fine tuning of the programmable genetic circuitry I developed is now underway within the Whitney laboratory.

The SRS program was a great opportunity to learn a lot of new techniques and skills inside a well-equipped lab in addition to meeting a diverse group of people from across ANU that were involved in the summer scholar program. There was a great balance of working and socialising with many activities organised over the 8-week program. I met so many terrific people and had a great summer exploring all that Canberra had to offer.

PROBING THE EVOLUTION OF A PEST MOTH AROUND AUSTRALIA

Leo Featherstone, The University of Melbourne
Supervisor: Angela McGaughran, Moritz group.



My project was about using the DNA of a pervasive pest-moth to learn about how it evolved pesticide resistance around Australia. The moth is a widespread pest called *Helicoverpa armigera* and it is thought to be resistant to more pesticides than any other agricultural pest. Combined with the fact it can feed off of at least 200 kinds of crop, this means it is a major pest and we need to understand how it rapidly evolves pesticide resistance to better manage it.

Beginning with DNA sequences from hundreds of samples of these moths, I wrote code to run programs that allowed us to assess the quality of the data and filter it for reliable signals in the DNA. We then used these DNA signals to analyse how related *H. armigera* populations are around the country. This in turn could inform us about how pesticide resistant genes might be able to spread between populations.

The main finding of the project was that there is a high degree of relatedness between Australian *H. armigera* populations, as well as more relatedness to populations from other countries than we previously thought. My supervisor and I are continuing to work on this to see if we can more closely identify the change in DNA as we move from global to Australian populations.

I found the opportunity to present my work to my lab group really rewarding and I enjoyed expanding my programming repertoire to achieve the aims of this project. It was also valuable to experience research outside of my home university. I am extremely grateful for having had this opportunity and would recommend the program to anybody else who gets the chance!

Helicoverpa armigera.

WHERE THE WILD THINGS ARE – USING NUTRITIONAL QUALITY TO ASSESS KOALA HABITAT SUITABILITY

Mahalia Kingsley, University of Tasmania
Supervisors: Kara Youngentob and Karen Ford, Foley group.

Tree use by koalas (*Phascolarctos cinereus*) is not only determined by the physical attributes of Eucalyptus trees (e.g. size, age, species), but it is also influenced by the chemical composition of their leaves. In order to deter herbivory, Eucalyptus trees produce various types and concentrations of plant secondary metabolites (PSMs) which may prevent protein digestion by herbivores, or act as toxins that cause discomfort, pain, nausea, or death. Traditional koala habitat surveys tend to ignore PSMs in

their assessment of habitat suitability, instead focusing on mapping the patchy distribution of koala food species within a region. My project involved comparing koala location data to indicators of forage quality in a forest area near Newcastle in the aim of demonstrating how measures of foliar quality can be incorporated into future koala habitat assessment surveys.

Leaf samples were collected from four focal trees and their three nearest neighbours along nine 250m transects (144 samples total), with an additional 57 spot-samples collected from areas of interest. Leaf samples were obtained by shooting a branch from the top half of the canopy using a modified slingshot system. Samples were freeze-dried and ground to pass a 0.5mm sieve before spectra were obtained by scanning with a Near Infrared Spectrophotometer. Spectral data was analysed for concentrations of UBFs, FPCs, and total and available nitrogen. The preliminary results show that the majority of trees sampled demonstrated above 1% total nitrogen content which is noteworthy as it has been proposed that this is the critical value below which koalas are incapable of maintaining positive nitrogen balance when feeding only on foliage of this nutritional quality. Additionally, most trees sampled had FPC content less than 15mg/g dry weight, which was observed to be the level at which koalas are deterred from feeding on such trees when more nutritious trees are also available. Our methodology also allows us to map nutritional quality in greater detail over the property surveyed, and hence would give a better indication of the areas that koalas feed from within the landscape if incorporated into future koala habitat surveys.

I have had many great experiences through this program, such as using slingshots to collect leaf samples from the canopy of the trees or using the blade of the cutting pole to open icy poles on the hotter days during field work (see image above). Field work also had its challenges, such as the constant struggle to deter biting bugs like march flies and mosquitos, and the wear-and-tear associated with stripping branches from over 200 trees. It all seemed worth it, however, when we spotted a koala as we walked back to the car (see image below).

