2016 AGTA Conference

Australasian Genomic Technologies Association

held in partnership with **the 8th Annual New Zealand Next Generation Sequencing Conference**

HANDBOOK

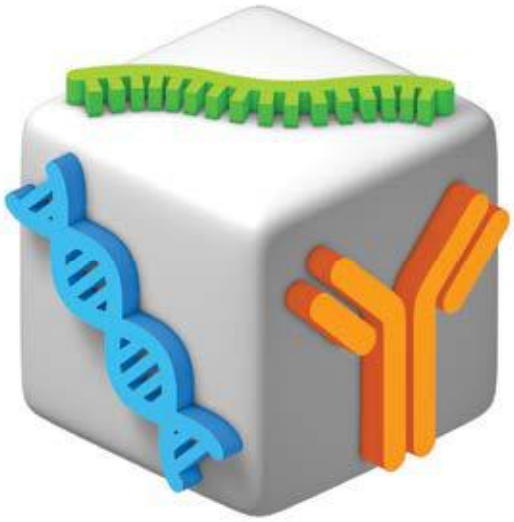
9–12 October Pullman Hotel, Auckland, New Zealand

www.agtaconference.org

Molecules That Count®

Translational Research • Gene Expression • miRNA Expression • Protein Detection • Copy Number Variation

nanoString TECHNOLOGIES



3D BIOLOGY™

DNA - RNA - Protein
All digital All at once

Hyb & Seq™ An All-New Single-Molecule NGS Platform

Come along to our talk and learn about **Hyb & Seq™**, an amazing new technology:
No enzymes, *no amplification* and *no library prep*, plus *the highest accuracy ever achieved.*
Sequencing will never be the same!

WHEN: Monday 4:40 – 5:00 PM, session 3 "New Technologies"

WHAT: ***Enzyme-Free, Amplification-Free, Hybridization Based Single Molecule Sequencing Technology Using Fluorescent Optical Barcodes - A First-In-Class Chemistry with Several Unique Features***

WHO: Presented by Dr Michael Rhodes, Director of Advanced Applications & Sequencing Commercialisation, NanoString Technologies, Seattle USA

bio-strategy
delivering technology

Visit us at BOOTH 1 for more information!

www.bio-strategy.com

CONTENTS

AGTA EXECUTIVE TEAM

Dr Mark Waltham (President)
University of Melbourne

Dr Carsten Kulheim (Vice President)
Australian National University

Dr Ruby C Y Lin (Convenor)
Asbestos Diseases Research Institute

Mr Mark van der Hoek (Treasurer)
South Australian Health and Medical Research Institute

Ms Vikki Marshall (Secretary/Convenor)
Melbourne Neuroscience Institute, University of Melbourne

Dr Alicia Oshlack
Murdoch Children's Research Institute

Dr Richard Tothill
Peter MacCallum Cancer Centre

Professor Erik (Rik) Thompson (Founding AMATA President)
Queensland Institute of Technology

Dr Jac Charlesworth
Menzies Institute for Medical Research, University of Tasmania

Professor Ryan Lister
The University of Western Australia

Mr Liam Williams (Convenor)
Auckland University

Associate Professor Marcel Dinger
The Garvan Institute of Medical Research

Dr Kate Howell
The University of Western Australia

Dr Kirby Siemering
Australian Genome Research Facility

Dr Andreas Schreiber (Membership Secretary/Vice-Treasurer)
Centre for Cancer Biology

Dr Nicole Cloonan (Membership Secretary)
QIMR Berghofer Medical Research Institute

Dr Helen Speirs
University of New South Wales

Dr Robert Day (Convenor)
University of Otago

AGTA/NZ NGS 2016 CONFERENCE CONVENORS

Dr Robert Day
University of Otago

Associate Professor Ruby Lin
Asbestos Diseases Research Institute

Ms Vikki Marshall
University of Melbourne

Dr Donia Macartney-Coxson
The Institute of Environmental Science and Research

Professor Cristin Print
University of Auckland

Dr Jo-Ann Stanton
University of Otago

Mr Liam Williams
University of Auckland

CONFERENCE MANAGERS

Leishman
it's all about people
CONFERENCE • EVENT • ASSOCIATION MANAGEMENT

Leishman Associates

113 Harrington Street,
Hobart TAS 7000

170 Elgin Street,
Carlton VIC 3053

P. 03 6234 7844 | F. 03 6234 5958
E. emma@leishman-associates.com.au
W. www.leishman-associates.com.au

WELCOME FROM THE AGTA CONFERENCE CONVENORS

On behalf of the Executive Committee and the 2016 Organising Committee, we are delighted to welcome you to the 16th annual conference of the Australasian Genomic Technologies Association (AGTA) and the 8th Annual New Zealand Next Generation Sequencing Conference (NZ NGS), at the Pullman Hotel, Auckland, New Zealand.

The AGTA meeting crosses the Tasman for the second time, and joins with the NZ NGS. We believe that holding the AGTA meeting alongside the NZ NGS will provide the best possibility for interactions between a wide cross section of Australasian researchers, industries and disciplines.

AGTA/NZ NGS 2016 attracts staff and clients from all genomics core service providers in the region, and beyond, eager to be exposed to new and evolving technologies, and to hear how those technologies are being applied to genomics research.

AGTA/NZ NGS 2016 will showcase dynamic cross-discipline-omics research, and uniquely allow genomics and bioinformatics researchers, as well as industry, to interact over a wide range of topics; these will include technologies focused on genome regulation: biology, phenotype and utility, genomics and agriculture, genome diversity and evolution, human disease and clinical genomics, plant genomics, microbial and single cell genomics and transcriptomics. We hope you will enjoy the exciting line-up of topics and presenters we have assembled.

The session organisers have invited scientists at the cutting edge of integrative -omics research to present keynote lectures at AGTA/NZ NGS 2016. These amazing high profile speakers will be joining us in Auckland 2016. They include Christopher Mason (Cornell University), founder of Genome Liberty (crowd funded genomics) and leader on many groundbreaking -omics projects including space missions, cancer, brain development, undiagnosed genetic disorders and the metagenomics of New York City. Gregory Gibson (Georgia Tech), Director of the Center for Integrative Genomics and Lance Millar (Wake Forest School of Medicine), Director of the Comprehensive Cancer Center.

The Pullman Hotel is located in central Auckland, close to transport links, the CBD, waterfront entertainment areas, and the university precinct. It offers excellent conference and social facilities.

Welcome to our conference!

Dr Robert Day, Associate Professor Ruby Lin, Ms Vikki Marshall,
Dr Donia Macartney-Coxson, Professor Cristin Print, Dr Jo-Ann Stanton, Mr Liam Williams

Conference Convenors

WELCOME FROM THE NZ NGS CHAIR

Dear Delegates and Invited Guests

Welcome to the 8th New Zealand Next Generation Sequencing and 16th Australasian Genomic Technology Association Conference, 2016.

Recently I have been asked whether genome technology-focused conferences still have relevance. An understandable question, particularly given the wide use of genomic technology in fields as disparate as mining and health care coupled with many genomic technologies having been formulated for use by non-experts. But a good example of why technology meetings have relevance is clearly demonstrated by the nine years since Next GenerationSequencing burst onto the New Zealand science scene (remember when 100Mb in a run was amazing?). Just as we thought NGS was going mainstream a new platform would appear that produced sequence information in a completely different way opening access to new questions and new uses for the information provided by genomics. On a personal level my work is currently focused on resource-challenged regions of the world. I have seen how genomic technology has the potential to significantly improve health, wellbeing and the environment globally. This will only happen, however, if we continue to innovate and develop new genomic technologies. So, yes, I am convinced the relevance of technology-focused meetings remains strong.

It has been an extremely positive experience working with the AGTA team to bring about this year's conference. Combining resources has strengthened our 2016 offering. This is reflected in the quality of the program, the speakers and posters that will be presented over the next three days. We hope you will enjoy the opportunity to network with old colleagues and new-found friends.

Thank you to the sponsors, particularly our medal-level sponsors. Without their generous support we would not be able to present such a high quality program or invite our distinguished international speakers to share their research with us.

We hope you enjoy the meeting and, for our international colleagues, the opportunity to visit and explore New Zealand.

Jo Stanton
Chair, NZ NGS

GENERAL INFORMATION

REGISTRATION DESK

Please direct any questions you may have regarding registration, accommodation or social functions to Leishman Associates staff at this desk.

Registration Desk Opening Times:

Sunday 9 October	2.00pm – 6.00pm
Monday 10 October	7.30am – 5.30pm
Tuesday 11 October	7.30am – 5.30pm
Wednesday 12 October	7.30am – 2.30pm

ACCOMMODATION

If you have any queries relating to your accommodation booking first speak to the staff at your hotel or alternatively Leishman Associates staff at the Registration Desk.

Your credit card details were supplied to the hotel you have selected, as security for your booking. If you have arrived 24 hours later than your indicated arrival day you may find that you have been charged a fee. You will be responsible for all room and incidental charges on check out and may be asked for an impression of your credit card for security against these charges. This is standard policy in many hotels.

DELEGATE NAME BADGES

All delegates, speakers, sponsors and exhibitors will be provided with a name badge, which must be worn at all times within the conference venue as it is required for access to all the conference sessions and social functions.

With thanks to our Name Badge Sponsor:

illumina®

CONFERENCE PROCEEDINGS

Speaker PowerPoints and abstracts will be available on the AGTA website following the conclusion of the conference. Speakers will be requested to sign a release form. This is not compulsory.

CONFERENCE WIFI

Wireless internet will be available throughout the conference venue for the duration of the conference.

Username: The Pullman Auckland
Password: AGTA16

TWITTER

Join the conversation at #AGTA16

DRESS CODE

Dress throughout the day is smart casual or informal business.

EMERGENCY MEDICAL CARE

For any medical emergency please telephone 111. The staff at your hotel will have information if you require contact details for a doctor, dentist or other health professional.

GENERAL INFORMATION

EXHIBITOR PRIZE DRAW

An exhibitor passport will be given to all delegates at registration. The AGTA/NZ NGS 2016 organising committee encourages you to visit each trade exhibitor and have your passport stamped, to go into the draw to win some great prizes!

SATCHEL

Each attendee will receive a satchel at registration.

Proudly sponsored by

bio-strategy
delivering technology

BARISTA COFFEE

Barista style coffee will be available during the refreshment breaks.

Proudly sponsored by

ThermoFisher
SCIENTIFIC

STUDENT FUNCTION

All conference students and early career researchers are invited to a casual function on Monday 10 October. Please forward enquiries to the registration desk.

Further information about this event can be found on page 21.

SOCIAL PROGRAM ENTRY

The Welcome Reception is included in the cost of each full conference registration.

The Conference Dinner IS NOT included in any registration type. Social events ARE NOT included in the cost of day registrations or for accompanying partners. Places for day registrants and additional guests for these events may still be available at an additional cost. Bookings can be made at the registration desk subject to availability.

All delegates who are registered to attend the dinner will receive a named sticker at registration. You MUST place your sticker on a table located on poster boards next to the registration desk. You must allocate yourself to a table no later than 11.00 am Tuesday 11 October.

PHOTOGRAPHS, VIDEOS, RECORDING OF SESSIONS

Delegates are not permitted to use any type of camera or recording device at any of the sessions unless written permission has been obtained from the relevant speaker.

ORAL PRESENTATIONS

Please refer to the program for the time allocated for each presentation, as these do vary. The chairperson for your session will give you a 2 minute warning, however you are asked to stick to your time allocation so that the program remains on schedule.

POSTER PRESENTATIONS

Posters will be displayed in the Exhibition Centre and the Gallery for the duration of the conference. There will be a poster session on Monday 10 October from 2.30 to 3.15pm and on Tuesday 11 October from 2.15pm to 3.00pm

GENERAL INFORMATION

SPEAKERS AND SPEAKER'S PREPARATION ROOM

All speakers should present themselves to the Speaker's Preparation Room, located in Princes Lounge B at least 4 hours before their scheduled presentation time, to upload their presentation.

Speakers are requested to assemble in their session room 5 minutes before the commencement of their session, to meet with their session chair and to familiarise themselves with the room and the audio visual equipment. For information on the chairperson attending your session, please see the program.

A technician will be present in the speaker's preparation room at the following times:

Sunday:	4:00pm – 7:00pm
Monday:	7:30am – 4:00pm
Tuesday:	7:30am – 4:00pm
Wednesday:	8:00am – 2:00pm

There will be facility to test and modify your presentation as required.

SPECIAL DIETS

All catering venues have been advised of any special diet preferences you have indicated on your registration form. Please identify yourself to venue staff as they come to serve you and they will be pleased to provide you with all pre-ordered food. For day catering, there may be a specific area where special food is brought out, please check with catering or conference staff.

SECURITY

The members of the conference organising committee, Leishman Associates and The Pullman Hotel accept no liability for personal accident or loss or damage suffered by any participant, accompanying person, invited observer or any other person by whatever means. Nor do we accept liability for any equipment or software brought to the conference by delegates, speakers, sponsors or any other party.

Please protect your personal property. Do not leave laptops, cameras, and other valuable items unsecured. Be conscious of individuals who appear out of place and do not wear a conference name badge. Advise Leishman Associates staff if this does occur.

DISCLAIMER

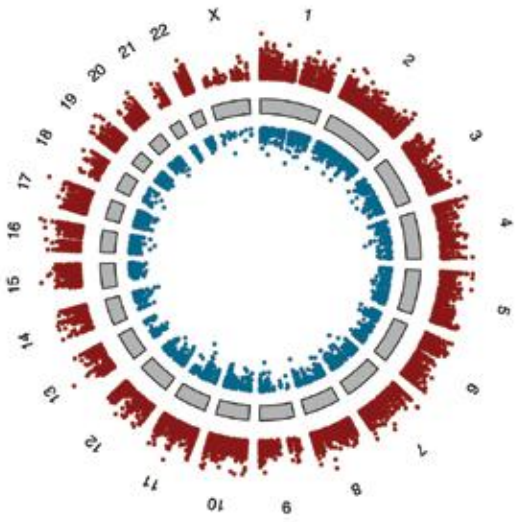
The 2016 AGTA/NZ NGS Conference reserves the right to amend or alter any advertised details relating to dates, program and speakers if necessary, without notice, as a result of circumstances beyond their control. All attempts have been made to keep any changes to an absolute minimum.

CHANGING THE FACE OF SCIENCE!

From research and development through to service delivery, ESR is involved in a broad range of disciplines including forensic, social science, radiological, environmental, microbial and epidemiological expertise. We deliver solutions for local and central government, industry and other science organisations.

Genomics at ESR includes: service delivery for microbial surveillance and outbreak investigations; utilising metagenomics for pathogen discovery and biological source tracking, and research into both microbial evolution and the role of epigenetic mechanisms in complex disease.

In the forensic setting our research teams are working on the introduction of amplicon sequencing for DNA identification and the use of RNA as a cell-type specific marker.



University of Otago Te Whare Wānanga o Otāgo New Zealand

Massey University Te Kunenga ki Pūrehuroa University of New Zealand

The University of Auckland New Zealand Te Whare Wānanga o Tāmaki Makaurau

Ministry of Business, Innovation & Employment Hikina Whakatutuki

CONFERENCE PROGRAM

SUNDAY 9 OCTOBER 2016

Old Government House, Auckland (OGH), University of Auckland.

