



CONNECTIONS

Bioplatforms Australia Newsletter – March 2013



NEW DATASET SUPPORTS GREAT BARRIER REEF

Bioplatforms Australia will create large scale genomics datasets for coral and their symbionts to help scientists protect coral reefs from climate change.

Climate change is considered to be the greatest risk facing the Great Barrier Reef through warmer water temperatures and more acidic oceans. In response to this threat, leading researchers from Australia and Saudi Arabia are coming together to understand the genetic makeup of corals together with their response and resilience to climate changes.

The first phase of the program is the world first 'Sea-quence' project. Ten important Australian corals and their symbionts will be sequenced by Bioplatforms Australia facilities in collaboration with the coral research community. Globally, only two corals have ever been sequenced one of which, *Acropora millepora* was completed in 2011 by Australian researchers led by James Cook University. The DNA sequencing for this project was undertaken by the Bioplatforms Australia node, Australian Genome Research Facility (AGRF).

The Sea-quence project will considerably extend the work completed on corals to date and will be used to identify which genes help climate change adaptation and which species contain those genes.

Sea-quence also has a parallel element in Saudi Arabia where genomic data will be generated on Red Sea corals which experience higher water temperatures than those of the Great Barrier Reef. The comparative information between the two environments will enhance the understanding of how and why some corals are more resilient and hopefully promote new protection and preservation strategies.

The sequencing elements of the project have now commenced. In the last couple of months, project collaborators confirmed extraction and sequencing protocols and collected the first samples during the last spawning season. So far, DNA samples from five species situated around Orpheus and Heron Islands in the Great Barrier Reef have been collected and are now ready for sequencing. Researchers are also collecting microbial samples from each of these species which will also be sequenced by Bioplatforms Australia facilities.

In addition to Bioplatforms Australia, the project is proudly supported by Rio Tinto and ReFuGe 2020, a consortium of national and international experts established by the Great Barrier Reef Foundation. ReFuGe 2020 partners are investing resources and a range of expertise relevant to reef management and genomics to collect and analyse

coral samples and, ultimately, to fast-track genomics-based research into coral resilience and adaptation.

ReFuGe 2020 partners include the Great Barrier Reef Foundation, James Cook University, the Australian Institute of Marine Science, the University of Queensland, the Great Barrier Reef Marine Park Authority, the King Abdullah University of Science and Technology (Saudi Arabia) and the Australian National University.

The datasets created by Bioplatforms Australia will serve several ReFuGe 2020 projects and will also be made publically available for broader scientific research. The Sea-quence launch was covered by 7 News last November and can be viewed at www.bioplatforms.com.au/about-us/news/239-great-barrier-reef

PERSONALISED CANCER VACCINES

Researchers at the Queensland Institute of Medical Research (QIMR) are investigating peptide-based immunotherapy approaches for new personalised cancer treatments.

This important project aims to develop cancer vaccines that utilise a patient's own immune system to fight cancer and stop it from spreading through the body. The work is being supported by funds raised through the QIMR-Rio Tinto Ride to Conquer Cancer event together with Bioplatforms Australia.

The project combines next generation sequencing, state of the art mass spectrometry and classical immunological techniques for the development of novel predictive tools critical for personalised peptide-based cancer immunotherapies. The multidisciplinary approach employed in the research is also supported by collaboration across a number of QIMR research teams lead by Professor Jeffrey Gorman (Protein Discovery Centre), Professor Nick Hayward



Researchers are combining genomic technology with more traditional techniques to gain deeper insights into the evolution of living systems.

(Oncogenomics) and Dr Chris Schmidt (Cancer Immunotherapy).

In cancer cells, like other cells, HLA proteins display peptides on the cell's surface which allows the immune system to identify and respond to infections or cancer. If the specific peptides that signal an immune response can be identified these can then be used as a basis of a personalised vaccine.

The project will utilise samples from patients that were previously involved in clinical trials of a novel melanoma vaccine and demonstrated a positive immune response. The peptides present on the cells of these patients will be analysed using a state of the art Orbitrap Elite mass spectrometer located at Bioplatforms Australia's proteomic node at QIMR. In parallel, next generation sequencing information from these patients will be analysed using a range of algorithms to identify possible mutated peptides. Both datasets will be compared in order to identify which of the predicted mutated peptides was actually displayed by the patient's cells.