There was some tedium, such as when I had to find approximately 900 samples in the storeroom, or when I spent my Christmas break grinding leaf samples. I believe my most memorable experience was having to carry 900 leaf samples in two large suitcases so that I could travel with them on the train to Sydney. I did this so that I could scan the ground samples using the Near Infrared Spectrophotometer at the Hawkesbury Institute of the Environment. I have had a wonderful time meeting and collaborating with other researchers at Western Sydney University and the Australian National University over the summer of 2018/19.

THE LOCALISATION OF DEPHOSPHO-COENZYME A KINASE (DPCK) IN PLASMODIUM FALCIPARUM

Ella Forward-Yang, ANU
Supervisors: Christina Spry and Kevin Saliba

Dephospho-coenzyme A kinase (DPCK) is an enzyme that is important in the conversion of pantothenate (a vitamin) into coenzyme A (an essential cofactor for many metabolic processes in *Plasmodium falciparum*). The objective of my project was to use molecular biology techniques to determine localisation of STREP-tagged DPCK in *P. falciparum*, which is believed to exist in an organelle called the apicoplast.

Roughly half a million people die each year from malaria, due to infection of *Plasmodium falciparum* parasites. *P. falciparum* parasites were once very responsive to drugs, but with resistance developing in the parasites, past drugs have since become ineffective. This project is important because the findings may lead to a new, effective antimalarial drug target.

The first approach we took before starting localisation of DPCK was to confirm that parasites transfected with a plasmid encoding STREP-tagged DPCK were expressing the STREP-tagged DPCK protein. To do this, protein was extracted from the cells and western blotted using anti-STREP antibodies. However, this proved to be difficult, as the experiment did not produce the expected bands. Because of this, different experimental approaches were carried out to troubleshoot the issues we encountered. For example, we tried increasing the volumes of lysis buffer and introducing a mechanical disruption step to ensure sufficient lysis of the apicoplast. We also performed PCR analyses to ensure that the cells still maintained the plasmid encoding the STREP-tagged DPCK protein.

The results from my project were an important step forward. For instance, the findings from my experiment confirmed that the parasites still retained the plasmid encoding the STREP-tagged DPCK protein. In addition, detection of the Acyl Carrier Protein (ACP), showed that proper lysis of the apicoplast occurred under the optimised conditions.

The SRS program gave me an opportunity to experience day-to-day life working in a research lab, which is important in deciding what I want to do in the future. A big part of the SRS program for me was growing as an individual in science; how to move on from failed experiments, learning how to troubleshoot my experiments and refining my molecular biology techniques, which I can carry with me in this field. I very much valued meeting and getting to know the people in the team I was working with! Thank you to everyone in the team for making me feel so welcomed!

INSIGHTS INTO PATHOGENIC PROTEIN RECOGNITION

Carl McCombe, Flinders University
Supervisor: Simon Williams

Plant pathogens secrete proteins called effectors into host cells to aid infection and increase disease susceptibility. However, if an effector protein is recognised by the host plant, effector-triggered immunity (ETI) is activated, resulting in an immune

response and disease resistance. In order to develop resistant plants capable of detecting effector proteins, we need to understand how plant immunity receptors recognise effectors.

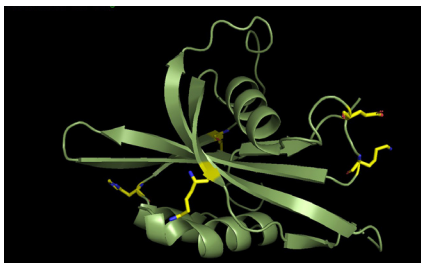


Figure 1: AvrM14 3-D structure. All amino acids that vary between the avirulent and virulent form are shown in yellow.

In the summer project at ANU I aimed to characterise the structural differences between two allelic variants of an effector protein secreted by flax rust fungi. The avirulent type (AvrM14) is recognised by flax resistance proteins M1 and M4, and this results in plant immunity. In contrast, the protein product of the virulent allele (avrM14) is not recognised by M1 and M4, and acts to promote disease. Previous work by Dr Williams and collaborators have solved the structure of AvrM14 (Figure 1), however the structure of avrM14 remains unknown. Determining the avrM14 structure will allow us to identify the specific protein regions required for recognition by the M1 and M4 flax resistance proteins, potentially providing novel insights into effector recognition.