PRE-CONFERENCE WORKSHOPS

Sponsored by: illumina®

	WORKSHOP ONE	WORKSHOP TWO
1000 - 1630	**Bisulfite Sequencing: Genomewide methylation sequencing** *Sponsored by:* NZGL NEW ZEALAND GENOMICS LIMITED	**Metagenomics: Understanding microbial communities with next-generation sequencing**

Pullman Hotel Auckland

1400 - 1800	Conference Registration Open

OPENING ORATION

CHAIRS: DR MARK WALTHAM & DR JO-ANN STANTON

1700 - 1800	**Associate Professor Christopher Mason** Cornell University, USA	●
1800 - 2000	Welcome Reception & Trade Exhibition	

SPEAKER KEY

- ● Guest Speaker
- ● Invited Speaker
- ● Sponsored Speaker
- ● Keynote Speaker
- ● Invited Abstract
- ● Early Career Researcher

MONDAY 10 OCTOBER 2016

Pullman Hotel Auckland

Today's Refreshment Breaks & Lunch sponsored by: NZGL

0730 - 1730	Conference Registration Open
0800 - 1730	Trade Exhibition Open *Barista Coffee Served, sponsored by:* ThermoFisher SCIENTIFIC
0830 - 0845	**Official Welcome and Conference Opening** **Mr Liam Williams** University of Auckland, New Zealand **Dr Donia Macartney-Coxson** The Institute of Environmental Science and Research, New Zealand

SESSION 1: ANIMAL & PLANT GENOMICS #1

CHAIRS: DR SHANNON CLARKE & DR PHILLIP WILCOX

Sponsored by: beef+lamb GENETICS

0845 - 0930	**FAANG INITIATIVE: IDENTIFICATION OF REGULATORY ELEMENTS IN LIVESTOCK SPECIES** **Associate Professor Huaijun Zhou** University of California, USA
0930 - 1000	**THE EXTENT OF CIS-REGULATION OF GENE EXPRESSION AND ITS INFLUENCE ON COMPLEX TRAIT VARIATION IN CATTLE** **Dr Amanda Chamberlain** Department of Economic Development, Jobs, Transport and Resources, Australia
1000 - 1030	**PNEUMONIA, PURKINJE CELLS AND PITHOMYCES: GENOMICS AND ANIMAL HEALTH IN SHEEP** **Dr Kathryn McRae** AgResearch, New Zealand
1030 - 1045	**SECRET IDENTITY OF THE SUPERPRAWN: NOVEL TRANSCRIPTOME ASSEMBLY AND ANNOTATION OF THE BLACK TIGER PRAWN (PENAEUS MONODON)** **Dr Roger Huerlimann** James Cook University, Australia

Pullman Hotel Auckland

1045 - 1100	**EXPRESSION QUANTITATIVE TRAIT LOCI MAPPING IN BOVINE MILK CELLS** **Dr Christy Vander Jagt** Agriculture Victoria, Australia
1100 - 1130	Morning Refreshments & Trade Exhibition

SESSION 2: ANIMAL & PLANT GENOMICS #2

CHAIRS: PROFESSOR ROGER HELLENS & DR ELENA HILARIO

1130 - 1215	**RADIATA PINE: A LARGE GENOME, GENE EDITING AND LEGISLATION EXPERIENCES** **Dr Elspeth MacRae** Scion, New Zealand
1215 - 1245	**WHERE ARE ALL THE SMALL PEPTIDES AND WHAT MIGHT THEY DO?** **Professor Roger Hellens** Queensland University of Technology, Australia
1245 - 1315	**DOES LENGTH MATTER?** **Professor Dave Edwards** University of Western Australia, Australia
1315 - 1330	**JUMPING ON THE EPIGENETICS BANDWAGON THROUGH HYPOMETHYLOME SEQUENCING** **Natalie Graham** Scion, New Zealand
1330 - 1345	**ANALYSIS OF GENOME-WIDE INHERITANCE PATTERNS USING RAD SEQUENCING IN BIVALVE SPECIES** **Carolina Penaloza** Roslin Institute, United Kingdom
1345 - 1430	Lunch & Trade Exhibition
1430 - 1515	Poster Session One & Afternoon Refreshments

SESSION 3: NEW TECHNOLOGIES

CHAIRS: DR JO-ANN STANTON & MR PETER SMITH

MONDAY 10 OCTOBER 2016

Pullman Hotel Auckland

1515 - 1600	**Associate Professor Christopher Mason** Cornell University, USA
1600 - 1620	**A FLEXIBLE SEQUENCE ANALYSIS PIPELINE PLATFORM FOR NON-BIOINFORMATICIANS** **Cameron Jack** The Australian National University, Australia
1620 - 1640	**SWIMMING IN CIRCLES** **Dr Katherine Pillman** University of South Australia, Australia
1640 - 1700	**ENZYME-FREE, AMPLIFICATION-FREE, HYBRIDIZATION-BASED SINGLE MOLECULE SEQUENCING TECHNOLOGY USING FLUORESCENT OPTICAL BARCODES: A FIRST-IN-CLASS CHEMISTRY WITH SEVERAL UNIQUE FEATURES** **Dr Michael Rhodes** NanoString Technologies, USA. *Sponsored by:* nanoString TECHNOLOGIES
1700 - 1715	**CHROMATIN-AWARE CLOUDBASED CRISPR TARGET SITE PREDICTOR** **Dr Denis Bauer** CSIRO, Australia
1715 - 1730	**THE EXTREME MICROBIOME PROJECT PRESENTS: METAGENOMICS ON THE MINION-NANOPORE SEQUENCING OF ENVIRONMENTAL MICROBIAL COMMUNITIES** **Dr Ken McGrath** AGRF, Australia
1730	Free evening for delegates

Old Government House, Auckland

1830 - onwards	**Student Function** *Students and early career researchers are invited to join fellow delegates for a fun night of genomics trivia. Finger food and limited beverage service will be provided. Bookings are essential.*

TUESDAY 11 OCTOBER 2016

Pullman Hotel Auckland

Today's Refreshment Breaks & Lunch sponsored by: Agilent

0730 - 1730	Conference Registration Open
0800 - 1730	Trade Exhibition Open *Barista Coffee Served, sponsored by:* ThermoFisher SCIENTIFIC
0825 - 0830	**Welcome to Day Two** **Dr Rob Day** University of Otago, New Zealand **Dr Ruby Lin** Asbestos Diseases Research Institute, Australia

SESSION 4: HUMAN DISEASES & CLINICAL GENOMICS

CHAIRS: DR LOGAN WALKER & ASSOCIATE PROFESSOR RUBY CY LIN

0830 - 0915	**UNDERSTANDING THE INTERPLAY BETWEEN TUMOR-IMMUNE INTERACTIONS AND PATIENT OUTCOMES THROUGH GENOMIC CORRELATES** **Associate Professor Lance Miller** Wake Forest School of Medicine, USA
0915 - 0945	**INSIGHTS FROM A GENETIC ISOLATE** **Dr Miles Benton** Institute of Health and Biomedical Innovation, Queensland University of Technology, Australia
0945 - 1015	**THE POWER OF A SINGLE PATIENT AND A SEQUENCING MACHINE: GENETIC INSIGHTS INTO HUMAN GROWTH** **Dr Louise Bicknell** University of Otago, New Zealand
1015 - 1030	**UNRAVELLING THE MULTIPLE SCLEROSIS COMPLEX DISEASE TRAIT THROUGH AN IMMUNE TRANSCRIPTIONAL REGULATORY NETWORK APPROACH** **Dr Margaret Jordan** James Cook University, Australia

Pullman Hotel Auckland

1030 - 1045	**IDENTIFICATION OF THREE NOVEL ENHANCERS THAT REGULATE HUMAN SOX9 IN THE GONAD: IMPLICATIONS FOR DISORDERS OF SEX DEVELOPMENT** **Dr Thomas Ohnesorg** Murdoch Childrens Research Institute, Australia
1045 - 1115	Morning Refreshments & Trade Exhibition

SESSION 5: GENOME REGULATION: BIOLOGY, PHENOTYPE AND UTILITY

CHAIRS: DR MILES BENTON & DR DONIA MACARTNEY-COXSON

1115 - 1200	**TRANSCRIPTOMICS FOR PERSONALIZED MEDICINE** **Professor Gregory Gibson** Georgia Institute of Technology, USA *Sponsored by:*
1200 - 1230	**UTILISING GENOMIC TECHNOLOGIES TO EVALUATE GERMLINE VARIANTS OF UNKNOWN CLINICAL SIGNIFICANCE** **Dr Logan Walker** University of Otago, New Zealand
1230 - 1300	**EXTRACELLULAR VESICLES – THEIR ROLE IN NEURODEGENERATIVE DISEASES AND POTENTIAL SOURCE OF RNA DISEASE BIOMARKERS** **Professor Andrew Hill** La Trobe Institute for Molecular Science, La Trobe University, Australia
1300 - 1315	**MECHANISMS AND REGULATION OF MRNA TRANSLATION INITIATION REVEALED BY TCP-SEQ** **Dr Stuart Archer** Monash Bioinformatics Platform, Australia
1315 - 1330	**CHARACTERISING EPIGENOME DYNAMICS DURING THE REPROGRAMMING OF SOMATIC CELLS TO IPS CELLS** **Dr Sam Buckberry** The University of Western Australia, Australia

Pullman Hotel Auckland

1330 - 1415	Lunch & Trade Exhibition
1415 - 1500	Poster Session Two & Afternoon Refreshments

SESSION 6: GENOMIC DATA INVESTIGATION

CHAIRS: DR AUSTEN GANLEY & DR GRETCHEN POORTINGA

1500 - 1545	**EMERGING LANDSCAPE OF REGULATORY RNAs** **Professor Piero Carninci** RIKEN Center for Life Science Technologies, Japan *Sponsored by:* dnature • DETECT • DISCOVER • GeneTargetSolutions
1545 - 1615	**USING THE STEMFORMATICS STEM CELL ATLAS TO MINE STABLE SIGNATURES OF STEM CELL IDENTITY** **Professor Christine Wells** The University of Melbourne, Australia
1615 - 1645	**THE GENOMIC RESPONSES TO POLYPLOIDY IN EUKARYOTES** **Dr Austen Ganley** University of Auckland, New Zealand
1645 - 1700	**COMPREHENSIVE MACHINE LEARNING ANNOTATION OF HUMAN SPLICING BRANCHPOINTS** **Beth Signal** Garvan Institute, Australia
1700 - 1715	**SYSTEMATIC EVALUATION OF GENETIC CORRELATIONS BETWEEN EXPRESSED TRANSCRIPTS IN PERIPHERAL BLOOD** **Dr Joseph Powell** The University of Queensland, Australia
1715 – 1730	**NEXT GENERATION MAPPING REVEALS NOVEL LARGE GENOMIC REARRANGEMENTS IN PROSTATE CANCER** **Professor Vanessa Hayes** Garvan Institute of Medical Research, Australia
1730 - 1815	AGTA Annual General Meeting

The Northern Club, 19 Princes St, Auckland

1900 - 2300	**Conference Dinner** *Optional function. Bookings are essential.*

Pullman Hotel Auckland

Today's Refreshment Breaks & Lunch sponsored by: dnature • DETECT • DISCOVER • GeneTargetSolutions

0730 - 1730	Conference Registration Open
0800 - 1530	TTrade Exhibition Open *Barista Coffee Served, sponsored by:* ThermoFisher SCIENTIFIC
0855 - 0900	**Welcome to Day Three** **Ms Vikki Marshall** University of Melbourne, Australia

SESSION 7: POPULATION GENETICS

CHAIRS: PROFESSOR NEIL GEMMEL & CATHERINE COLLINS

0900 - 0945	**PASSENGER PIGEON PALEOGENOMES REVEAL THE GENOMIC CONSEQUENCES OF LONG-TERM EXTREMELY LARGE EFFECTIVE POPULATION SIZES** **Associate Professor Beth Shapiro** University of California Santa Cruz, USA *Sponsored by:* geneious
0945 - 1015	**GENOMIC BASIS OF TOOL MANUFACTURE AND USE IN NEW CALEDONIAN CROWS** **Dr Nicolas Dussex** University of Otago, New Zealand
1015 - 1045	**RECONSTRUCTING THE SETTLEMENT OF EAST POLYNESIA** **Professor Lisa Matisoo-Smith** University of Otago, New Zealand
1045 - 1100	**FAMILY VALUES: SEQUENCING LARGE FAMILIES AS AN ENRICHMENT STRATEGY FOR RARE VARIANTS** **Jac Charlesworth** Menzies Institute For Medical Research, Australia
1100 - 1115	**RNA SEQUENCING IN FORENSIC SCIENCE** **Dr Rachel Fleming** Institute of Environmental Science and Research Ltd, New Zealand

Pullman Hotel Auckland

1115 - 1145	Early lunch & Trade Exhibition

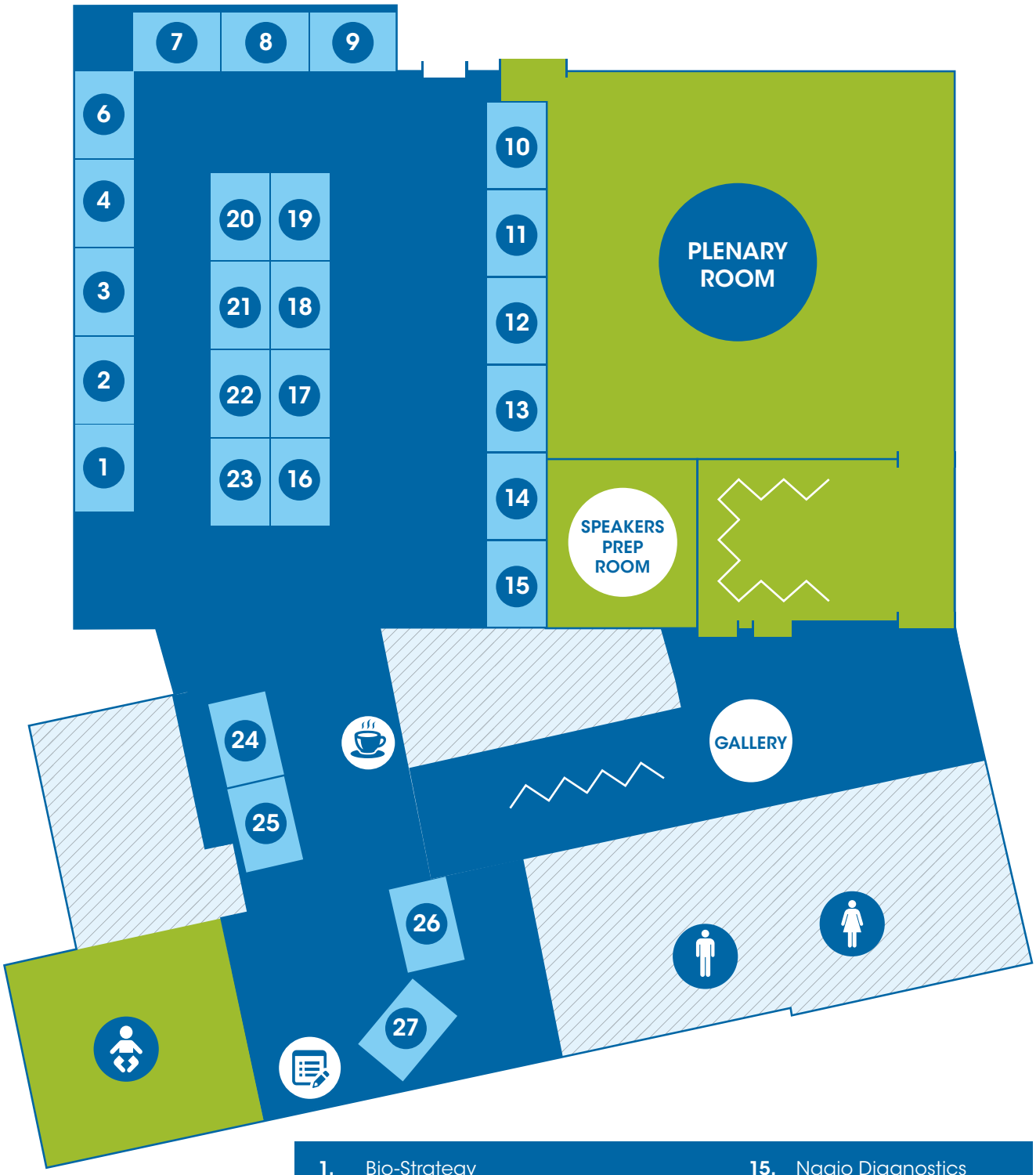
SESSION 8: GENOME DIVERSITY & EVOLUTION

CHAIRS: PROFESSOR LISA MATISOO-SMITH & DR NIC DUSSEX

1145 - 1230	**USING PROXIMITY LIGATION DATA FOR EFFICIENT DE NOVO GENOME ASSEMBLY** **Dr Richard Edward Green** University of California Santa Cruz, USA *Sponsored by:* NZGL NEW ZEALAND GENOMICS LIMITED
1230 - 1300	**GENOMIC INSIGHTS INTO CANCER AND DIVERSITY IN TASMANIAN DEVILS: LESSONS FOR CONSERVATION** **Dr Catherine Grueber** The University of Sydney, Australia
1300 - 1330	**THE TUATARA GENOME PROJECT — UNLOCKING THE GENOME OF A LIVING FOSSIL** **Professor Neil Gemmel** University of Otago, New Zealand
1330 - 1345	**SHIGATOXIN TYPE 1 IS LIKELY TO BE VERTICALLY TRANSMITTED IN WAIKATO DAIRY FARMS** **Paul Maclean** Agresearch Ltd NZ, New Zealand
1345 - 1400	**BREAKING THE CURSE OF DIMENSIONALITY FOR MACHINE LEARNING ON GENOMIC DATA** **Adrian O'Brien** CSIRO, Australia
1400 - 1420	Awarding of Prizes, 2017 Conference Launch and Conference Close

The 2016 AGTA / NZ NGS Conference reserves the right to amend or alter any advertised details relating to dates, program and speakers if necessary and without notice, as a result of circumstances beyond their control. All attempts will be made to keep any changes to an absolute minimum.