The peptides identified as being mutated due to the cancer and also presented by the HLA molecules of patient cells will be assayed using standard immunological techniques to determine their ability to elicit an immune response. This knowledge will in turn facilitate development of improved algorithms for predicting immunogenic peptides from DNA sequencing data.

Melanoma is being used as the model for this work, however, the results should be applicable to other spontaneous cancers that respond to immunotherapy. The project involves a novel tie-in between proteomics, genomics and classical immunology techniques in developing personalised cancer vaccines and will hopefully lead to outcomes of clinical significance.

Please contact **Prof. Jeff Gorman** at Jeffrey.Gorman@qimr.edu.au for further information.



GERANIUMS OFFER NEW INSIGHTS ON GENETIC EVOLUTION

The rapid advances and growing accessibility of next-generation sequencing technology are constantly being exploited to open new avenues of research.

In the fast emerging field of genome evolution and biodiversity, researchers are combining genomic technology with more traditional techniques to gain deeper insights into the evolution of living systems.

A research team lead by Justin Borevitz and Adrienne Nicotra at the Australian National University are incorporating the latest genomics technology as part of a multi-faceted investigation into the evolutionary lineages of Australian *Pelargonium*, the flowering plant most commonly known as the geranium.

This widespread plant can reveal much about evolution as there are many aspects regarding dispersal and habitat that have prompted questions about the process of evolution and species adaptation.

There are dozens of described entities of Australian *Pelargonium* which display significant diversity in their

physical features and a remarkable ability to flourish and adapt to a wide range of habitats. Given the huge range of morphological differences, Australian varieties are considered to be taxonomically complex and much controversy has surrounded attempts to classify the species on a morphological basis alone. Borevitz and Nicotra's research has combined morphological and reproductive data with DNA sequencing in order to define and compare different species. This genetic investigation has revealed that the Australian ancestry of *Pelargonium* is actually quite recent and may encompass only eight species.

Pinpointing evolutionary origins confirms the significant phenotypic plasticity of Australian *Pelargonium* already revealed by abundant physical variations and diverse growth habitats. In defining the actual genetic differences within and among species, the research team has been able to establish the lineage boundaries of species with highly diverse phenotypic traits.

Bioplatforms Australia investments in next-generation sequencing technology have been fundamental to this work and used by Borevitz and Nicotra to sequence and analyse portions of over 300 individual plant genomes. As a result, the research team has uncovered unique insights into the relationship between environmental variation and genetic factors that give rise to a new species. At the same time, the project is

also helping to resolve many arguments related to defining the individual lineage boundaries of Australian *Pelargonium*.

This project is also part of a broader effort in quantifying Australian biodiversity and is adding to emerging research on evolutionary history being generated by the new Centre of Biodiversity Analysis located at the ANU.

GENOMICS VIRTUAL LABORATORY

Bioplatforms Australia is co-investing in the development of a Genomics Virtual Laboratory (GVL) to further enable eResearch capabilities in Australia and facilitate broad access to Bioplatforms large-scale datasets.

The data-intensive nature of modern genomics research relies on large-scale computational infrastructure together with sophisticated software for analysis and visualisation. The GVL responds to the growing demands for data and bioinformatics capabilities and will provide a central resource of hardware, software and human expertise.

Researchers will be able to access best practice genomics tools and data repositories while also linking to Australia's eResearch and high computing infrastructure. Important new facilities such as the new NeCTAR Research Cloud, Research Data Storage Infrastructure, National Computational Infrastructure Specialised Facility in Bioinformatics (NCI-SFB), the Australian National Data Service and the EMBL Australia mirror of EMBL-EBI will all become much more accessible to Australian researchers and offer a new forum for collaboration. Likewise, the GVL can provide a vehicle for training and support and allow researchers to share data and workflows.