My project was focused on expressing, purifying, and crystallising the avrM14 protein to enable structural determination by X-ray crystallography. Recombinant protein expression was carried out in *Escherichia coli*. Purification was achieved using Ni-affinity chromatography followed by gel filtration chromatography, and the purified protein was used in crystallisation experiments. We were successful in producing avrM14 protein crystals (Figure 2), which can now be used for structure determination utilising X-ray sources from the Australian synchrotron.

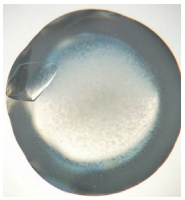


Figure 2: An arvM14 protein crystal

For me the summer research project was a fantastic opportunity to experience research in world-class facilities, and work alongside experienced staff and students. In addition to the meaningful research, the summer scholar experience resulted in the formation of various friendships with other promising science students from around Australia and New Zealand. I would like to thank ANU, the Research School of Biology, and Dr Williams for providing me with this opportunity.

CHARACTERISING ERROR PROFILES IN THE OXFORD NANOPORE SEQUENCING DATA USED FOR FUNGAL SPECIES IDENTIFICATION

Tavish Eenjes, ANU
Supervisor: Benjamin Schwessinger.

Fungi can pose a major threat to crop yield, and understanding the fungi present on and around crops enables more effective control measures. While early detection is preferable, most genetic sequencers are bulky, expensive and limited to lab-spaces far from the land of farmers. On the contrary the MinION sequencer, from Oxford Nanopore Technologies, is portable and runs off a standard laptop. This enables close to real-time analysis and data streaming. However, it has a trade-off, with a relatively low per read sequencing accuracy between 80-90%.

Internal transcribed spacer (ITS) regions, of about 500 bases in size, within the ribosomal RNA loci in fungi is often used for fungal species identification with other more accurate sequencing platforms. Here I studied the error profile of the whole ribosomal RNA locus, of about 3 kilo bases in size (long ITS), on the MinION sequencing platform. I compared the DNA read sequences of long ITS amplicons of multiple fungi with the corresponding locus in publicly available genome sequences of the corresponding fungus. This was done bioinformatically, using a combination of pre-packaged modules and self-written scripts to read, filter, slice and compare the reference genome to the read data. Through this I developed a deeper understanding of both python coding and, through reading papers, the current practices and methods of the field.

An understanding of the error rate and characterisation is required to determine how useful the MinION is for detection of unknown samples when running against a database of known species.

The results indicate that the sequencer has an accuracy of approximately 81.5%, with the majority (approximately 15%) of inaccuracies due to apparent insertion and deletion errors. This is likely caused by the difficulties the nanopore approach has with reading sequences with repeated motifs of a single base. Additionally, there was an increased rate of purine transition errors (between Adenine and Guanine), but no evidence of an increased rate of pyrimidine transition errors. This suggests a bias towards this type of error in the nanopore sequencing.

This was an amazing experience, and I'm now looking forward to continued study with the Schwessinger lab and beyond.

WAHLENBERGIA CERACEA F3 GERMINATION TRIAL THE VIABILITY OF SEEDS DEVELOPED AND GROWN IN HOT VS. COLD ENVIRONMENTS

Pamudika Kiridena, ANU
Supervisor: Pieter Arnold, Nicotra group.

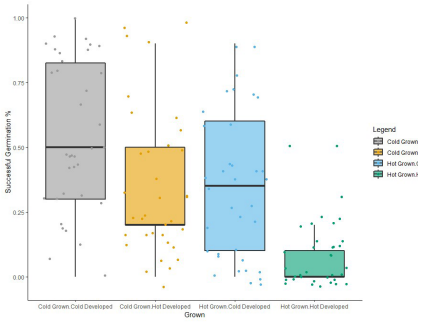
The overall aim of this study was to investigate the viability of *Wahlenbergia ceracea* F3 seeds developed and grown in hot vs. cold environments. A key objective was to assess the viability of *W. ceracea* F3 seed through visual assessment and germination success.

Given rising temperatures across many regions of the world, it is imperative we develop a deeper understanding of the effects on at-risk regions such as the Australian alps where temperatures are expected to rise 4-5˚C by 2100. Therefore, it is important to study alpine species such as *W. ceracea* and their response to different climatic conditions.