CONFERENCE FLOOR PLAN



- Registration Desk
- Barista Cart
- Toilets
- Parents Room
- Poster Boards

1. Bio-Strategy
2. Macrogen Oceania
3. Qiagen
4. Partek
6. Ramaciotti Centre for Genomics
7. Kinghorn Centre for Clinical Genomics
8. medi'Ray
9. QFAB Bioinformatics
10. BGI: Your Science. Our Solution
11. Tecan Australia
12. dnature & Gene Target Solutions
13. Millennium Science
14. AGRF
15. Ngaio Diagnostics
16. Illumina
17. Roche Diagnostics
18. PerkinElmer
19. Agilent Technologies
20. Genesearch
21. Biomatters
22. Agena Bioscience Inc
23. NZ Genomics Ltd
24. Thermo Fisher Scientific
25. AGTA
26. ZyGEM
27. Custom Science

SOCIAL PROGRAM

WELCOME RECEPTION & MEET THE EXHIBITORS

Date: Sunday 9 October 2016
Venue: Exhibition Centre, Pullman Hotel, Auckland
Time: 6.00pm – 8.00pm
Dress: Smart Casual

Join us for the official Welcome Reception for the 2016 AGTA/NZ NGS Conference. Enjoy networking with old and new acquaintances, and familiarising yourself with the trade exhibitors, whilst enjoying drinks and canapés.

The Welcome Reception is included in a full registration only. Additional tickets can be purchased at $70.00 per person.

STUDENT FUNCTION

Date: Monday 10 October 2016
Venue: VC Suite, The Northern Club, Auckland
Time: 6.00pm onwards

Students and early career researchers are invited to join fellow delegates for a fun night of genomics trivia. Test your knowledge of genomics technologies, bioinformatics tools, the history of sequencing and much more! As in real life genomics, your team will have a diverse range of backgrounds, so will have to work together to win the prize! Teams will form on the night, and points will be issued for both correct and creative answers.

Finger food and limited beverage service will be provided. Bookings are essential.

CONFERENCE DINNER

Date: Tuesday 11 October 2016
Venue: The Wintergarden, The Northern Club, Auckland
Time: 7.00pm – 11.00pm
Dress: Smart Casual
Cost: $130 per ticket. Tickets are not included in any registration type.

Registration for the Conference Dinner could be made through the online registration form. Registration can still be made during the Conference, subject to availability.

The conference dinner is the social highlight of the conference and should not be missed.

Come and join us for another chance to network and meet with colleagues, whilst enjoying a great night of food, wine and dancing.

IMPORTANT INFORMATION
Seating and table allocation for the AGTA Dinner will be by way of sticker allocation. All delegates registered to attend the Conference Dinner will receive a sticker to be placed on the table allocation sheets near the Registration Desk. These sheets will be available from Sunday 9 October and will be taken down at the end of morning refreshments on Tuesday 11 October or as they become full.

If you do not have a sticker please see the Registration Desk staff, DO NOT write your name directly on the board, as you will NOT be allocated a seat.

ion torrent

The S is for Simplicity

The new Ion S5™ System. Targeted sequencing has never been simpler.

Simple library prep tools, cartridge-based reagents and automated data analysis have reduced DNA-to-data hands-on time to less than 45 minutes. So you'll spend less time doing routine molecular biology, and more time informing time-sensitive decisions.

Ion AmpliSeq™ technology

As low as 10ng low-quality DNA sample input for library prep

Cartridge-based reagents

Less than 15 minutes of sequencing setup time

2.5 to 4 hours of run time

Fastest run time of any benchtop sequencer

Watch the Ion S5 System in action at
thermofisher.com/ionS5

ThermoFisher SCIENTIFIC

PRE-CONFERENCE WORKSHOPS

WORKSHOP 1

BISULFITE SEQUENCING: GENOMEWIDE METHYLATION SEQUENCING

This workshop is brought to you by New Zealand Genomics Limited

Date: Sunday 9 October
Venue: VC Suite, Old Government House
Time: 10.00am – 4.30pm

This is an optional workshop at $55.00 per person, which includes lunch and afternoon refreshments. Bookings are essential.

This hands-on workshop will cover:

- Aspects of library preparation, particularly for RRBS
- The consequences of bisulphite treatment on read quality
- Downstream read processing (adaptor trimming, quality trimming)
- Information on mapping strategies for BS reads
- Examples of work mapping RRBS reads
- Hands on work mapping human RRBS and WGBS data
- Differential methylation analysis from mapped data including DMAP (Differential Methylation Analysis Program)
- Visualisation of mapped data using IGV

Registrants will require their own laptop (64bit with 4Gb RAM or more)

This is entirely optional and is NOT necessary to take to part in the session.

WORKSHOP 2

METAGENOMICS: UNDERSTANDING MICROBIAL COMMUNITIES WITH NEXT-GENERATION SEQUENCING

Date: Sunday 9 October
Venue: Upstairs Lecture Room, Old Government House
Time: 10.00am – 4.30pm

This is an optional workshop at $55.00 per person, which includes lunch and afternoon refreshments. Bookings are essential.

Next generation sequencing has enabled detailed analysis of microbial communities and uncultured microorganisms at far greater depth that was previously possible with other methods.

Metagenomics analysis is now an integral part of medical, agricultural and environmental research.

This workshop will highlight different experimental and analytical approaches used in metagenomics studies.

Leading experts in the field will present their research project and will cover various methodologies including amplicon-based profiling of microbial communities, shotgun metagenomics and metatranscriptomics.

Pre-conference workshops proudly sponsored by Illumina

ABSTRACTS & BIOGRAPHIES

OPENING ORATION

CHAIRS: DR MARK WALTHAM & DR JO-ANN STANTON

1700 – 1800

ASSOCIATE PROFESSOR CHRISTOPHER MASON
Cornell University, U.S.A

Dr. Christopher Mason completed his dual B.S. in Genetics and Biochemistry from University of Wisconsin-Madison (2001), his Ph.D. in Genetics from Yale University (2006), and then completed his dual post-doctoral training at Yale Medical School in genetics and a fellowship at Yale Law School (2009). He is currently an Associate Professor at Weill Cornell Medicine, with appointments at the Tri-Institutional Program on Computational Biology and Medicine between Cornell, Memorial Sloan-Kettering Cancer Center and Rockefeller University, the Sandra and Edward Meyer Cancer Center, and the Feil Family Brain and Mind Research Institute.

The Mason laboratory develops and deploys new biochemical and computational methods in functional genomics to elucidate the genetic basis of human disease and human physiology. We create and explore novel techniques in next-generation sequencing and algorithms for: tumor evolution, genome evolution, DNA and RNA modifications, and genome/epigenome engineering. We also work closely with NIST/FDA to build international standards for these methods, to ensure clinical-quality genome measurements/editing. We also work with NASA to build integrated molecular portraits of genomes, epigenomes, transcriptomes, and metagenomes for astronauts, which help establish the molecular foundations and genetic defenses for enabling long-term human space travel.

He has won the NIH's Transformative R01 Award, the Pershing Square Sohn Cancer Research Alliance Young Investigator award, the Hirschl-Weill-Caulier Career Scientist Award, the Vallee Foundation Young Investigator Award, the CDC Honor Award for Standardization of Clinical Testing, and the WorldQuant Foundation Research Scholar Award. He was named as one of the "Brilliant Ten" Scientists by Popular Science, featured as a TEDMED speaker, and called "The Genius of Genetics" by 92Y.

He has >125 peer-reviewed papers that have been featured on the covers of Nature, Science, Nature Biotechnology, Nature Microbiology, Neuron, and Genome Biology and Evolution, as well as cited by the U.S. District Court and U.S. Supreme Court. His work has also appeared on the covers of the Wall Street Journal, TIME, and the New York Times, and across many media (ABC, NBC, CBC, CBS, Fox, CNN, PBS, NASA, NatGeo). He has co-founded three biotechnology start-up companies and serves as an advisor to many others. He lives with his daughter and wife in Brooklyn, NY.

SINGLE-CELL, CITY-SCALE, AND INTER-PLANETARY GENOMICS

The avalanche of easy-to-create genomics data has impacted almost all areas of medicine and science, and here we report the implementation of genomics technologies from single cells, to cancer patients, and into space. Recent methods and algorithms enable single-cell and clonal resolution of phenotypes as they evolve, for cancer samples as well as metagenomics. We show that evolution moves at the genetic, epigenetic, transcriptional, and epitranscriptional level, enabling many means by which cancer can resist therapy. Notably, some of these changes can be resolved by single-cell analysis and enable prognostic relevance. We reveal new biochemical methods (eRRBS, MeRIP-seq) and algorithms (methylKit, eDMR, methclone) to examine these changes. Finally, pilot data will be shown for enabling patients to become more involved in their 'omics data, including an integrative genomics view of entire cities (MetaSUB.org) that leverages longitudinal metagenome and microbiome profiles of the world's subway systems and cites to map global dynamics of DNA/RNA. All of these methods and molecular tools work together to guide the most comprehensive, longitudinal, mutli-omic view of human physiology in the NASA Twins Study and the NASA Biomolecule Sequencer Mission to enable new technologies that can sequence, quantify, and engineer nucleic acids and entire genomes for long-term human space travel.

SESSION 1: ANIMAL & PLANT GENOMICS #1

CHAIRS: DR SHANNON CLARKE & DR PHILLIP WILCOX

This session is sponsored by Beef and Lamb New Zealand

0845 – 0930

ASSOCIATE PROFESSOR HUAIJUN ZHOU
University of California, U.S.A.

Zhou, an associate professor and Chancellor's Fellow in the Department of Animal Science, specialises in immunology genetics, genomics and bioinformatics. Zhou completed his Ph.D. in Immunogenetics and Molecular Genetics, and M.S. in Bioinformatics and Computational Biology at Iowa State University, and joined the UC Davis faculty in 2011 from Texas A&M University.

His research interests include animal disease, genetic resistance, poultry (chicken), genomics, bioinformatics, bacterial and viral pathogens, identifying host and pathogen genes that are involved in the interaction between animals and pathogens, and functionally annotate livestock genomes.

FAANG INITIATIVE: IDENTIFICATION OF REGULATORY ELEMENTS IN LIVESTOCK SPECIES

Regulatory elements play an essential role in understanding how an organism's genotype determines the phenotype. The technologies and assays developed in human and mouse ENCODE projects provide a solid foundation to functionally annotate farm animal genomes. Chicken, pig and cattle are major farm animals in providing the world's food production. Robust functional annotations of their genomes could be leveraged to improve their production efficiency. Recent international FAANG (Functional Annotation of ANimal Genomes) initiative has stimulated such great efforts on livestock species. UC Davis is leading such an effort to functionally annotate regulatory elements in three livestock species. The first key step is to identify regulatory elements in the genomes by integrating RNA-seq, DNase-seq and ChIP-seq data (4 histone modification marks and CTCF) from each tissue.

We present the current progress in generating and analyzing data from these three important species, including analysis of forty-eight RNA-seq libraries (sixteen per species) collected from two biological replicates across eight tissues: adipose, cerebellum, cortex, hypothalamus, liver, lung, muscle and spleen. Using the ChromHMM statistical prediction model developed by the ENCODE project, we integrated the DNase-seq and ChIP-seq libraries to generate tissue-specific chromatin state predictions across the genome and generate preliminary landscape of regulatory elements in chicken genome.

0930 – 1000

DR AMANDA CHAMBERLAIN
Department of Economic Development, Jobs, Transport and Resources, Australia

*Dr Amanda Chamberlain undertook her Bachelor of Agricultural Science (Hons) at University of Melbourne in Australia and went on to do her PhD, also at University of Melbourne, under the supervision of Professor Michael Goddard on mapping QTL for milk production in dairy cattle and using them in marker assisted selection. She joined what is now known as the Department of Economic Development, Jobs, Transport and Resources (DEDJTR) as a post-doctoral scientist where she has been involved in the implementation of genomic selection in the Australian dairy industry, the 1000 bull genomes consortium and more recently the Functional Annotation of Animal Genomes (FAANG) consortium. Dr Chamberlain is currently working on sequencing the bovine genome and identifying regions involved in r*egulating gene expression

THE EXTENT OF CIS-REGULATION OF GENE EXPRESSION AND ITS INFLUENCE ON COMPLEX TRAIT VARIATION IN CATTLE

Amanda Chamberlain[1], Majid Khansefid[2,3], Christy J Vander Jagt[1], Benjamin J Hayes[1,4], Leah C Marett[5], Yizhou Chen[6], Sunduimijid Bolormaa[1], Catriona Millen[2,3], Thuy Nguyen[1], Michael Goddard[1,3]

1. Agriculture Victoria, Department of Economic Development, Jobs, Transport and Resources, Agribiosciences Building, Bundoora, Australia
2. Dairy Futures Cooperative Research Centre, Agriobiosciences Building, Bundoora, Australia
3. Department of Agriculture and Food Systems, University of Melbourne, Parkville Australia
4. Queensland Alliance for Agriculture and Food Innovation, Centre for Animal Science, University of Queensland, St Lucia, Australia
5. Agriculture Victoria, Department of Economic Development, Jobs, Transport and Resources, Ellinbank, Australia
6. NSW Department of Primary Industries, Elizabeth Macarthur Agricultural Institute, Menangle, Australia

Cis-regulatory genetic variation can cause changes in gene expression between individuals and tissues. Allele specific expression (ASE) and expression quantitative trait loci (eQTL) mapping have been used to locate cis-regulation. RNAseq data from two studies was used to gain insight into genetic variation in complex traits, including gene expression and production traits, resulting from cis-regulation. The first tested ASE in 18 tissues taken from a lactating cow and two tissues from 20 lactating cows. 89% of all genes tested in the single cow contained at least one SNP with significant ASE ($p<0.01$), in at least one tissue. There were some genes that displayed ASE consistently across tissues however many were tissue specific. The level of ASE and tissue specificity were validated in the two tissues from 20 cows. The second study compared SNP significantly associated with gene expression with those significantly associated with variation in complex traits. ASE was tested and local eQTL mapped in four RNAseq datasets and combined in a meta-analysis. Genome wide association studies were conducted in three populations. Results showed that SNP driving ASE were also often local eQTL implying that they were cis-eQTL. These SNP often affected gene expression in more than one tissue and the allele increasing expression was usually the same. Also, SNP significantly associated with gene expression were more likely to influence complex traits. Identification of cis-regulatory variants responsible for phenotypic variation in cattle production traits may lead to rapid identification of causative mutations affecting complex traits, and more accurate genomic selection.

1000-1030

DR KATHRYN MCRAE
AgResearch, New Zealand

Dr Kathryn McRae completed a Masters on selective sweeps in parasite selection line sheep in 2011 (University of Otago/AgResearch), and a PhD in 2015 (Dublin City University/Teagasc), where her research focussed on investigating variation in the sheep genome controlling resistance to gastrointestinal nematodes in Scottish Blackface sheep. Kathryn is currently a post-doctoral research scientist in the Animal Genomics team at AgResearch Invermay, where she is working primarily on the genetics and genomics of animal health traits in sheep.

PNEUMONIA, PURKINJE CELLS AND PITHOMYCES: GENOMICS AND ANIMAL HEALTH IN SHEEP

Kathryn M McRae[1], Suzanne J Rowe[1], Patricia L Johnson[1], Hayley J Baird[1], Shannon M Clarke[1]

1. AgResearch, Invermay Agricultural Centre, Private Bag 5034, Mosgiel, 9053, New Zealand

Genetic and infectious diseases are of major importance to livestock production worldwide. There is well-documented evidence for between-animal variation in the ability of sheep to resist multiple infectious diseases of economic importance, including both pneumonia and facial eczema. These heritable differences mean that improvement of animal health through genetic selection for enhanced resistance can be used as a complementary approach to current methods for disease control. Genomics can also be used to further increase our understanding of disease, be it differences in the ability to withstand infection, or causative mutations for inherited disease. Multiple genomic tools, including SNP chips and whole-genome sequencing, are available to pinpoint the small genetic differences that produce a variety of animal health traits in livestock. Large phenotypic datasets have been used in conjunction with high and low density genotyping to interrogate the sheep genome for regions associated with variability in resistance or tolerance to pneumonia and facial eczema, respectively. In an alternative approach, whole genome-sequencing has been used to search for a causative mutation for Cerebellar Cortical Abiotrophy (CCA), a genetic neurological disease in Wiltshire sheep.

1030 - 1045

Dr Roger Huerlimann
James Cook University, Australia

Roger Huerlimann is a post-doctoral fellow at the Centre for Sustainable Tropical Fisheries and Aquaculture at James Cook University since 2014. His main interests lie in the development and application of genomic and bioinformatics tools to answer research questions. He has been involved in a variety of research projects including 16S metagenomics, metabarcoding, transcriptomes and genomics. Currently he is part of the ARC Research Hub for Advanced Prawn Breeding, where he is in charge of the de novo assembly and annotation of the Tiger Prawn (Penaeus monodon) genome and transcriptome.