THE CHALLENGES OF BIG DATA

Bioplatforms Australia's Melanoma Initiative is a primary example of 'big data' - a computer science buzz word to describe large datasets of a size and complexity that need custom strategies for processing and interpretation.

The Melanoma Initiative represents an Australian born, large-scale genomics project that has drawn together a number of partners to investigate the common gene mutations that lead to melanoma. Bioplatforms Australia is funding sequencing elements of the project which will enable researchers to more deeply investigate the cause of melanoma cancer and pinpoint new drug targets.

At this stage of the project, DNA sequencing is well underway and is expected to include up to 1000 whole genome sequences from both tumours and normal tissue. Approximately 150 of these genomes are being supported by Bioplatforms Australia. The dataset being created will definitely be 'big' and with total storage needs of 200+ terabytes and computation time of 10 CPU years, the timely analysis of this data is a significant challenge.

Pipelines to manage and extract information from sequencing data are becoming compulsory components of large genomics projects. In this project, Matthew Field and his Genome Bioinformatics team at the John Curtin School of Medical Research, have built a bespoke bioinformatics system that can detect the DNA mutations present in melanoma tumours. Naturally, this work involved the development of a high throughput pipeline to allow researchers to extract meaningful information from the raw data.

An additional challenge with data of this size is to enable researchers to access relevant aspects of the melanoma genomes without having to store and manage the data themselves. Open-source solutions are emerging in this area such as the Galaxy genomics platform which can be customised for computational analysis allowing specifically defined and reproducible workflows. Similarly, the availability of high performance visualisation tools can allow researchers to remotely view large and diverse datasets.

For the Melanoma Initiative, Field's bioinformatics team has addressed data access through a Galaxy portal, hosted by the new Genomics Virtual Laboratory. Data files and genome alignments will be accessible via the web and will be visible with the IGV (Integrative Genomic Viewer) and UCSC (University of California Santa Cruz) genome browsers. Utilities to help summarise variation data



A new study aims to identify metabolite biomarkers that can be used in the early diagnosis of malaria.

and aid data mining will also be built into the system.

As modern genomics technologies become cheaper and more accessible, it is also likely that 'big data' will become the norm. The bioinformatics solutions needed for the analysis of melanoma genomes is a clear illustration of the growing computational requirements of laboratory-based research. The expertise and capabilities of Field's team will soon be core to large-scale genomic analyses given the consistent need for reproducible workflows; fast, robust data storage; and high performance computing.

METABOLOMICS IN BIOMARKER DISCOVERY

Metabolomics is increasingly becoming a powerful tool in identifying early diagnostic markers for infectious diseases as well understanding the physiological changes that can underlie disease pathology.

A new project that hopes to identify biomarkers associated with infections caused by the malarial parasite, *Plasmodium falciparum*, will measure changes in blood and urine metabolite levels of people who have been infected



with the parasite. The research will apply state of the art mass spectrometry to simultaneously monitor the changes in hundreds to thousands of metabolites that occur during the course of a malaria infection and subsequent drug treatments. Ultimately, these unique studies aim to identify metabolite biomarkers that can be used in the early diagnosis of malaria in individuals who might be exposed to many other infectious agents.

The research will also measure the effectiveness of current and new drug treatments and should shed light on the potential role of patient metabolism in determining the outcome of malaria infections.

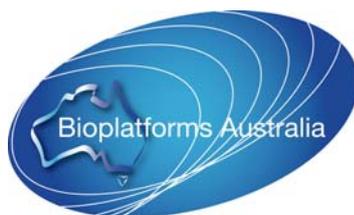
The research is being undertaken as a collaborative project between Bioplatforms Australia's metabolomics node at the University of Melbourne and malaria researchers at the Queensland Institute of Medical Research.

2012 ANNUAL REPORT

Download our 2012 Annual Report at www.bioplatforms.com.au or request a copy from Anita Tymkiw 0403 889 796.

BIOPPLATFORMS AUSTRALIA

2012 ANNUAL REPORT



Bioplatforms Australia Ltd

Level 4, Building F7B Research Park Drive, Macquarie University NSW 2109 Australia
P: +61 410 538 648
F: +61 2 9850 6200
www.bioplatforms.com