To address the aims of this project I used controlled germination trials on seed visually scored for viability. High, medium and low quality seeds were determined based on colour and fullness, with darker browns and fullness indicating better quality. F2 of *W. ceracea* capsules that had developed seed (the F3 seeds used here) in hot and cold environments were harvested. These hot and cold developed seeds were collected from 28 lines which were represented in both growing environments. Agar was prepared, autoclaved and poured into 55mm petri dishes in a sterilised laminar flow. For each line, 60 seeds were randomly selected for use in 6 petri dishes with 10 seeds sown in each. Half of the petri dishes were placed in a cold and the other half in a hot environment to produce cold and hot grown seeds.

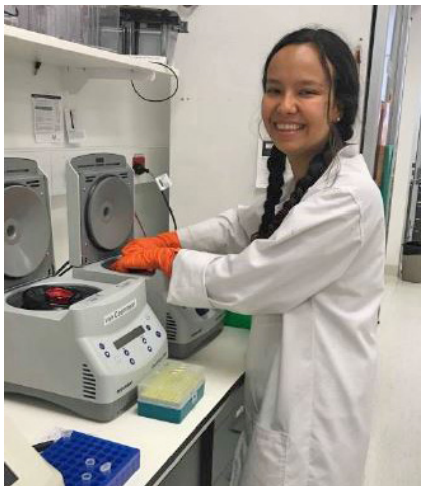
A key finding of the study was the poor correlation (R2 = -0.006) between seed

viability scores for the two methods used. Instead, most of the variation in germination rates was explained by differences among F3 lines (R2 = 0.388). This indicates that visual assessment may not be used as a credible approach to measuring seed viability. Additionally, hot developed seeds showed low germination rates regardless of the environment they were grown in ($F_{1,22} = 9.408$, $p = 0.006$). Low germination rates were also seen in hot grown seeds ($F_{1,30} = 4$, $p < 0.001$), indicating that both hot and cold developed seeds have higher germination success in cold environments. Overall, seed viability decreases when developed and/or grown in warmer environments. This may result in population decline for *W. ceracea* and therefore higher temperatures are a serious concern going forward.



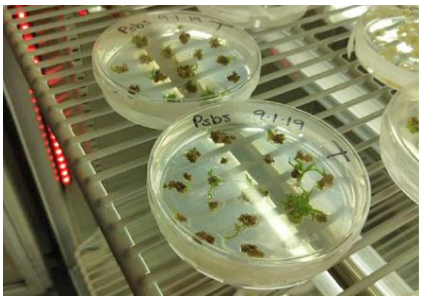
ESTABLISHING A CRISPR/CAS9 MEDIATED GENE KNOCK-OUT SYSTEM FOR SETARIA VIRDIS

Emily Watson, University of New South Wales
Supervisor: Maria Ermakova, von Caemmerer group.



How do you feed a growing population with finite land? With the earth's population predicted to hit 9.8 billion by 2050 we need a way to increase our food production capabilities.

The C4 photosynthetic pathway offers a potential solution. It is a biochemical carbon concentrating mechanism which increases rates of CO2 assimilation, making C4 plants some of the most efficient terrestrial primary producers. Engineering the C4 photosynthesis pathway into common C3 crop species such as rice is a long term goal which could help solve the world's food shortage issues. However, before we can engineer C4 photosynthesis, we need to understand the basic mechanisms of this complex process.



Model plant *Setaria viridis* is a popular choice for C4 research due to its small stature, simple growth requirements, rapid life cycle, relatively simple genome and established transformation techniques. My project aimed to establish knock-out lines of key genes involved in photosynthesis in *S. viridis*, to study the role of these genes in the C4 pathway. I designed plasmids which incorporate the CRISPR/Cas9 system, to make targeted double stranded breaks in the genomic DNA of *S. viridis* and produce gene deletions. Using the Golden Gate cloning system, I assembled plasmids to target four genes of interest, and have currently transformed two of these constructs into *S. viridis* with the rest to follow. Once these plants have regenerated they can be used to study key processes involved in the C4 pathway.

My time in the von Caemmerer lab has been a wonderful introduction to life in the lab, made even more memorable by the friendly and supportive community of researchers and summer scholars.