SECRET IDENTITY OF THE SUPERPRAWN: NOVEL TRANSCRIPTOME ASSEMBLY AND ANNOTATION OF THE BLACK TIGER PRAWN (PENAEUS MONODON)

Dr Roger Huerlimann[1,2], Dr. Lavinia Gordon[1,3], Mr. Jake Goodall[1,4], Dr. Kirby Siemering[1,3], Dr. Melony Sellars[1,4], Dr. Greg Coman[1,4], Dr. Nick Wade[1,4], Prof. Herman Raadsma[1,5], Mr. Dallas Donovan[1,6], A/Prof Kyall Zenger[1,2], Dr. Gregory Maes[1,7], Prof. Dean Jerry[1,2]

1. ARC Research Hub for Advanced Prawn Breeding
2. Centre for Sustainable Tropical Fisheries and Aquaculture, College of Science and Engineering, James Cook University, Townsville, QLD 4811, Australia
3. Australian Genome Research Facility Ltd, Royal Parade, Parkville, VIC 3050, Australia.
4. CSIRO Agriculture and Food, Qld Bioscience Precinct, St Lucia, QLD 4067, Australia.
5. Animal Bioscience Group, Faculty of Veterinary Science, The University of Sydney, Camden, NSW 2570, Australia
6. Seafarms Pty Ltd, Brisbane, QLD, Australia
7. Laboratory of Biodiversity and Evolutionary Genomics, KU Leuven, B-3000 Leuven, Belgium

Continued next page

The black tiger prawn (Penaeus monodon) is one of the major farmed marine prawn species globally, with an annual value of 4.5 B USD. At present, expansion of the industry is hampered by a lack of selective breeding programs that capitalise on the genetic potential of the species to improve growth, survival and tolerance or resistance to commercially devastating pathogens. The construction of a complete transcriptome and genome of the black tiger prawn are a key foundation to enable a genomics-assisted breeding program which has the highest potential for the selection of traits that cannot be directly measured on the selection candidates (eg pathogen tolerance, meat quality).

There is a deficiency of sequence information for crustaceans, and the available transcriptome assemblies for prawns cover few tissue types and genearly use a single transcriptome assembler. To remedy this lack of data, and produce a thorough transcriptome assembly, samples from nine different adult black tiger prawn tissues and nine early life-stages have been sequenced using Illumina HiSeq, resulting in approximately 30 M reads per sample. Furthermore, muscle tissue from an additional eight commercially important prawn species have been sequenced and assembled to identify novel genes of importance shared within the family. To allow for differences introduced by the assembly algorithms, the assembly was carried out using four different assemblers. Transrate was then used to assess the quality of each assembly, to filter bad contigs, and to merge the assemblies.

The discovery of novel genes involved in life history transitions and tissue specific gene expression has been combined to generate a complete transcriptome for the black tiger prawn that will be an invaluable resource for further research and genomics-assisted breeding programs. The information will be used to annotate the complete genome of the black tiger prawn which we are currently assemblingcurrently under construction.

1045 - 1100

Dr Christy Vander Jagt
Agriculture Victoria, Australia

Dr Christy Vander Jagt completed her Bachelor of Science (Hons)/Bachelor of Bioinformatics in 2004 (LaTrobe University), and a PhD in 2012 (University of Melbourne), where her project involved the comparative analysis of marsupial and eutherian genomes to identify genes critical in milk production. Christy is currently a research scientist with the Victorian Department of Economic Development, Jobs, Transport and Resources (DEDJTR) Computational Biology department. Her current work involves the analysis of RNA-sequencing data and assisting with the day-to-day running of the '1000 Bull Genomes' project pipeline.

EXPRESSION QUANTITATIVE TRAIT LOCI MAPPING IN BOVINE MILK CELLS

Dr Christy Vander Jagt[1], Dr Amanda Chamberlain[1], Professor Michael Goddard[1,3], Dr Iona MacLeod[1], Dr Coralie Reich[1], Claire Prowse-Wilkins1, Brett Mason[1], Josephine Garner[2], Professor Ben Hayes[1,4]

1. Agriculture Victoria, Department of Economic Development, Jobs, Transport and Resources
2. Agriculture Victoria, Department of Economic Development, Jobs, Transport and Resources
3. Faculty of Veterinary & Agricultural Science, The University of Melbourne
4. Centre for Animal Science, The University of Queensland

Continued next page

It is now widely regarded that much of the genetic variation affecting complex traits lies in regions of the genome regulating gene expression. Expression quantitative trait loci (eQTL) mapping is one way of finding cis-regulatory genetic variation. Gene expression in the bovine mammary gland is likely to affect milk production- and health related-traits such as mastitis and fertility. However, due to cost and ethical considerations, mammary glands are not ideal tissues to sample. We instead propose that the analysis of RNA obtained from bovine milk cell samples may be used as a substitute to aid in the identification of regulatory variants affecting gene expression in the bovine mammary gland. In this study we have used RNA-sequencing (RNAseq) data from milk samples taken from 29 Jersey and 112 Holstein dairy cows in an attempt to identify eQTL that affect gene expression in the bovine mammary gland.

Paired end RNAseq reads were generated using the Illumina platform for each of the 141 milk samples. High quality read pairs were then aligned to the Ensembl UMD3.1 bovine genome assembly using Tophat2 (Kim et al., 2013) allowing for two mismatches. Samples with >25 million read pairs also having a mapping rate of >80% were retained for gene count matrix generation. On average 126 million reads per library were generated with 88.4% reads passing quality control and a mapping rate of 91.73%. Gene counts were generated for 114 samples using HTSeq (Anders et al., 2014) and then combined and normalised using DESeq (Anders and Huber, 2010) to form a gene count matrix. As expected, the most highly expressed genes (most sequence reads after normalisation) included the four casein genes (alpha-s1, alpha-s2, beta and kappa).

Using BovineSNP50 genotypes whole genome sequence data was imputed for each animal utilising the 1000 bull genome sequences (Run5) as reference population (Daetwyler et al., 2014). After removing variants that had a minor allele frequency less than 0.05 in the experimental cows, 11,115,780 variants remained. Only genes that were expressed in more than 25% of the cows were analysed, to avoid spurious associations due to very low read counts. Gene counts for 12,075 genes were tested for association with all the variants on the chromosome that contained that gene tested. That is, 12,075 genome wide association analyses were run, with up to 690,000 variants per gene tested (for genes on chromosome 1).

Association testing was performed with EMMAX (Kang et al., 2010) fitting the genome relationship matrix among cows to control for population structure, and fixed effects of breed (Holstein or Jersey), parity, days in milk, sampling day and RNAseq batch. Read counts were transformed as log(x+1), where x is the read count of a particular gene for a cow.

eQTL analysis revealed 5827 variants significantly associated with 132 genes ($p<1x10^{-6}$, false discovery rate of 0.016), indicating that multiple variants, in high linkage disequilibrium are detecting the same eQTL. A trend was also observed for the most significant variants being in closer proximity to genes compared to less significant variants. Interestingly, the 132 genes included four solute carrier genes. Recently a mutation in a solute carrier has been proposed as a causative mutation affecting milk production by Kemper and colleagues (2016).

The most significant variants associated with these 132 genes have been included in the next design for a low density SNP chip for genomic testing of young dairy bulls and heifers. These variants can then be genotyped in a large number of animals and tested to determine if they improve the reliability of genomic breeding values for dairy traits of economic importance.

SESSION 2: ANIMAL AND PLANT GENOMICS

CHAIRS – PROFESSOR ROGER HELLENS & DR ELENA HILARIO

1130–1215

Dr Elspeth MacRae
Scion, New Zealand

Elspeth MacRae leads Scion's wood and fibre processing activities including biorefinery, biodiscovery and bioplastic, clean technology and biotech/omics activities (>100 staff). She leads Portfolio 5 (Materials Manufacturing and Applications) for the NZ National Science Challenge in Science and Technology for Industry, and overviews the 3-D printing project. She belongs to the International Advisory Group for the Global Bioeconomy Forum. She is Deputy Chair of the Biopolymer Network Ltd Board, and a member of the Bioprocessing Alliance and the New Zealand Packaging Council governance groups. She is a member of the EPA's BRMAC (advisory) group 2011-2016, and recently moderated the Advanced Manufacturing session in the Auckland Tripartite Economic Summit with Guangzhou and Los Angeles. She has represented Scion in biotechnology policy issues for 10 years and has sat on multiple grants panels.

RADIATA PINE: A LARGE GENOME, GENE EDITING AND LEGISLATION EXPERIENCES

Elspeth MacRae, Alex Hennebry and Glenn Thorlby, Scion

The radiata pine genome of ~25 billion base pairs, 8x the size of the human genome, is packaged into 12 chromosomes[1]. Most of the size can be explained by divergence of retrotransposons[2] rather than duplication of genes (polyploidy). To make transgenic pine requires the transformation of cell lines derived from the embryos contained within immature green cones and the subsequent production of plants from these transformed cells via somatic embryogenesis. To establish trees in soil takes 18 months with an additional 4 months needed before trees are ready to plant in the field.

Using conifer-verified promoters we have optimised a CRISPR-cas9 vector system to implement transgenic gene editing in radiata pine and have carried out initial transformations. Non-transgenic gene editing will require the direct delivery of the gene editing complex into cells. Preliminary work has demonstrated the import and expression of proteins into radiata pine protoplasts and import into embryo-derived cells in culture is being tested. This work is discussed in the context of the NZ HSNO Act and impact on research and deployment.

1. Wilcox and Macdonald 2015. NZJ For. 60:17-22.
2. Morse, Peterson et al 2009. PLoS ONE e4332.

1215–1245

PROFESSOR ROGER HELLENS
Queensland University of Technology, Australia

Professor RP Hellens is currently the acting Assistant Dean (Research and Innovations) for the Science and Engineering Faculty and Deputy Director for The Institute for Future Environments. He joined Queensland University of Technologies' Centre for Tropical Crop and Biocomodities in March 2014.

From 2000 to 2014 he worked in New Zealand's crown research institute; Plant&Food in Auckland, where he had a number of role including leading the genomics research and the kiwifruit breeding programme.
His research interests were the development of red-fleshed apple and kiwifruit varieties and exploiting Next Generation Sequencing (NGS) techniques in breeding programmes. He has also maintained a keen interest in post transcriptional gene regulation and this has become relevant in work to understand the regulation of vitamin C.

Continued next page

Prior to his move to New Zealand, Professor Hellens worked at the John Innes Centre in Norwich, UK. Here he developed the first genetic maps in pea (including his PhD on the molecular basis of Mendel's white flower phenotype), He developed the pGreen plant transformation vector and a project on gene silencing (RNAi) in petunia.

WHERE ARE ALL THE SMALL PEPTIDES AND WHAT MIGHT THEY DO?

While genome sequencing is now commonplace, genome annotation and functional characterisation of components of the genome remains an enormous challenge. Like small regulatory RNAs, short open reading frames (sORFs) often reside in 'non-coding' regions of the genome that have long been considered to be 'junk DNA [1].

Translatable sORFs of less than 100 amino acids are extremely difficult to predict from genome sequences as the number of potential ORFs increases exponentially as the potential peptide lengths get smaller [2]. This challenge is made more complicated by mounting evidence that these short proteins do not always comply with genetic convention, and are frequently encoded by short ORFs that use a translation start codon other than AUG [3].

1. Waterhouse PM, Hellens RP (2015) Plant biology: Coding in non-coding RNAs. Nature 520: 41-42.

2. Hellens RP, Brown CM, Chisnall MA, Waterhouse PM, Macknight RC (2015) The Emerging World of Small ORFs. Trends Plant Sci.

3. Laing WA, Martinez-Sanches M, Wright M, Bulley S, Brewster D, et al. (2015) A non-canonical upstream open reading frame is essential for feedback regulation of ascorbate biosynthesis The Plant Cell 27: 772-786.

1245 – 1315

Professor Dave Edwards

University of Western Australia, Australia

David Edwards is a Professor at the University of Western Australia. He gained an Honours degree in agriculture from the University of Nottingham and a PhD from the Department of Plant Sciences, University of Cambridge. He has held positions within academia, government and industry, working in the UK, Canada and Australia. He currently leads a bioinformatics research group with interests including the structure and expression of plant genomes, the discovery and application of genome variation and applied bioinformatics, with a focus on plants, understanding their evolution and accelerating crop improvement in the face of climate change.

DOES LENGTH MATTER?

Next generation DNA sequencing has revolutionised biology. Draft genome assemblies are being produced for a diverse range of species, with many recent genome assemblies being produced using data the common short read Illumina sequencing platform. While long read sequencing has been around for a few years, the low accuracy of these platforms has limited their practical application. Recent advances have seen an increase in the use of long read sequencing for genome assembly, competing with advanced algorithms for the assembly of large insert short read data. With the rapid changes in sequencing technology, researchers need to make decisions as to which is the best tool to ask their biological questions. In this presentation, I will outline the recent changes in DNA sequencing and comment on approaches for the application of this data.

1315 – 1330

Natalie Graham
Scion, New Zealand

Natalie Graham is a scientist in the Forest Genetics team at Scion, New Zealand's forest research institute. With a background in molecular biology, she has an interest in applying new molecular techniques to solving industry challenges. Previous work has included candidate gene-based SNP discovery and association genetics research in radiata pine. She is currently involved in a Phytophthora research programme to identify signature gene expressions profiles for resistant and susceptible trees. She is also part of a partnership programme with the Radiata Pine Breeding Company to develop and apply genomic selection tools in radiata pine breeding programmes.

JUMPING ON THE EPIGENETICS BANDWAGON THROUGH HYPOMETHYLOME SEQUENCING

Natalie Graham[1], Catherine Reeves[1], Cathy Hargreaves[1].

1. Forest Genetics, Scion

If nucleotides and genes are like letters and words, then epigenetic modifications have been compared to the punctuation marks that impart sense and meaning. Methylation is seen as the most common epigenetic modification in plants, and is probably the most widely-studied. Active gene regions tend to have low levels of methylation, while heavily-methylated gene regions tend to be under-expressed. Methylation is also very common in repetitive regions of the genome.

While genomic tools are being developed for radiata pine (Pinus radiata D. Don) breeding, we know that many characteristics of this species are greatly influenced by physiological age and the environment, suggesting a strong involvement of epigenetic modifications. Previous studies in radiata pine have used more broad scale approaches such as global methylation indicators, or narrowed their focus to a handful of candidate genes. However, these approaches have their limits - global methylation can only detect very large differences and provides no information about where in the genome changes might be occurring, and targeted gene assays require a level of prior knowledge to inform candidate selection. We wanted to determine whether high resolution next generation sequencing (NGS)-based assays could be used to prospect for candidate genomic regions under epigenetic regulation in radiata pine. Although whole-genome bisulfite sequencing (WGBS) is seen as the gold standard for determining methylation status, the depth of coverage required (ideally 15-20x), the size of the genome (25 Gb) and access to a good reference genome (currently in ~27 million scaffolds) make this approach unsuitable for radiata pine. Reducing the complexity of a sample can decrease the amount of sequencing needed to achieve the required depth of coverage. However, many of these methods enrich for the methylated portions of the genome, which includes much of the repetitive regions – since more than 80% of the 25 Gb radiata pine genome is estimated to be repetitive, the effective reduction in complexity would likely be insufficient to overcome the challenges posed by WGBS. Methylation-sensitive restriction enzyme-based sequencing (MRE-Seq) is an alternative that allows for enrichment of the non-methylated regions of the genome, termed the hypomethylome. This method also lends itself to investigating non-model species with complex genomes, and does not rely on the pre-existence of a reference genome.

Using two genetically identical cell-lines with different histories and divergent phenotypes, we investigated the suitability of MRE-Seq for radiata pine. Radiata pine clonal lines, used for somatic embryogenesis, are an essential component of clonal forestry, as well as a valuable research tool. The ability to quantify epigenetic modifications could have tremendous value if the performance of a cell-line or the resulting somatic embryos could be predicted by its epigenetic status. We report on the potential for MRE-Seq to assess methylation status in radiata pine and compare this to RNA expression information as an indicator of cell-line performance.

1330 – 1345

Carolina Peñaloza
Roslin Institute, United Kingdom

Ph.D. candidate at the Roslin Institute, University of Edinburgh, working in shellfish genetics.

ANALYSIS OF GENOME-WIDE INHERITANCE PATTERNS USING RAD SEQUENCING IN BIVALVE SPECIES

Peñaloza C., Bishop S.C., Houston R.D.
The Roslin Institute, University of Edinburgh, Edinburgh, UK

Selective breeding of bivalves (e.g. mussels and oysters) for aquaculture is at an early stage. Incorporation of genomic markers into breeding schemes has major potential for both family assignment and genomic selection. However, genetic markers in bivalve species show unusual segregation patterns, which may potentially limit their applicability. In most species, molecular markers follow predictable Mendelian patterns of inheritance within families. In bivalves, however, genotypic ratios of offspring significantly deviate from Mendelian expectation at polymorphic loci. This phenomenon of segregation distortion is widespread among other shellfish species, and usually affects most of the markers under study. Despite the importance of understanding the phenomenon of segregation distortion in shellfish, the origin of these unusual inheritance patterns remains unclear.

To gain further insight into this ubiquitous phenomenon we studied the inheritance pattern of SNP markers in bivalve families using Restriction-site-Associated-DNA sequencing (RAD-seq); a powerful tool for genome-wide marker discovery and genotyping in non-model species. The DNA from two Chilean blue mussel (Mytilus chilensis) (n=96) and three oyster (Crassostrea gigas) families (n=72) was extracted and quantified using a fluorometric assay. Two SbfI RAD libraries (one for each species) were prepared following published protocols. Each individual in the sequencing library had a unique barcode combination of P1 and P2 adapter sequences that allowed sample identification after sequencing. RAD libraries were sequenced on two lanes of a HiSeq 2500 Illumina platform.

De novo assembly and SNP discovery and genotyping were performed using Stacks version 1.40.

A total of 4,500 and 4,304 putative SNPs were identified in the mussel and oyster genome, respectively. A very high rate of polymorphism was observed for both species: an average of 1 SNP each 30bp for the mussel genome and 1 SNP each 25bp for oysters. The analysis of inheritance across bivalve families showed high levels of segregation distortion among SNP markers. For instance, distorted loci comprised 72% of the discovered SNPs in mussels, and showed a clear tendency towards a deficiency of heterozygous genotypes in the offspring. In addition, alleles were detected in the offspring that were apparently not present in their parents. The presence of these unexpected alleles was confirmed by Sanger sequencing.

In conclusion, genetic marker in the bivalve shellfish species evaluated is complex. We are currently undertaking experiments to investigate the origins and possible causes of these highly unusual segregation patterns.

SESSION 3: NEW TECHNOLOGIES

CHAIRS – DR JO-ANN STANTON & MR PETER SMITH

1515 – 1600

Associate Professor Christopher Mason
Cornell University, U.S.A

Dr. Christopher Mason completed his dual B.S. in Genetics and Biochemistry from University of Wisconsin-Madison (2001), his Ph.D. in Genetics from Yale University (2006), and then completed his dual post-doctoral training at Yale Medical School in genetics and a fellowship at Yale Law School (2009). He is currently an Associate Professor at Weill Cornell Medicine, with appointments at the Tri-Institutional Program on Computational Biology and Medicine between Cornell, Memorial Sloan-Kettering Cancer Center and Rockefeller University, the Sandra and Edward Meyer Cancer Center, and the Feil Family Brain and Mind Research Institute.

The Mason laboratory develops and deploys new biochemical and computational methods in functional genomics to elucidate the genetic basis of human disease and human physiology. We create and explore novel techniques in next-generation sequencing and algorithms for: tumor evolution, genome evolution, DNA and RNA modifications, and genome/epigenome engineering. We also work closely with NIST/FDA to build international standards for these methods, to ensure clinical-quality genome measurements/editing. We also work with NASA to build integrated molecular portraits of genomes, epigenomes, transcriptomes, and metagenomes for astronauts, which help establish the molecular foundations and genetic defenses for enabling long-term human space travel.

He has won the NIH's Transformative R01 Award, the Pershing Square Sohn Cancer Research Alliance Young Investigator award, the Hirschl-Weill-Caulier Career Scientist Award, the Vallee Foundation Young Investigator Award, the CDC Honor Award for Standardization of Clinical Testing, and the WorldQuant Foundation Research Scholar Award. He was named as one of the "Brilliant Ten" Scientists by Popular Science, featured as a TEDMED speaker, and called "The Genius of Genetics" by 92Y.

He has >125 peer-reviewed papers that have been featured on the covers of Nature, Science, Nature Biotechnology, Nature Microbiology, Neuron, and Genome Biology and Evolution, as well as cited by the U.S. District Court and U.S. Supreme Court. His work has also appeared on the covers of the Wall Street Journal, TIME, and the New York Times, and across many media (ABC, NBC, CBC, CBS, Fox, CNN, PBS, NASA, NatGeo). He has co-founded three biotechnology start-up companies and serves as an advisor to many others. He lives with his daughter and wife in Brooklyn, NY.

"SINGLE-MOLECULE SEQUENCING ON EARTH AND IN SPACE."

Several new technologies enable high-resolution characterization of single cells and populations of cells, such as patients' response to chemotherapy and viruses as they evolve. Here, we describe new methods to characterize modified bases of DNA (epigenomes) and RNA (epitranscriptomes) with single-molecule and enrichment-based methods, and their impact on tracing clinical dynamics during care, as well as their applications for global monitoring of cross-kingdom, shifting alleles. We also show how portable DNA sequencing can enable rapid diagnostics and also present data from the first-ever DNA sequencing experiments in microgravity and in space. These results show the promise of an emerging era of ubiquitous sequencing and enable novel applications for human health and long-term survival.

1600 – 1620

CAMERON JACK
The Australian National University, Australia

Mr Jack has been a bioinformatician at the Australian National University (ANU) since 2011 and is the manager of the ANU Bioinformatics Consultancy - a cost recovered team of bioinformaticians providing infrastructure and analysis capabilities to ANU, the Canberra region, and beyond. His team's analysis work spans the gamut; covering fields as wide as plant genetics, immunological response, fundamental cell biology and epigenetics. They are now focusing their efforts on improving the interface between biologists and bioinformatic tools. Mr Jack completed his undergraduate degree in computer science at Victoria University of Wellington (VUW) in 2003 and went on to work as a technician, microscopist, and bioinformatician at VUW's School of Biological Sciences.

A FLEXIBLE SEQUENCE ANALYSIS PIPELINE PLATFORM FOR NON-BIOINFORMATICIANS

Camero Jack[1], Aaron Chuah[1], Robert E. Buckley[1], Bhavik Mehta[2], Runa Daniel[3], Dennis McNevin[4]

1. ANU Bioinformatics Consultancy, Australian National University, Acton, ACT, 0200, cameron.jack@anu.edu.au
2. Specialist Operations - Forensics, Australian Federal Police, Majura, ACT, 2609
3. Office of the Chief Forensic Scientist, Victoria Police Forensic Services Department, Macleod, Victoria, 3085
4. National Centre for Forensic Studies, Faculty of ESTeM, University of Canberra, Bruce, ACT, 2617

Many bioinformatics workflow managers already exist to aid genomics researchers in building reproducible pipelines. However, practical difficulties remain in translating general-purpose workflow managers from bioinformatic toolkits to discipline-specific, production-ready applications. We present our flexible graphical pipeline platform, commissioned by the National Centre for Forensic Studies at the University of Canberra as a customisable open-source data processing platform for discovering Single Nucleotide Variants and Short Tandem Repeats in Massively Parallel Sequencing data.

This was conceived after popular existing workflow tools such as Galaxy were found to be insufficient in addressing the end user requirements of the forensics community for a stable, production-ready graphical analysis suite that could be installed and customized without the need for trained computer operators. We have made a deliberate trade-off in limiting new pipeline design to programmers, whilst retaining full custom run-time and sample filtering by untrained users, in return for "single-click" installation and operation. Using Python and JavaScript on an Ubuntu Linux operating system, we ensure that all components are open source and widely available. Integrating data processing and visualisation capabilities, our pipeline platform offers a unique and new way to make complex analytical pipelines readily accessible to previously non-catered-for communities.

1620 – 1640

DR KATHERINE PILLMAN
University of South Australia, Australia

Dr Pillman is a bioinformatician with a broad interest in many areas of gene regulation. She began her career as a wet-bench biologist, working on exploring regulation of gene transcription in barley plants (Australia) and stress-responsive gene regulatory networks in potato plants (Oregon State University). Her experience with RNA-seq in the latter project fuelled her interest in bioinformatics. In 2012, she returned to Australia to take up the role of lead bioinformatician with Prof Greg Goodall, working on gene regulatory mechanisms and networks in cancer with a focus on the epithelial-to-mesenchymal transition in breast cancer. Their recent seminal paper identified the first protein known to control the formation of circular RNAs and characterising circular RNAs during the epithelial-to-mesenchymal transition. Her current work involves using a range of genomics next-generation sequencing data types to dissect gene regulation, including analysis of circular RNAs, alternative splicing, microRNA biology and targeting, epigenetic modifications, gene regulatory networks and expression.

Continued next page

SWIMMING IN CIRCLES

Katherine A. Pillman [1,2,3], Simon J. Conn[1], John Toubia[1,2,3], Gregory J. Goodall[1,2]

1. Centre for Cancer Biology, SA Pathology, University of South Australia, SA, 5000, Australia
2. University of Adelaide, SA, 5000, Australia
3. ACRF Cancer Genomics Facility, SA, 5000, Australia

Circular RNAs (circRNAs) are a single stranded RNA species which have recently been at the centre of an explosion of interest in the RNA community. Produced by back-splicing the exons of linear mRNAs, circRNAs were originally overlooked as rare by-products of canonical splicing. However, recent advances in next generation sequencing technology and associated bioinformatic methods revealed a surprising pervasiveness. They have since been identified in their thousands in humans and across the eukaryotic domain.

Recently, high profile research has made significant discoveries, casting light on the cell type-specific expression profiles of circRNAs, mechanisms of biogenesis and cellular function of a handful of circular RNAs. Our research determined that hundreds of circRNAs are regulated during human epithelial to mesenchymal transition and that production of many of these is regulated by the alternative splicing factor, Quaking. This was the first conclusive mechanistic evidence that RNA-binding proteins can affect circRNA abundance in human cells.

Despite the recent surge of interest on circular RNAs, there are still many important and potentially high-impact questions to be answered. Arguably, two of the most valuable will be determining a) their broad and context-specific functions and b) whether they can be used as molecular biomarkers for disease. For many researchers, the ability to use their existing total RNA-sequencing datasets to identify and quantify circular RNAs makes this an especially appealing area of investigation. The bioinformatic analysis of circular RNAs brings its own challenges. With a focus on bioinformatics, I will discuss what we know about this new class of non-coding RNAs and outline approaches to identify and analyse them and make recommendations from our experience.

1640 – 1700

DR MICHAEL RHODES
NanoString Technologies, U.S.A.

Graduated from York University with a degree in Genetics, did a Ph.D. in Bioinorganic Chemistry at University of London. After a post doc in Chicago working on genetics of metal transport in P. aeruginosa, returned to UK to work at United Kingdom Human Genome Mapping Project Resource Centre finishing as Operation Manager in charge of four teams: - Mouse resequencing, linkage Hotel, Academic Services and Custom services. Joined Applied Biosystems in 1999, worked on Genotyping, sequencing, qPCR and finally Next Generation Sequencing. Joined Nanostring in 2012 after seeing the potential of the nCounter technology to take the discoveries from NGS and apply them to translational research.

ENZYME-FREE, AMPLIFICATION-FREE, HYBRIDIZATION-BASED SINGLE MOLECULE SEQUENCING TECHNOLOGY USING FLUORESCENT OPTICAL BARCODES: A FIRST-IN-CLASS CHEMISTRY WITH SEVERAL UNIQUE FEATURES

Hyb & SeqTM is an enzyme-free, amplification-free, single molecule sequencing chemistry using cyclic nucleic acid hybridization of fluorescent molecular barcoded n-mers. Targets are sequenced using barcodes with the structure: (nnb1b2b3b4b5b6nn—linker—r1r2r3r4r5r6), where: n = mixture of all 4 bases, b1 through b6 are the 6 contiguous bases that form the complement to the target DNA/RNA to be sequenced, and r1 through r6 is a sub-diffraction-limited cyclically-read optical barcode that encodes the identity of the six-bases hybridized to that target via 4-color fluorescence (~25 fluorophores per "r").

Continued next page

Genomic DNA or RNA is first "gapped" to generate a single-stranded region and captured onto a flow-cell. Optical barcodes are hybridized to these single-molecule targets, and bases at each hybridized target are read through a series of 6 "fast-exchange" hybridizations to the reporter (r)-regions. Each cycle yields a 6-base read. The hybridized probes are eluted, and the cycle is repeated until all regions have sufficient coverage. Proof-of-concept was achieved using a prototype sequencer based on a modified nCounter Sprint Profiler instrument and BRAF V600E model system. Fast reporter readout-cycling was confirmed (<10 seconds). The raw single-pass error rate averaged 2.1% (min = 0.66%, max = 3.72%) and did not reveal any systematic bias. Based on this data, less than 5x coverage from a single molecule would be required to reach a consensus sequence accuracy of 99.99% (Q40).

Hyb & Seq has many potential advantages over other sequencing chemistries. Read lengths can be as long as the gapped single-stranded regions. Sequencing error is independent of read length. No covalent bonds are formed, making each cycle non-destructive and allowing multiple reads on the same native single molecule, drastically reducing error rates. The workflow does not require enzymes, amplification, or libraries, opening the possibility of a sample-to-answer sequencing instrument appropriate for clinical use.

1700 – 1715

Dr Denis Bauer
CSIRO, Australia

Dr. Denis Bauer is the team leader of the transformational bioinformatics team in CSIRO's ehealth program. Her expertise is in high throughput genomic data analysis, computational genome engineering, as well as Spark/Hadoop and high-performance compute system. She has a PhD in Bioinformatics and has done her Postdoctoral training in machine learning and human genetics, respectively. Her collaborators include Prof Simon Foote on mammalian susceptibility to infectious diseases, Prof Ian Blair on molecular mechanisms on motor neuron disease, and Prof Rodney Scott on obesity-driven cancer. She has 25 peer-reviewed publications (9 first author, 4 senior author) with four in journals of IF>8 (e.g. Nat Genet.) and H-index 9. To date she has attracted more than AU$25Million in funding.

CHROMATIN-AWARE CLOUDBASED CRISPR TARGET SITE PREDICTOR

Denis C. Bauer[1], O. Wilson[1], Aidan O'Brien[1], Oscar J. Luo[1], Robert Dunne[2]

1. CSIRO, Health and Biosecurity, North Ryde, Sydney
2. CSIRO, Data 61, North Ryde Sydney

Genome Engineering holds great promise for personalized medicine and the CRISPRCas9 system is one of the most widely adopted genome editing mechanisms. For medical applications being able to accurately score targets for their sensitivity and offtarget specificity is critical. This, however, is hampered by uncertainty around the influence of the chromatin environment on the in vivo binding activity.

We hence investigate whether chromatin marks are able to predict CRISPRCas9 activity by performing a meta study over the in vivo binding activity of sgRNAs. While looking at DNase Hypersensitivity marks as a proxy for accessibility we indeed see little correlation with CRISPRCas9 functionality. However, we find that specific histone marks are significantly associated with sgRNA activity (pvalue < 0.05).

We hence designed GTScan2, which is the first method to combine information about the sgRNA structure and the chromatin environment to predict CRISPR-Cas9 target sites. Our method shows up to a 37% improvement over previously published methods when tested on two independent datasets. By leveraging the Roadmap Epigenomics Project, GTScan2 is readily applicable to all human tissues/cell types and provides a full endtoend service from identifying target sgRNAs to evaluating their effectiveness. GTScan2 is available as an Amazon Lambda function, which allows server-less continuouslyscalable applications, that can be shared and built upon without barrier to access.

1715 – 1730

Dr Ken McGrath
AGRF, Australia

Ken McGrath is the National Sanger Sequencing Manager at the Australian Genome Research Facility, based in Brisbane, Australia. He obtained his PhD studying Molecular Pathology in 2005 from the University of Queensland, and has a research background in microbial community genomics, including human and environmental microbiomes and metagenomics analysis. Ken is currently involved with several research projects, including the US-based eXtreme Microbiome Project (XMP), as well as evaluating emerging technologies that can be used to profile the diversity of microbial communities.

THE EXTREME MICROBIOME PROJECT PRESENTS: METAGENOMICS ON THE MINION -NANOPORE SEQUENCING OF ENVIRONMENTAL MICROBIAL COMMUNITIES

Ken McGrath[1], Rachael McNally[1], Lavinia Gordon[1], Alexis Lucattini[1], Kirby Siemering[1]

1. Australian Genome Research Facility

The eXtreme Microbiome Project (XMP) is a global scientific collaboration to characterize, discover, and develop new pipelines and protocols for studying novel microorganisms in extreme environments.

While many genomics platforms are used in this project, the bioinformatics analysis of short-read metagenomics data is complicated by conserved gene families, gene duplications/ repeats, and homology between closely related species. In these circumstances, long-read sequencing technologies are proving to be a valuable tool for detecting novel and rare organisms in complex communities where reference genomes are unavailable.

We have used the Oxford Nanopore MinION platform to sequence a metagenomic mock community (gDNA pool), to benchmark the current performance of the technology (R9 cells). The analysis of the error rates and accuracy of the assembled data from the control community demonstrate the utility of this platform for rapid microbial identification.

In addition, we have used the MinION to sequence environmental samples from the eXtreme Microbiome Project's Lake Hillier, a bright-pink hypersaline lake located on a remote island off the Western Australian coast. The results obtained from the water and sediment samples of the lake reveal a wide range of microbial diversity in the lake, and identify the algal, bacterial, and archaeal halophiles that contribute to the pigmentation of this distinctive feature of the Australian landscape.