USING PATTERN RECOGNITION TO IDENTIFY WILD SIGANUS DOLIATUS

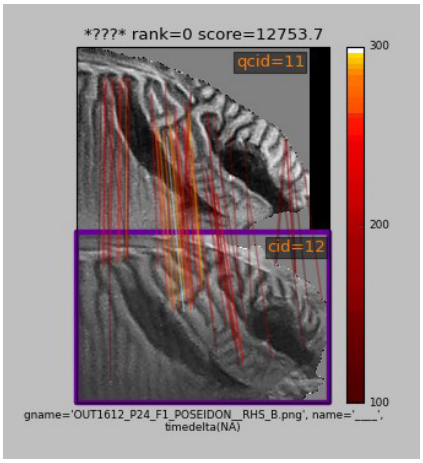
Kwan Lung Elroy Au, ANU
Supervisors: Megan Head, Rebecca Fox

Siganus doliatus, Scribbled rabbitfish.

Monitoring of individuals in a population is a powerful tool for wildlife biology, providing important information about population dynamics such as survival, abundance, recruitment and growth. However, traditional mark-recapture studies for population monitoring can be costly and time intensive, while also requiring invasive tagging of animals. However, for animals that exhibit unique facial or body patterns such as zebras, pattern recognition from field photographs represents an alternative to physical tagging. This process of photographic mark-recapture often employs pattern recognition programs to identify individuals over time using field photographs, and is often more cost effective.

We aimed to test the potential of using pattern recognition programs to identify *Siganus doliatus* (Scribbled rabbitfish) in the wild by using their facial patterns. In doing so, we hoped to be able to create a photographic mark-recapture protocol that could be utilised in future monitoring studies of this species.

Using pattern recognition programs as an identification method requires patterns to be unique to individuals and remain stable over time. Thus, we aimed to investigate whether the facial patterns of the scribbled rabbitfish were indeed unique between individuals using a database of photographs taken of dead *S. doliatus*. As underwater field photographs suffer from reduced quality, the second step was to examine whether these facial patterns were still distinguishable enough to identify live individuals in the wild.



Identification of live individuals from field photographs using a pattern recognition program.

Analysis of photographs of dead *S. doliatus* reveals the facial patterns to be highly unique and distinguishable between individual fish. Coupled with an identification rate of around 84% for live *S. doliatus* in the wild indicates that photographic mark-recapture using

pattern identification programs has strong future potential to be applied to monitoring of *S. doliatus* in the wild.

I would like to thank Dr. Megan Head and Dr. Rebecca Fox for their help and supervision in this project. Being a summer research scholar at ANU has been a profoundly rewarding and educational experience, and has allowed me to further explore doing research as a career path while meeting new people.

INCREASING GAS EXCHANGE BY OVEREXPRESSING PIP1;3 IN TOBACCO

Hanh Vo, ANU
Supervisor: Tory Clarke, von Caemmerer group.

Hanh Vo, picking out homozygous plants for gas exchange

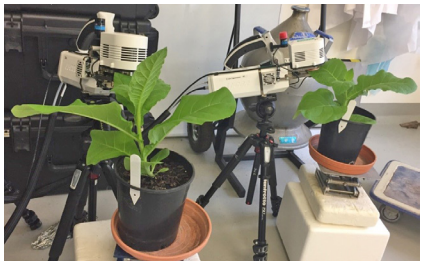
There is a mutation in the tobacco plants I worked with: they are overexpressing the *Arabidopsis thaliana* PIP1;3 protein. This protein is an aquaporin, and may transport CO2 into mesophyll cells, which could enhance carbon fixation. Therefore, by overexpressing it, we may increase mesophyll conductance, and increase photosynthesis and yield. This protein later can be expressed in potatoes, tomatoes, etc. to increase yield of those plants.

Wild type tobacco plants had already been transformed with the PIP1;3 gene, along with a selectable marker, hygromycin. I grew the T2 seeds on antibiotic agar plates for selection. By counting the number of resistant (big) and non-resistant (small) seedlings for the segregation ratio, I got an idea of whether there is a single or double insert in the lines. After the counting, I harvested the seedlings, extracted their RNA and used qRT-PCR to determine the expression level of the inserted PIP1;3 gene. This helped me to pick out three best lines (highest expression) which I grew again for phenotyping.

From these plants, I extracted DNA and used RT-PCR to determine the zygosity – either azygous (no insert), hemizygos (one copy) or homozygous (two copies). Using this information, I selected only the homozygous plants for phenotyping by gas exchange with the LiCOR 6800. The gas exchange was to see if there are any differences in CO2 assimilation between the mutants and the wild type.