SESSION 4: HUMAN DISEASES & CLINICAL GENOMICS

CHAIRS: DR LOGAN WALKER & ASSOCIATE PROFESSOR RUBY CY LIN

0830–0915

Associate Professor Lance Miller
Wake Forest School of Medicine, U.S.A

Dr. Miller is the Mary Kirkpatrick Associate Professor of Breast Cancer Research, Department of Cancer Biology, Wake Forest School of Medicine in Winston-Salem, North Carolina, USA. Within the Wake Forest Comprehensive Cancer Center, Dr. Miller is the Director of the NCI-supported Cancer Genomics Shared Resource, and Director of the Breast Cancer Center of Excellence. Dr. Miller received his Ph.D. in Genetics and Molecular Biology from UNC-Chapel Hill in 2001 under the mentorship of renowned cancer researcher, Dr. Edison T. Liu, MD. Professionally, Dr. Miller has served as staff scientist at the National Cancer Institute in Bethesda, Maryland (1999-2001) and as Senior Group Leader at the Genome Institute of Singapore (2001-2008). Dr. Miller joined the Department of Cancer Biology at Wake Forest in 2008. Dr. Miller's research focuses on questions related to the functional genomics of breast cancer and other cancer types. Using bioinformatics methods to quantify and interpret genome-scale data from large patient cohorts, Dr. Miller's laboratory studies how unique genomic and transcriptomic properties of tumors relate to cancer behavior and clinical outcomes of patients. This information, in turn, drives the discovery of novel molecular interactions that govern aspects of cancer growth and progression, as well as the development of new biomarkers to aid in clinical decision making. Dr. Miller is currently funded by the American Cancer Society and the National Cancer Institute, and has multiple long-standing collaborations with cancer researchers in New Zealand.

UNDERSTANDING THE INTERPLAY BETWEEN TUMOR-IMMUNE INTERACTIONS AND PATIENT OUTCOMES THROUGH GENOMIC CORRELATES

While immunotherapies that augment anti-tumor immune responses are advancing in the clinic, the ability to predict responsive and nonresponsive patients remains a major challenge. Central to this challenge is a lack of understanding of how tumor molecular and genetic factors interact with the host immune system to influence patient outcomes. Recently, we and others have shown that immune gene signatures deduced from tumor expression profiles can be used to deconstruct aspects of the immune system's involvement in cancer control. In this presentation, I will discuss how bioinformatics approaches are providing new insights into fundamental principles that govern the immunogenic potential of cancer

0915 - 0945

DR MILES BENTON
Institute of Health and Biomedical Innovation, Queensland University of Technology, Australia

Dr Benton is computational geneticist interested in all facets of biology and technology. He completed his masters exploring mitochondrial genetics in 2010 (Victoria University of Wellington/Environmental Science and Research) and a PhD exploring the underlying genetic susceptibility to obesity and metabolic disorders in the Norfolk Island population isolate (Griffith University, Gold Coast). He is currently working at Queensland University of Technology as a postdoctoral research fellow. His main areas of interest are Bioinformatics and Computational Genetics, he also has an active interest in population genetics. Dr Benton's current work involves the creation of methods to deal with, and analyse, large genomic data sets, including incorporation of multiple layers of both phenotypic and genomic data. He is also a big advocate of reproducible research and open-source science.

INSIGHTS FROM A GENETIC ISOLATE

It is well established that genetic isolates have distinct advantages for mapping gene variants that influence complex traits. The Norfolk Island population is one such isolate. Located off the east coast of Australia, the original population was founded by 11 British mutineers of the HMS Bounty and 6 Polynesian women in the late 1700s. This admixture gives rise to a very unique pedigree structure, with upwards of 6000 members spanning 11 generations, or 200 years.

I have been fortunate enough over the last 6 years to have been involved in the analysis and interpretation of a wealth of genomic data from this fascinating population. In this talk I will highlight what we have discovered thus far using a range of different genomic techniques in a series of case studies, including: the development of an integrated 'phenomic' signature approach; whole genome sequencing leading to the identification of enriched functional founder effect variants associated with metabolic traits, and the recent identification of multi-generational allelic-specific methylation profiles.

0945-1015

DR LOUISE BICKNELL
University of Otago, New Zealand

Dr Bicknell completed her PhD in human genetics with Professor Stephen Robertson in 2007 before undertaking a postdoctoral position studying the genetic basis of growth disorders with Professor Andrew Jackson at the MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, supported by Medical Research Scotland. In 2015, she repatriated with a Rutherford Discovery Fellowship.

Her broad research theme is understanding how variation present in our genomes makes people differ in their development and their lifetime health, with a focus on rare Mendelian human disorders. Her research is supported by the Marsden Fund and the University of Otago.

THE POWER OF A SINGLE PATIENT AND A SEQUENCING MACHINE: GENETIC INSIGHTS INTO HUMAN GROWTH

Genetic mutations occurring in humans are an unparalleled resource to utilise in human biology, and by studying these mutations, we can gain great insight into how organisms develop and what cellular pathways contribute to a disease phenotype. We have utilised exome sequencing to study a Mendelian disorder of growth, primordial dwarfism, characterised by extreme global growth failure. Our genetics research harnessing NGS technology has established how components of DNA replication and DNA repair molecular pathways contribute to organism growth. Of particular excitement is the impact that single patients, with a novel mutation illuminated by NGS technology and a unique set of clinical features, can have to dramatically advance our research and understanding.

1015-1030

Dr Margaret Jordan
James Cook University, Australia

Dr Jordan (PhD in Immunogenetics) is a Senior Research Fellow, Director of Research (Molecular and Cell Biology) and Honours' Co-ordinator in the Comparative Genomics Centre at JCU, Townsville, Australia. She is a current recipient of the prestigious NHMRC/MSRA Betty Cuthbert Fellowship. Her research explores the genetics of autoimmune diseases, including Type 1 Diabetes, Multiple Sclerosis, Type A Gastritis and Systemic Lupus Erythematosis, both in humans and in mouse models of disease. She is the recipient of grants from the National Health and Medical Research Council (NHMRC), Multiple Sclerosis Research Australia (MSRA), Rebecca L. Cooper Foundation, Rising Stars Program, JCU, the Logan Foundation, Australasian Society for Immunology (ASI) and Juvenile Diabetes Research Foundation (JDRF). Dr Jordan is a current committee member of the Queensland branch of ASI and has editorial responsibilities for scientific journals and reviews grant applications for NHMRC, Diabetes Research Australia Trust (DART), Logan Foundation, Rebecca L. Cooper Foundation and MSRA. She also reviews grant applications for Competitive JCU Scholarships as well as for the Graduate Research Scheme Grants-in aid Conference Funding (College of Public Health, Medical and Veterinary Sciences, JCU, Townsville, Australia). She has published 14 primary manuscripts, 9 reviews and 7 abstracts as well as a book chapter in "The Autoimmune Diseases", a well-regarded comprehensive work for basic and clinical scientists.

UNRAVELLING THE MULTIPLE SCLEROSIS COMPLEX DISEASE TRAIT THROUGH AN IMMUNE TRANSCRIPTIONAL REGULATORY NETWORK APPROACH

Dr Margaret Jordan[1], Dr Melissa Gresle[2], Dr Louise Laverick[2], Dr Dragana Stanley[3], Ms Letitia Smith[1], Dr Tim Spelman[2], Dr Judith Field[4], Dr Laura Johnson[4], Ass/Professor Helmut Butzkueven[2], Professor Alan Baxter[1]

1. Comparative Genomics Centre, James Cook University
2. The Department of Medicine, University of Melbourne
3. Central Queensland University
4. The Florey Neuroscience Institute

Multiple Sclerosis (MS), the most common disabling neurological disease affecting young adults in developed countries, is a complex genetic disease associated with both environmental and genetic risk factors. In most cases, the risk factors' individual associations with MS are so weak that any meaningful understanding of the disease will require the identification of molecular pathways that contribute to MS liability. We therefore hypothesised that the complex genetic phenotype is driven by a co-ordinated expression of transcriptional regulatory networks. To test this, we generated a weighted gene co-expression network based on 712 pooled Affymetrix Human Gene 1.0 ST array analyses of magnetic bead sorted B cells, CD4 and CD8 T cells, NK cells and monocytes, from 67 untreated relapsing/remitting MS patients and 102 Healthy Controls (HC). Sixteen relatively independent gene modules were identified. For each leukocyte population, the strength of differential expression between patients and HC was assessed, by ranking genes by Mann Whitney U test and ANOVA, and each transcript was tested across the network to identify modules of interest. A major networked module of genes involved in cell-mediated cytotoxicity that is downregulated in the monocytes of MS patients was identified. While monocyte-derived cytokines have been reported to persist throughout disease progression, a role for monocytes in immunomodulation by cytotoxicity is innovative. A dissection of this "killer module" and the elucidation of its role in MS and EAE may thus provide a new perspective on the aetiology of, not only MS but, other complex autoimmune diseases and thereby offer novel interventions.

1030-1045

Dr Thomas Ohnesorg
Murdoch Childrens Research Institute, Australia

Thomas Ohnesorg received his PhD from the Technical University of Munich, Germany and is currently working as a Postdoctoral Research Fellow in the lab of Professor Andrew Sinclair at the Murdoch Childrens Research Institute in Melbourne, Australia. His research focuses on genes involved in mammalian sex determination, differentiation, and development, in particular how these genes are regulated at the transcriptional level and their role in the regulation of other genes.

IDENTIFICATION OF THREE NOVEL ENHANCERS THAT REGULATE HUMAN SOX9 IN THE GONAD: IMPLICATIONS FOR DISORDERS OF SEX DEVELOPMENT.

Thomas Ohnesorg[1], Jo Bowles[2], Peter Koopman[2], Andrew Sinclair[1,3]

1. Murdoch Childrens Research Institute, Melbourne, VIC, Australia
2. Institute for Molecular Bioscience, University of Queensland, Brisbane, QLD, Australia
3. Department of Paediatrics, University of Melbourne, VIC, Australia

Disorders of Sex Development (DSD) encompass a wide spectrum of conditions that most commonly manifest with atypical gonads or genitalia. DSDs represent a significant healthcare issue due to the direct clinical management of these conditions and also their common association with gonadal cancer and infertility. The underlying cause is often a breakdown in the complex network of genes that regulate the development of testes or ovaries. The majority of DSD patients cannot be given an accurate diagnosis, severely compromising their clinical management. While mutations in coding regions of gonad genes have been important in understanding the etiology of some DSD, little attention has been focussed on the regulatory regions of these genes.

We and others have identified 46,XX testicular DSD (female-to-male sex reversal) patients with duplications and 46,XY gonadal dysgenesis (male-to-female sex reversal) patients carrying deletions upstream of SOX9 that indicate the presence of gonad specific enhancers. Using a comprehensive tiling luciferase assay and bioinformatic approaches we have identified three novel enhancers within this region. Enhancers that showed the strongest activity in vitro were used to generate transgenic mice. The enhancers showed expression in embryonic mouse gonads at the time of gonad differentiation. Furthermore, there appears to be a sequential regulatory cascade: one enhancer requires the binding of SF1 and SRY to initiate SOX9 up-regulation, while the second requires SF1 and SOX9 binding to further increase SOX9 expression while the final enhancer is auto-regulated by SOX9 to maintain high levels of expression in the testis. Our results strongly suggest that CNVs affecting these novel enhancers are causative for DSD

SESSION 5: GENOME REGULATION: BIOLOGY, PHENOTYPE AND UTILITY

CHAIRS: DR MILES BENTON & DR DONIA MACARTNEY-COXSON

1115-1200

PROFESSOR GREGORY GIBSON
Georgia Institute of Technology, U.S.A

Sponsored by ESR

A graduate of Sydney University undergraduate and the University of Basel PhD, followed by a Post-Doc at Stanford, Greg has been at a Professor in the School of Biology at Georgia Tech since 2009.

View his full CV here: http://www.agtaconference.org/wp-content/uploads/2016/01/gibson_cv_jan2016.pdf

TRANSCRIPTOMICS FOR PERSONALIZED MEDICINE

While exome and genome sequencing are quickly finding their way into precision medical diagnosis of the causes of congenital abnormalities and tumor progression, applications of transcriptome profiling for personalized medicine are just beginning to emerge. I will discuss some theoretical aspects of how expression quantitative trait locus (eQTL) and transcriptional risk score (TRS) analyses have the potential to improve on genome sequencing for molecular pathology, then present three case studies: longitudinal profiling of peripheral blood in autoimmune disease, ileal biopsies to define high and low risk patients for Crohns Disease progression, and targeted RNASeq for discovery and/or validation of neuromuscular disease mechanisms.

1200 - 1230

Dr Logan Walker
University of Otago, New Zealand

Dr Walker's primary research is focused on developing better methods for identifying individuals at high-risk of breast and ovarian cancer, and understanding how genetic variations cause an increased risk of cancer in these individuals. Utilising genomic- and transcriptomic-based technologies, we are uncovering new genetic and molecular variation that contributes to tissue-specific cancer risk. This research is being carried out in collaboration with international cancer genetics-based consortia, including CIMBA (The Consortium of Investigators of Modifiers of BRCA1/2) and ENIGMA (Evidence-based Network for the Interpretation of Germline Mutant Alleles). View more here: http://www.otago.ac.nz/christchurch/research/mackenzie/ourpeople/otago011991.html).

UTILISING GENOMIC TECHNOLOGIES TO EVALUATE GERMLINE VARIANTS OF UNKNOWN CLINICAL SIGNIFICANCE

Routine diagnostic BRCA1 and BRCA2 gene screening for pathogenic variants is typically performed for individuals from suspected high-risk breast-ovarian cancer families. However, for most breast and ovarian cancer patients, the genetic change(s) contributing to their disease remain poorly understood. Massively parallel sequencing technologies are now being adopted by research and diagnostic laboratories worldwide, enabling screening of BRCA1 and BRCA2, and other cancer related genes in a greater number of people. Determining the clinical meaning of newly discovered variants is a central challenge for genomic health.

Continued next page

To address this challenge for high-risk breast-ovarian cancer families, the international ENIGMA (Evidence-based Network for the Interpretation of Germline Mutant Alleles) consortium was founded in 2009 and is now considered an expert panel by ClinVar. ENIGMA has received >6000 submissions of BRCA1/2 variants of unknown clinical significance (>3000 unique variants) from laboratories worldwide. I will highlight how new and old genomic approaches are being implemented by ENIGMA to help determine the pathogenicity of these variants.

1230 – 1300

PROFESSOR ANDREW HILL
La Trobe Institute for Molecular Science,
La Trobe University, Australia

Professor Andrew Hill gained his BSc(Hons) in Biochemistry and Molecular Biology from Victoria University of Wellington in New Zealand and his PhD at Imperial College, London. He held post-doctoral positions in the MRC Prion Unit (London) and in the Department of Pathology at the University of Melbourne as a Wellcome Trust Prize Travelling Research Fellow. Andrew joined the Department of Biochemistry and Molecular Biology at the University of Melbourne in 2002 and moved his lab into the Bio21 Institute when it opened in 2005. In 2015, he moved his laboratory to the La Trobe Institute of Molecular Sciences (LIMS) at La Trobe University where he is also Head of the Department of Biochemistry and Genetics. Andrews research team investigates the mechanisms involved in the spread of misfolded proteins associated with neurodegenerative disorders such as Alzheimer's, Parkinson's and prion diseases. His laboratory studies the role that extracellular vesicles such as exosomes play a role in this process. Over the last five years his laboratory has developed an interest in extracellular vesicle RNA and investigated its potential as disease biomarkers for neurological and other conditions. He has been the recipient of several awards and prizes including a Victorian Young Tall Poppy Award in 2006, and the Edman Award (2005) and, in 2010, the Merck Research Excellence Medal from the Australian Society for Biochemistry and Molecular Biology. He has held a number of competitive fellowships including an ARC Future Fellowship (FT3) and NHMRC Senior Research Fellowship.

In 2016, Andrew was elected President of the International Society of Extracellular Vesicles (ISEV).

EXTRACELLULAR VESICLES – THEIR ROLE IN NEURODEGENERATIVE DISEASES AND POTENTIAL SOURCE OF RNA DISEASE BIOMARKERS

Neurodegenerative disorders such as Alzheimer's (AD), Parkinson's (PD) and prion diseases are associated with proteins that misfold and deposit in the brain. Many cell types, including neurons, release extracellular vesicles (EVs) which include microvesicles and exosomes. Roles for these vesicles include cell-cell signalling, removal of unwanted proteins, and transfer of pathogens (including prion-like misfolded proteins) between cells. In addition to their protein content these vesicles have recently been shown to contain genetic material in the form of protein coding (mRNA) and noncoding RNA species including miRNAs. We have analysed the protein and genetic cargo of EVs from a number of cell types and using deep sequencing, characterised the RNA cargo of these vesicles. As exosomes can be isolated from circulating fluids such as serum, urine, and cerebrospinal fluid (CSF), they provide a potential source of biomarkers for neurological conditions. This talk will review the roles these vesicles play in neurodegenerative disease and highlight their potential in diagnosing these disorders through analysis of their RNA content.