When I analysed the results, there were no significant differences in assimilation rate between my wild-type controls, and the T2 overexpression lines. This might be because my estimation of zygosity based on qPCR was incorrect and there were some hemizygous plants, thus diluting any improved phenotype. There might also be transgene silencing due to interactions between the transgene inserts and/or with the host genes, thus the overall assimilation was the same. I could have some leaf samples to a digital PCR and confirm my zygosity prediction. After gas exchange measurements, I continued to grow my plants to collect their seeds as further phenotyping, including measuring mesophyll conductance directly, could be done on the stable T3 lines. Moreover, instead of using a transgene from A. thaliana, I could over-express the host gene for the same protein instead to avoid transgene silencing.

I'm grateful that I have Tory as my supervisor. At first, I got a little scared since I don't know what to do with all the seeds. However, Tory always made sure I know what each step is for, what to expect and what's the next step. Without her guidance, I would probably have no idea of what to do most of the time!



Measuring gas exchange using LiCor 6800 machines

GENETIC IMPUTATION IN BREAD WHEAT

Wenjun Liu, University of Adelaide
Supervisor: Justin Borevitz

Wheat is a worldwide staple food, but the size of its genome (six times larger than the haploid human genome) and its complexity made it challenging to understand the association between wheat traits and genetics.

My project was part of the Genome-wide Association Study within the Seri/Babax recombinant line population, which aims to find genetics information associated with drought adaption.

The population was consisted of 175 6th generation inbred offspring and two parents: Seri and Babax. Whole genome was sequenced using Illumina. The offspring were sequenced with extremely low coverage (x 0.1), while the parents were sequenced with higher coverage, at x 0.5. Because of the low coverage used, most of the samples were not genotyped for the single nucleotide polymorphisms (snps) we are interested at, so the aim of my project was to computationally impute the missing data, which counts for above 90% of the dataset.

My works were done mostly on Raijin using Bash and R. I tested two established imputation methods: Beagle 4.1 and LB-impute. Neither of them gave the anticipated results. Beagle was designed for imputing population sequencing data, where samples are independent from each other. The samples in our data, on the other hand, are too closely related to each other. So Beagle would impute most samples to have the same genotype for a certain snps, while we are expecting a ratio of 1:1 for genotypes. LB-impute, was designed to impute low-coverage sequencing datasets generated from biallelic populations, and the imputation is basing on the allelic depth of coverage information. However, because the sequencing depth of my data was too low, LB-impute would only impute part of the samples for each position.

So I choose to impute the data manually. Each chromosomes were divided into 10 Mb windows, and the genotyped data each sample has within that window are compared to the genotype of the two parents to decide which parent is the sample closer to in the window. The analysis were done and visualized using R. The imputed data showed unexpected recombination pattern and an overestimated heterozygosity, which might also be a result of low coverage in sequencing. For further studies, more designed imputation packages can be experimented on and the project group is also thinking of resequencing the population with another technique or higher coverages.

The 8-weeks summer project in RSB gave me chance to practice my programming skills and learn useful agricultural biological concepts. My supervisor and everyone else in Borevitz lab are very supportive and welcoming. They did not only give me suggestions on my projects, but also shared their own career pathway with me, which inspired me a lot. The project itself was very rewarding as I could see myself making progress and my works being acknowledged by other scientists. I really appreciate this opportunity and benefited a lot from it!

THE EFFECT OF TEMPERATURE AND THE CIRCADIAN CLOCK ON CARBON METABOLISM IN HORDEUM VULGARE

Aisha Ward, ANU
Supervisor: Anna Flis, Furbank group.

Objective: To determine if temperature has any significant effect on the night-time degradation of carbon storage reserves and to explore how the circadian clock exerts its regulation over the night-time carbon metabolism in different temperature regimes.

Why is it important: Agriculture is economically dominated by the need to grow food, with the majority of plants grown being monocot grains, such as barley, wheat, and rice. As population size increases, so does the demand for these plants. In the model plant species Arabidopsis thaliana, recent research has shown that the circadian clock has strong control over the regulation of carbon metabolism. It is theorised that being able to manipulate the control of this clock has the potential for improving plant performance, which would improve the efficiency of commercial plant growth and harvest. However, in monocots, the operation of the circadian clock and its co-regulation of carbon metabolism is obscure. Consequently, the aim of this research to be to expand our knowledge about the two processes in barley (*Hordeum vulgare*), which, as a model species for wheat could present the opportunity to increase the efficiency of these monocots.