1300 – 1315

Dr Stuart Archer
Monash Bioinformatics Platform, Australia

Dr. Stuart Archer has focused on RNA biology for most of his postdoctoral career, more recently moving into the bioinformatics realm. His first postdoctoral project, undertaken in Heidelberg, Germany, examined mRNA regulation in parasitic protists, where he uncovered a network of cell-cycle regulated genes controlled post-transcriptionally by PUF RNA-binding proteins. He returned to Australia in 2010 to study basic aspects of eukaryotic RNA biology, such as 5'-3' interactions and translation initiation, at the Australian National University in Canberra. He began to develop more data analysis skills to take greater advantage of RNA-seq and other new high-throughput technologies, and now works as a senior bioinformatician at the Monash Bioinformatics Platform in Melbourne.

MECHANISMS AND REGULATION OF MRNA TRANSLATION INITIATION REVEALED BY TCP-SEQ

Stuart K. Archer[1], Nikolay E. Shirokikh[2], Traude H. Beilharz[3] and Thomas Preiss[2]

1. Monash Bioinformatics Platform, Monash University, Melbourne, Australia
2. The John Curtin School of Medical Research, The Australian National University, Canberra, Australia
3. Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Australia

Eukaryotic translation initiation is a fundamental component of the central dogma of molecular biology, and an important point of gene expression regulation, however there are large blind spots in our knowledge of both the basic mechanism and the points in the process at which regulatory intervention can influence gene expression. I will describe the first results from TCP-seq, a novel variant of RNA-seq that gives a transcriptome-wide snapshot of translation initiation events. By analysing the position and length of RNase-protected 'footprints' left behind by the small ribosomal subunit, which is the first ribosomal subunit to bind mRNA, general features of initiation and start-codon recognition can be discerned. Transcript-specific departures from the normal initiation processes are also apparent in many cases, indicative of cis-regulation by 5' UTR elements. These results underpin mechanistic models of translation initiation and termination, built on decades of biochemical and structural investigation, with direct in vivo evidence on a transcriptome-wide scale.

1315 – 1330

Dr Sam Buckberry
The University of Western Australia, Australia

Sam Buckberry is a genome and computational biologist in the Lister Lab at The University of Western Australia. His research is focused on deciphering the role of the epigenome in neuronal gene regulation and stem cell reprogramming by combining cutting-edge DNA sequencing with integrative bioinformatics and statistical methods.

CHARACTERISING EPIGENOME DYNAMICS DURING THE REPROGRAMMING OF SOMATIC CELLS TO IPS CELLS

Sam Buckberry[1], Anja Knaupp[2], Jose Polo[2], Ryan Lister[1]

1. The University of Western Australia, Perth, WA, Australia
2. Monash University, Melbourne, VIC, Australia

Reprogramming of somatic cells into induced pluripotent stem cells (iPSCs) can be achieved through the forced expression of defined transcription factors. This induction of pluripotency requires global change of the somatic cell epigenome patterns into a state similar to embryonic stem cells. It has become apparent that epigenetic mechanisms play important roles in cellular reprogramming. However, the reprogramming process is inefficient and only a minority of cells complete the transition to pluripotency, which has been a limiting factor in studying the epigenomic changes that occur during reprogramming.

To further understand the epigenetic changes that occur during cellular reprogramming, we isolated sub-populations of reprogramming intermediates that were poised to become iPSCs and performed time-course genome-wide profiling of reprogramming transcription factor binding (ChIP-seq), chromatin accessibility (ATAC-seq), gene expression (RNA-seq) and DNA methylation (MethylC-seq) throughout the whole reprogramming process. Our results show that induction of reprogramming factors initiates an early wave of widespread chromatin reconfiguration coupled with transient gene expression change and transcription factor binding.

A second wave of reconfiguration occurs just prior to cells making the final transition to the pluripotent state, where a majority of early wave chromatin states are reverted and the majority of iPSC factor binding sites and chromatin states become established. The first wave of change is coincident with gradual change in DNA methylation, with the second wave of reconfiguration being coupled with widespread reduction in DNA methylation as the cells acquire pluripotency.

This integrative time-course analysis of cellular reprogramming provides a comprehensive molecular picture of the genomic changes that occur during reprogramming.

SESSION 6: GENOMIC DATA INVESTIGATION

CHAIRS – DR AUSTEN GANLEY & DR GRETCHEN POORTINGA

1500 – 1545

PROFESSOR PIERO CARNINCI
RIKEN Center for Life Science Technologies, Japan

Sponsored by Dnature and Gene Target Solutions

Born and Educated in Italy Professor Piero Carninci obtained his doctoral degree at the University of Trieste in 1989. From 1990 to 1995 he developed technologies for DNA extraction and DNA sequencing at a start-up company, Talent Srl.

Carninci moved to Japan in 1995 at RIKEN, Tsukuba Life Science Center as a postdoc to become soon a tenured researcher (1997) working to develop technologies to capture full-length cDNAs, which were later broadly used in projects like FANTOM.

From 2008 he became a Team Leader, a Unit Leader and a Deputy Project Director at the RIKEN Omics Science Center in Yokohama.

Carninci developed technologies to analyze the transcriptome, including the CAGE technology. These technologies have been broadly used in FANTOM project, allowing identification of non-coding RNAs as the major output of the mammalian genome and providing the most comprehensive maps of the mammalian transcripts and promoters, broadly used in the ENCODE and the FANTOM projects.

From April 2013, he became the Director of the Division of Genomic Technologies, Deputy Center Director at RIKEN Center for Life Science Technologies.

In the new division, Carninci is coordinating research teams developing the next generation of technologies to broadly study life sciences. They aim to address biological problems that could not be studied before, such as the study of homogeneous or heterogeneous cell populations.

He has published more than 200 papers.

EMERGING LANDSCAPE OF REGULATORY RNAS

In the Functional Annotation of the Mammalian Genome (FANTOM) 5 project, we have applied cap analysis of gene expression (CAGE) technology, which simultaneously maps mRNAs and long non-coding RNAs (lncRNAs) transcription starting sites (TSSs) and measure their expression at each promoter, on a comprehensive panel of human/mouse primary cells and other tissues, resulting in a broad map of the promoterome/regulatory networks1. Our map revealed the existence of 201,802 and 158,966 promoters and 65,423 and 44,459 enhancers, in human and mouse respectively, which are often tissue specific. The project also revealed complexity of genome activation hierarchy in which transcription initiates with enhancers followed by promoters and then other genes in high density of time course expression analysis profiling 33 mouse/human biological systems2. The FANTOM5 database is one of the broadest expression database available to the community (http://fantom.gsc.riken.jp/5/). Additionally, we have focused on nuclear RNAs and determined the pattern of expression of retrotransposon elements (RE), which are likely to have a regulatory role. As example, some families of LTR retrotransposon elements are specifically expressed in ES and iPS cells, where they have a role in maintenance of pluripotency3.

Ongoing FANTOM6 project is focusing on broad understanding for the function and the interaction with cell regulatory networks of these RNAs in several primary cells, with the purpose to create the broadest database of functional lncRNAs.

References:
1. Forrest et al., Nature 507, 462-470 (2014)
2. Arner et al., Science 347, 1010-1014 (2015)
3. Fort et al., Nature Genetics 46, 558-566 (2014)

1545 – 1615

PROFESSOR CHRISTINE WELLS
The University of Melbourne, Australia

Professor Christine Wells is an internationally recognised pioneer of genomics in its application to innate immunity and stem cell biology. She has driven programs to identify the genetic elements that define cellular states, contribute to the regulation of immune genes and describe the functions of new gene products. During the past two decades Professor Wells has made key contributions to several seminal papers that mapped out mammalian genome architecture and transcriptional complexity. Her laboratory focuses on data integration and meta-analysis, working to build platforms for collaboration and visualisation of multi-omic data for the stem cell community. Her laboratory works on computational and statistical method development, as well as projects aimed at understanding the function of innate immune genes in tissue injury and repair. In 2011, she established Stemformatics.org – a collaborative hub for Australian and international stem cell researchers. She is Director of the Centre for Stem Cell Systems at Melbourne University, the computational PI in Stem Cells Australia and leads the multi-omics stem cell program for BioPlatforms Australia.

USING THE STEMFORMATICS STEM CELL ATLAS TO MINE STABLE SIGNATURES OF STEM CELL IDENTITY

Genotyping-by-Sequencing (GBS) is a high throughput method for generating many thousands of genetic markers using next generation sequencing. The method was developed with simplicity of deployment and an open source philosophy in mind. We published the method in 2011 in an open access journal so that as many researchers as possible could benefit from its use and potentially modify or extend it. The method has been adopted by many researchers. In addition to plant breeders in commodity species, GBS is being used by breeders of orphan crops as well as in the ecological sciences. Simplicity, low cost and openness of the method have allowed many researchers to jump into the genomics era and leveled the research playing field. However, more than a good molecular method is needed to democratise genetic analysis.

The Biospectra by Sequencing project aims to develop a set of community resources. There are two core components. One is an interactive information repository to contain useful information that is not often published in the literature. The other is a Free / Libre Open Source Software (FLOSS) project developing the software tools for best practice in data quality, automation and reproducibility. Working together, we will reduce duplication of efforts, provide a high quality set of information and tools and further reduce the barriers to entry in genomics research. We aim to enable researchers to do better research at a lower price than previously possible.

1615 – 1645

DR AUSTEN GANLEY
University of Auckland, New Zealand

Austen Ganley is a Senior Lecturer in the School of Biological Sciences at University of Auckland. He has wide-ranging interests in genomics, and his work is characterised by the development of novel techniques to address fundamental questions in biology. He is an internationally recognised expert in the genomics of the ribosomal RNA gene repeats in eukaryotes, where he has made seminal advances in the relationship between transcription and recombination, and in the evolutionary dynamics of the ribosomal RNA gene repeats. His research involves a combination of bioinformatics and experimental approaches, and uses a variety of model systems, including yeast, filamentous fungi and mammalian systems.

THE GENOMIC RESPONSES TO POLYPLOIDY IN EUKARYOTES

Austen R. D. Ganley[1], Matthew A. Campbell[2], and Murray P. Cox[2]

1. School of Biological Sciences, University of Auckland, Auckland, New Zealand
2. Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand

Continued next page

Polyploidy, the increase of entire genome complements, has played a role in the evolution of many eukaryote lineages. It has been proposed to drive evolutionary diversification by creating genetic redundancy, and is associated with increased size and growth rate that has frequently led to domestication. However, the merger of two or more dissimilar genome sets, as occurs in allopolyploidy, also results in a variety of deleterious effects, collectively known as genome shock, and frequently involves loss of duplicated genetic material. Understanding the genomic responses that contribute to the beneficial and detrimental effects of polyploidy has been limited by a lack of tools. To overcome this limitation, we have developed computational approaches that allow us to dissect the fates of the parental genomes. We have used these approaches to investigate the genomic responses to allopolyploidy in a group of fungi that include many naturally allopolyploid species, the Epichloë endophytes. I will present data detailing the transcriptional responses to allopolyploidy, and imbalances resulting from differential loss of genomic regions. I will also present these results in a comparative framework with other polyploid species. This work provides a platform for developing a better understanding of the important yet elusive responses of genomes to polyploidy across different systems.

1645–1700

Beth Signal
Garvan Institute, Australia

Beth is a 2nd year PhD student. Her research involves using bioinformatic methods to understand genome biology. In particular, she is interested in applications of machine learning and temporal RNA-Sequencing to decipher functions of non-coding regions of the transcriptome contributing to cellular differentiation.

COMPREHENSIVE MACHINE LEARNING ANNOTATION OF HUMAN SPLICING BRANCHPOINTS

Beth Signal[1,2], Brian S Gloss[1,2], Marcel E Dinger[1,2], Tim R Mercer[1,2]

1. Genomics and Epigenetics, Garvan Institute of Medical Research, Sydney, Australia
2. St Vincent's Clinical School, University of New South Wales, Sydney, Australia

The branchpoint is a basal genetic element required for gene splicing. Despite a primary role in exon inclusion, current annotations of branchpoints are incomplete and limited to experimental catalogues. Due to difficulty in experimentally identifying branchpoints compared to other splicing elements, their contribution to normal and alternative splicing has been largely understudied. We have developed a machine-learning algorithm, using the most recent gold-standard human branchpoint annotations, to identify branchpoints from gene sequence alone. Using this approach, we are able to locate branchpoints in 85% of introns in current gene annotations. This near-complete annotation is unbiased towards gene type and expression levels, highlighting the advantage of developing predictive models in genome annotation. Several introns were found to encode multiple branchpoints, which may be a mechanism through which mutational redundancy is encoded in key genes. Branchpoint strength is associated with differing modes of alternative splicing, and has a distinct contribution to other splice elements. Notably, this annotation constitutes an invaluable resource for interpreting the mechanistic impact of common- and disease-causing human genetic variation on gene splicing.

1700-1715

Dr Joseph Powell
The University of Queensland, Australia

Dr Joseph Powell is a group leader at the Institute for Molecular Bioscience (IMB), University of Queensland. His lab develops and applies computational genomics approaches to investigate the genetic control of genome regulation and its role in contributing to the susceptibility to disease. He obtained his Ph.D. from the Roslin Institute, University of Edinburgh in 2010 and subsequently moved to Brisbane to work with Professor Peter Visscher, first as a research fellow and then as a team leader. During this time, he co-led the formation of the Consortium for the Architecture of Gene Expression, an international effort to share raw-level transcriptomic and genomics data. Joseph has worked on a range of research projects involving methods, theory, and application around the nexus of quantitative, statistical and population genetics. Joseph was awarded the 2016 Commonwealth Health Ministers Medal for excellence in Medical Research and the 2016 top CJ Wright CD Fellowship.

SYSTEMATIC EVALUATION OF GENETIC CORRELATIONS BETWEEN EXPRESSED TRANSCRIPTS IN PERIPHERAL BLOOD

Joseph Powell[1], Samuel Lukowski[1]

1. Institute for Molecular Bioscience, University of Queensland

Establishing the proportion of shared genetic control between expressed transcripts can be used to help identify pathways and genetic mechanisms that underlie disease susceptibility or severity. Using a bivariate GREML approach, we estimated the genetic correlations between each pairwise combination of 2,469 transcripts that are both highly heritable and expressed in whole blood in a cohort of 1,748 unrelated individuals of European ancestry. We identified 556 pairs with a significant genetic correlation at a Bonferroni study-wide threshold, of which 77% were located on different chromosomes to one another. Using eQTL data from an independent cohort (n=2,112) we subsequently identified 934 incidences where the eSNP for one probe had significant ($p<4.1*10-8$) shared effect on the paired probe, providing further verification of our results.

We investigated the genetic regulatory mechanisms underlying the co-regulation of highly correlated transcript pairs. Our findings reveal (i) a significant enrichment of highly interconnected transcription factors ($p=3.43*10-25$), (ii) shared eSNP-mediated transcriptional regulation and (iii) significant enrichment of genetically correlated transcripts pairs and regions of chromatin interaction (empirical $p<0.001$). We used estimates of genetic correlations to construct graph networks of interconnected transcripts, which revealed the direction of shared genetic control between transcript isoforms, and their correlated transcripts. Our findings demonstrate the utility of using genetic correlations to investigate transcriptional co-regulation and to gain valuable insight into the nature of the underlying genetic architecture of regulation. All results are made publicly available at http://computationalgenomics.com.au/shiny/rg/

1715–1730

Professor Vanessa Hayes
Garvan Institute of Medical Research,
Australia

NEXT GENERATION MAPPING REVEALS NOVEL LARGE GENOMIC REARRANGEMENTS IN PROSTATE CANCER

Complex genomic rearrangements are common molecular events driving prostate carcinogenesis. Clinical significance, however, has yet to be fully elucidated. Detecting the full range and subtypes of large structural variants (SVs), greater than one kilobase in length, is challenging using clinically feasible next generation sequencing (NGS) technologies. Next generation mapping (NGM) is a new technology that allows for the interrogation of megabase length DNA molecules outside the detection range of single-base resolution NGS. In this study, we sought to determine the feasibility of using the Irys (BioNano Genomics Inc.) nanochannel NGM technology to generate whole genome maps of a primary prostate tumor and matched blood from a Gleason score 7 (4+3), ETS-fusion negative prostate cancer patient. With an effective mapped coverage of 35X and sequence coverage of 60X, and an estimated 43% tumor purity, we identified 85 large somatic structural rearrangements and 6,172 smaller somatic variants, respectively. The vast majority of the large SVs (89%), of which 73% are insertions, were not detectable ab initio using high-coverage short-read NGS. However, guided manual inspection of single NGS reads and de novo assembled scaffolds of NGM-derived candidate regions allowed for confirmation of 94% of these large SVs, with over a third impacting genes with oncogenic potential. From this single-patient study, the first cancer study to integrate NGS and NGM data, we hypothesise that there exists a novel spectrum of large genomic rearrangements in prostate cancer, that these large genomic rearrangements are likely early events in tumorigenesis, and they have potential to enhance taxonomy.