What research approach: Wild-type barley, variety Bowman, and two mutants affected in the operation of the circadian clock (eam8 and eam10) were grown in controlled environment in a 12 h day - night cycles with temperature of 20°C during the day and 18°C during the night. At 14 days after sowing two sets of plants were exposed to two different night-time temperature regimes: warm night (18°C) and cold night (10°C). The plants were harvested at two-hour intervals, beginning directly before dusk and ending before dawn. Collected samples were grinded, aliquoted and metabolites were extracted by boiling in 80% ethanol. Soluble and non-soluble carbohydrates were then quantified using enzyme based spectrophotometric assay.

Findings: Consistent with results obtained from preliminary data, the non-storage metabolite (glucose and fructose) levels remained unchanged throughout the cold and warm night-time cycle while the storage metabolites (starch, sucrose, and soluble fructans) were degraded during the night-time cycle. In the warm night experiment, a significant difference in the levels of starch between the mutants and the wild-type at the end of the night was observed. This confirms the circadian regulation of night-time starch degradation, as the end-of-night levels of starch in the mutants were significantly lower than that of the wild-type. There was also an observed difference between the rates of starch degradation in the cold and warm night levels of the mutant, eam10, therefore showing evidence that temperature plays a major role in the regulation of the night-time degradation of storage compounds. Further analysis of this data, and repeated experiments will have to be conducted to confirm this in greater detail.

Highlights of the program: The internship was an unforgettable experience, through the support and passion of my supervisor. My experience in lab work has grown exponentially and I have been exposed to many new facets of research and careers that I had not previously considered for myself.

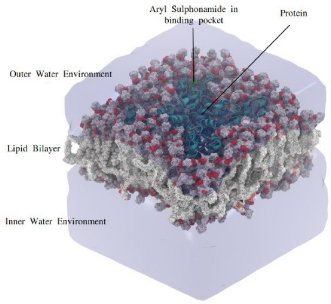
HOW DO SODIUM CHANNEL SUBTYPE SELECTIVE DRUGS FIND THEIR PREFERRED MATCH?

Josiah Bones, ANU
Supervisor: Ben Corry

Objective

My summer project was an in silico experiment aimed to determine the relative free energy difference in drug binding of various aryl sulfonamide compounds, known for their specificity towards sodium channel subtypes, at a specific protein residue location where these compounds bind, residue 1537.

The protein structure used was a human-bacterial hybrid sodium channel designed to mimic the human subtype sodium channel Nav1.7.



An example of a fully prepared Na_v 1.7 system with bound aryl sulfonamides compounds.

In sodium channel subtype Nav1.7 (linked to peripheral pain signaling), this residue of interest is a Tyrosine. In sodium channel subtype Nav1.3 (linked to neuropathic pain), it is a Serine. In my experiment I mutated the Tyrosine to a Serine. This was originally thought to negatively impact the binding of the various drugs, especially the Nav1.7 specific drugs, to this site, due to a loss of a favorable pi-pi interaction between the drug and residue. The results would indicate otherwise.

Importance

Development of subtype specific inhibition or modulation of the various sodium channels present in the human nervous system has proved to be a diffcult challenge, however attempts are still being made to understand these channels better to provide drug design opportunities that could target specific channels implicated in such pathological conditions seen in people with chronic pain or epilepsy, without shutting down the relatively similar channels present in the nerve cells of the heart, which would lead to a rapid and untimely death.

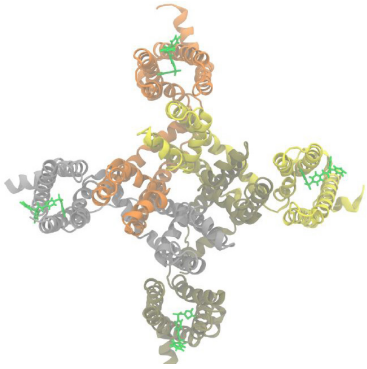
Method

To test the importance of the 1537 residue for selectivity of the various aryl sulfonamides compounds, a computational study was developed utilizing the FEP molecular dynamics method on the NAMD platform.
About 30 000 hours of computational time were required for the simulations if run on a modern 6 core CPU, or about 3.5 years. Fortunately, the supercomputer Raijin was utilized to complete these simulations in only days.

Results

I discovered that Tyrosine 1537 is actually a hindrance for all tested aryl sulfonamides binding in Nav1.7 unlike previously

thought, and that after mutation to Serine (A Nav1.3 subtype residue), all the drugs bound more easily, even the ones known to bind better to Nav1.7. This indicates that something else is going on, and some other residue candidates have been brought forward to be looked at further. This discovery is one piece in the puzzle of identifying the mechanism of subtype specific inhibition in sodium channels.



$Na_v1.7$ Protein, coloured by chain, in complex with aryl sulfonamides compounds shown in green, viewed from the top, or outer cell water environment. In the very center of the image, the ion tunnel, known to preferentially transfer sodium ions, is apparent as the intersection between all four chains..

PHYLOGENETIC ROOT FINDING USING CODON STRAND SYMMETRY

Thomas Hassall, University of Queensland
Supervisor: Gavin Huttley

Modern evolutionary taxonomy classifies organisms based on the evolutionary relationships between species.

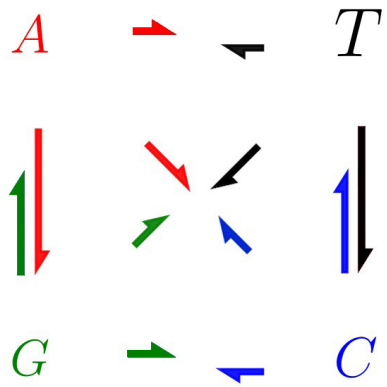
The construction of rooted phylogenetic trees – evolutionary trees that incorporate the concept of time flowing from a root – are necessary for determining the order of speciation events and the location of the last common ancestor of a clade; both of which are important results that inform taxonomic classification.

DNA substitution models offer a way to quantify evolutionary relationships. These models are mathematical processes that dictate how a DNA sequence is allowed to transform, via nucleotide substitutions, into another. Substitution models that have the mathematical property of time-irreversibility can be used to locate the root of phylogenetic trees to varying degrees of accuracy. The strand symmetric model is one such example.

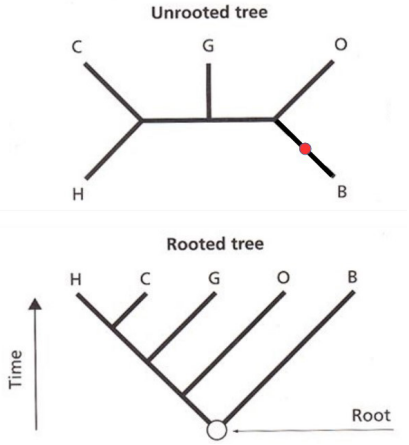
In addition to being time-irreversible, the strand symmetric model has also been proven to be statistically identifiable, i.e. the model's estimate of the root position converges to a single location, becoming more precise as more data becomes available. This is a desirable property of substitution models, and must be satisfied for precise inference to be possible.

It was hypothesised that by considering DNA substitutions of whole codons instead of single nucleotides, this strand symmetric model may potentially be improved to extract information about evolutionary events further back in time, and to greater temporal resolution than currently possible. Numerical experiments were run, with the help of the PyCogent software library, to investigate the plausibility of the strand symmetric codon model retaining the statistical identifiability property of the strand symmetric nucleotide model.

Preliminary findings using the codon model show promise in resolving the order of bifurcations that occurred in quick succession, even in cases where the nucleotide model could not. Additionally, the agreed upon outgroup of the human-chimpanzee-gorilla triad was correctly identified using mitochondrial DNA. This is unexpected as mitochondrial DNA often conflicts with the conventionally more reliable nuclear DNA results. If these results are successfully replicated and validated using more sequence data, this methodology may provide novel insight into other opaque evolutionary periods.



A transition graph of a nucleotide DNA substitution process. Arrow sizes represent the rate of simulated mutation between bases. Strand symmetric models (an example of which is pictured) have identical mutation rates between complementary bases (e.g. A→G = T→C, A→C = T→G, A→T = T→A, etc.). A codon substitution process has 61 (or 60 for vertebrate mtDNA) states instead of 4: one for each non-termination codon.



An unrooted phylogenetic tree (top) only considers the evolutionary distance between species, and cannot represent time in a meaningful sense. Using a time-irreversible substitution model allows for a root (red spot) to be positioned on the tree topology, permitting the formation of a rooted phylogenetic tree (bottom). Note that the rooted tree clearly shows that "B" is the outgroup of the "HCGO" clade; an observation not possible without the placement of the root.