SESSION 7: POPULATION GENETICS

CHAIRS – PROFESSOR NEIL GEMMEL & PROFESSOR LISA MATISOO-SMITH

0900-0945

ASSOCIATE PROFESSOR BETH SHAPIRO
University of California Santa Cruz, U.S.A.

Sponsored by Biomatters

geneious

Beth Shapiro is an evolutionary biologist who specializes in the genetics of ice age animals and plants. A pioneer in the young field called "ancient DNA," Beth travels extensively in the Arctic collecting bones and other remains of long-dead creatures including mammoths, giant bears, and extinct camels. Using DNA sequences extracted from these remains, she hopes to better understand how the distribution and abundance of species changed in response to major climate changes in the past. The results could be used to help develop strategies for the conservation of species that are under threat from climate change today. Associate director of the University of California Santa Cruz Genomics Institute and a research associate of the Denver Museum of Natural History, Shapiro has been widely honored for her research. She has been named a Royal Society University Research Fellow, Searle Scholar, Packard Fellow, and a National Geographic Emerging Explorer. In 2009, she received a MacArthur "genius" award. Her recent book, "How to Clone a Mammoth: The Science of De-extinction," won the 2016 AAAS/Subaru Prize in nonfiction science.

PASSENGER PIGEON PALEOGENOMES REVEAL THE GENOMIC CONSEQUENCES OF LONG-TERM EXTREMELY LARGE EFFECTIVE POPULATION SIZES

Shapiro, B, Green, RE, Murray G, Soares AER

The passenger pigeon was once the most abundant bird in North America, with flocks that, during the early and middle 19th century, were estimated to comprise up to three billion individuals. Less than 100 years later, however, passenger pigeons were extinct. To explore the genomic consequences of such a rapid population decline, we sequenced and assembled high-quality genomes from five passenger pigeons and two band-tailed pigeons (the closest living relative of passenger pigeons) and mitochondrial genomes from 41 passenger pigeons. The mitochondrial indicate that passenger pigeon populations had been large for at least the last 40,000 years, and were not experiencing a decline at the time of their extinction. The nuclear data reveal a striking, bimodal distribution of genetic diversity, confounding demographic inference from these data. We hypothesize that this pattern is a consequence of their long-term large effective population size and the consequent dominance of natural selection as an evolutionary force.

0945 – 1015

DR NICOLAS DUSSEX
University of Otago, New Zealand

Nic received his PhD from the University of Otago and his research interests pertain to conservation genetics, population dynamics and history and evolution.

Nic's work in the Gemmell lab focuses on the genomic basis of complex cognition in New Caledonian Crows (Corvus moneduloides). This research involves comparisons of multiple crow species genomes and the identification of genomic regions under selection.

His past research focused on the population histories of the endangered kea and kakapo, kaka population structure and translocations and the evolution of flightloss in alpine stoneflies.

GENOMIC BASIS OF TOOL MANUFACTURE AND USE IN NEW CALEDONIAN CROWS

Nic Dussex[1], R. Axel W. Wiberg[2], Verena E. Kutschera[3], Gavin Hunt[4], Russell D. Gray[4,5], Robert Fleischer[6], Christian Rutz[2], Michael G. Ritchie[2], Jochen B.W. Wolf[3], Neil J. Gemmell[1]

1. University of Otago, Dunedin, New Zealand,
2. University of St Andrews, St Andrews, UK,
3. University of Uppsala, Uppsala, Sweden,
4. University of Auckland, Auckland, New Zealand,
5. Max Planck Institute for Science of Human History, Jena, Germany,
6. Smithsonian Conservation Biology Institute, Washington, DC, USA

The discovery of avian cognitive abilities has revolutionised our understanding of the evolution of intelligence. However, the genetic basis of these abilities is unknown. One possibility is that numerous changes in genes across a wide range of functional domains are required for the evolution of complex intelligence. Alternatively, only a limited number of genetic tweaks might be required. Corvids are well known for their cognitive abilities such as episodic-like memory, problem solving, and tool use.

The New Caledonian crow (Corvus moneduloides) is particularly intriguing as it is one of the few non-human species to manufacture foraging tools, making it an ideal model to study the genetic basis of cognition. Here we present genome-wide (~18,000 protein coding genes) phylogenetic comparisons among 12 crow species including C. moneduloides and scrutinize the genome for signatures of selection. In order to test whether similar genetic changes may have arisen earlier in the evolutionary history of the lineage, we perform the same analysis focusing on the closely-related tropical but non-tool using white-billed crow (Corvus woodfordi). We then use transcriptome data from C. woodfordi and avian protein databases to link candidate genes under selection to the species biology. We expect to detect signatures of positive selection (i.e. dN/dS >1) in genes associated with brain function and bill morphology allowing tool use and manufacture. Our results will help uncover the evolution and genetic basis of cognition in the wild and will reveal the nature of changes required to evolve cognitive abilities.

1015 – 1045

PROFESSOR LISA MATISOO-SMITH
University of Otago, New Zealand

Lisa Matisoo-Smith is the Professor of Biological Anthropology in the Department of Anatomy at the University of Otago, and a Principal Investigator with National Geographic's Genographic Project and Director of the Allan Wilson Research Theme at Otago. She is interested in understanding human history and human variation in general, but her primary area of interest is in looking at the biological evidence for the human settlement of the Pacific. She applies both ancient and modern DNA techniques to reconstruct migration pathways, and is increasingly interested in how human history and population origins may be related to some of the health issues facing Pacific populations today. Lisa works throughout the Pacific, but her most recent work has focused on New Zealand, Tokelau, Papua New Guinea and South America. She is committed to working in close collaboration with the communities she studies and strongly believes that such relationships benefit the research in ways not often appreciated.

RECONSTRUCTING THE SETTLEMENT OF EAST POLYNESIA

Lisa Matisoo-Smith[1], Catherine Collins[1] and Anna Gosling[1,2]

1. Department of Anatomy, University of Otago
2. Department of Human Genetics, University of Chicago

The Polynesian Triangle was the last region to be settled by humans, with the settlement of East Polynesia occurring only within the last 1200 years. Genetic data have been used to identify population origins and reconstruct the process of Pacific settlement, yet most studies have focused on populations in the Western Pacific, with few East Polynesian samples included. Our previous studies on ancient and modern populations of commensal animals indicated that the settlement of Polynesia was perhaps not as simple as traditional models suggested. Our preliminary mitochondrial genome data from modern human populations from Polynesia raised similar issues and suggest that there was significantly more variation in East Polynesia than previously thought.

The current data for mitochondrial genome variation in Polynesia raise key questions regarding population origins. Here we will present both ancient and modern mitochondrial genomes from Polynesian and other Pacific populations, compare these to current models of population origins and discuss possible settlement scenarios for New Zealand and more widely in East Polynesia and the Pacific.

1045 – 1100

Jac Charlesworth
Menzies Institute for Medical Research, Australia

Dr Charlesworth is a statistical geneticist, experienced at working with a wide range of genomic data and complex phenotypes. She obtained her PhD in statistical genetics from the University of Tasmania in 2006 working on the genetics of Tasmanian glaucoma families. She then spent four years as a postdoctoral fellow at the Texas Biomedical Research Institute in the USA, training in computational genomics with Profs. John Blangero and Laura Almasy. In April 2010 she returned to Australia to start a computational genomics group at the Menzies Institute for Medical Research. She built and now directs the Menzies Genomics Computing Cluster which supports all genomic research at Menzies. Her research into complex disease genomics, particularly using family-based study designs, sits at the strategic interface between biology and informatics.

FAMILY VALUES: SEQUENCING LARGE FAMILIES AS AN ENRICHMENT STRATEGY FOR RARE VARIANTS

Jac Charlesworth[1], Nicholas Blackburn[1,2], Juan Manuel Peralta[1,2]

1. Menzies Institute for Medical Research, University of Tasmania, Australia
2. South Texas Diabetes and Obesity Institute, University of Texas Rio Grande Valley, USA

Continued next page

There is growing interest surrounding the role of rare variants in complex trait genetics. Large, population-based sequencing studies such as 1000 Genomes (1KGP) and the Exome Aggregation Consortium (ExAC) continue to demonstrate that the majority of human variation is rare or even private. 54% of variants from 60,706 exomes in ExAC are singleton variants with a staggering 72% absent from both 1000 Genomes and the Exome Sequencing Project databases (ExAC et al. bioRxiv 2016). As such, study design for variant enrichment is critical, in order to obtain sufficient copies of a variant to perform appropriate statistical tests against a trait of interest. GWAS has been successful in this arena for common variant associations but requires exponentially larger sample sizes for rarer variation.

In this study we sequenced 252 individuals from five extended pedigrees to demonstrate the distribution of rare, private, and also predicted deleterious variants as well as the enrichment of these classes of variants using a family-based study design.

Exome sequencing was conducted on the Illumina HiSeq platform using Nextera Expanded Exome enrichment; providing 62Mb of genomic content including exons and additional regulatory and non-coding regions. Sequence data were aligned, called and cleaned using an in-house pipeline, with BWA-MEM and the Churchill pipeline (Kelly et al. Genome Biology 2015) as the foundations, taking advantage of the pedigree structures throughout the calling and cleaning process. For this enrichment study we included only bi-allelic SNPs from autosomes for which the sequence data was of high confidence and quality. We used these data to determine the number and distribution of 'novel' variants not present in any population sequencing database (ie ExAC). We also calculated the distribution in our families of variants present in both ExAC and the UK10K databases at frequencies <1% – demonstrating clear enrichment of variants in the pedigrees that are rare in the general population. In addition, we used in silico tools to predict the functional impact of these variants and were able to show enrichment of rare and predicted functional variant classes – present in the families at minor allele copies numbers sufficient for statistical association with traits of interest.

These data demonstrate that family-based sequencing studies represent an efficient enrichment strategy for rare and predicted functional variants. Appropriate study design is still required to maximise these advantages in a disease or trait association context.

1100 – 1115

Dr Rachel Fleming
Institute of Environmental Science and Research Ltd, New Zealand

Rachel is a Science Leader in the Forensic Research and Development Team. Her main focus of research is in body tissue/fluid identification for forensic purposes using RNA. She is especially interested in new technology and how this can be applied to the challenging nature of forensic samples. Alongside leading the RNA body fluid project, Rachel also supervises PhD and MSc students.

RNA SEQUENCING IN FORENSIC SCIENCE

Rachel Fleming[1] and Meng-Han Lin[1]

1. Forensic Research and Development, Institute of Environmental Science and Research Limited, Mount Albert, Auckland, New Zealand

The dynamic nature of RNA expression in different tissues and cell types allows for messenger RNA profiling for the identification of forensically relevant body fluids. The very nature of forensic samples brings many challenges to RNA sequencing – from the way samples are collected at the crime scene through to the level of degradation and abundance of RNA. RNAseq has enabled the characterisation of the degradation state of transcripts targeted for body fluid identification. This new insight has enabled a novel approach to significantly improve the real world performance of RNA detection methods, regardless of RNA abundance and degradation levels. We have applied this knowledge to the detection of RNA from crime scene samples for forensic purposes, including the use of on-site rapid sequencing technologies.

SESSION 8: GENOME DIVERSITY AND EVOLUTION

CHAIRS – PROFESSOR LISA MATISOO-SMITH & PROFESSOR NEIL GEMMEL

1145 – 1230

DR RICHARD EDWARD GREEN
University of California Santa Cruz, U.S.A.

Sponsored by New Zealand Genomics Limited

Ed Green is an associate professor of biomolecular engineering in the Baskin School of Engineering. Prof. Green is perhaps best known for his work to determine the genome sequence of our closest evolutionary cousins, the Neandertals. Prof. Green began working on the Neanderthal genome project as a postdoctoral researcher at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany. He led a major bioinformatics effort, writing software and developing computational techniques needed to analyze ancient DNA extracted from fossil bones. Prof. Green's research now involves studying the genomic histories of many species, ranging from humans and other primates to alligators and birds. In addition to being widely honored as both a Searle Scholar and Sloan Fellow, Prof. Green is also a leading developer of biotechnology related to genome sequencing and assembly.

USING PROXIMITY LIGATION DATA FOR EFFICIENT DE NOVO GENOME ASSEMBLY

Long-range and highly accurate de novo assembly from short-read data is one of the most pressing challenges in genomics. We have developed a proximity-ligation method which uses in vitro assembled chromatin to scaffold long strands of DNA. In this approach, we generate connectivity information across a variety of genomic distances in a single library. These data can be used to scaffold contigs, dramatically increasing assembly contiguity. I will also show how this approach can be applied to metagenomic data from the human gut to increase assembly contiguity and resolve closely related strains.

1230 – 1300

DR CATHERINE GRUEBER
The University of Sydney, Australia

My passion is trying to understand how population genetics and evolutionary theory apply to the real-world management problems in threatened species conservation. By examining the genetic and fitness consequences of inbreeding, selection, and genetic drift in natural populations, I resolve to learn how we can better manage threatened species. I completed by PhD in 2010 at the University of Otago, where I studied the effects of inbreeding in a highly endangered bird, the takahe. Later, I conducted my first post-doc examining the roles of selection and drift (particularly on TLR immune genes) in 10 threatened birds, especially a bottlenecked population of NZ robin. In 2014, I joined Prof Kathy Belov's research group at the University of Sydney in a postdoctoral position sponsored by San Diego Zoo Global. We are using next-generation sequencing techniques to monitor and manage the processes that impact genetic diversity in captive and wild Tasmanian devil populations. I am pleased to work alongside conservation practitioners here in Australia and abroad, to help build creative questions and outcomes that influence both the conservation industry and the broader scientific community.

GENOMIC INSIGHTS INTO CANCER AND DIVERSITY IN TASMANIAN DEVILS: LESSONS FOR CONSERVATION

Catherine E. Grueber[1,2], Emma Peel[1], Rebecca Gooley[1], Carolyn Hogg[1,3], Katherine Belov[1]

1. School of Life and Environmental Sciences, Faculty of Veterinary Science, University of Sydney, NSW 2006, Australia
2. San Diego Zoo Global, PO Box 120551, San Diego, CA, 92112, USA
3. Zoo and Aquarium Association, Mosman, NSW 2088, Australia

Continued next page

The Tasmanian devil faces extinction in the wild due to a contagious cancer. Genetic and genomic technologies studies revealed that the disease arose in a Schwann cell of a female devil. Instead of dying with the original host, the tumour was passed from animal to animal, slipping under the radar of the immune system. Studying the genomes of the devil and the cancer has driven our understanding of this unique disease. From characterising immune genes and immune responses to studying tumour evolution, we have begun to uncover how a cancer can be 'caught' and are using genomic data to manage an insurance population of disease-free devils for the long-term survival of the species. Ongoing studies show how utilising genome resources enables us to target creative conservation-oriented questions that support captive breeding decisions, and population supplementation and reintroduction.

1300 - 1330

PROFESSOR NEIL GEMMEL
University of Otago, New Zealand

Professor Neil Gemmell is the AgResearch Chair in Reproduction and Genomics at the University of Otago and the Head of the Department of Anatomy. He leads a research group that blends ecology, population, conservation and evolutionary biology with technological spin-offs from the various genome projects. A recurring theme in his research is that of reproduction, with past and current projects spanning mating systems and mate choice sperm function, sex determination, sex allocation, and inter-sexual genomic conflict. Neil also has interests in several congruent fields of research, particularly the evolution of the mitochondrial genome, the evolution and functions of repetitive DNAs, the evolution of sex determining and sexual differentiation mechanisms and the processes that lead to speciation. Currently he heads a research consortium that is sequencing the genome of the tuatara in partnership with Ngatiwai iwi.

THE TUATARA GENOME PROJECT—UNLOCKING THE GENOME OF A LIVING FOSSIL

The tuatara (Spenodon punctatus) is iconic, unique to New Zealand and perhaps one of the most enigmatic of terrestrial vertebrates. Once widespread across the supercontinent of Gondwana, tuatara are now only found on a small number of offshore Islands in Cook Strait and the north of the North Island, New Zealand. We have now completed a de novo assembly for the 4.6-Gbp tuatara genome. In this presentation I will highlight some of the challenges associated with sequencing this genome and the novel insights emerging from the genome of this important linchpin in vertebrate evolution.

1330 - 1345

Paul Maclean
Agresearch Ltd, New Zealand

Paul Maclean started at AgResearch as a bioinformatician in 2008, after completing a Masters Degree in Bioinformatics at Auckland University. During his time at AgResearch, he has been involved in many projects ranging from genomics and transcriptomics to integrating 'omics on a variety of different eukaryotic and prokaryotic organisms. Most of these projects involve many different methods of next generation sequencing technology.

SHIGATOXIN TYPE 1 IS LIKELY TO BE VERTICALLY TRANSMITTED IN WAIKATO DAIRY FARMS

Paul H Maclean[1], Delphine Rapp[1], Colleen Ross[1]

1. AgResearch Limited, New Zealand

Shiga toxin-producing Escherichia coli (STEC) O26 is among the leading E. coli serogroups responsible for severe cases of diarrheoa and haemolytic uraemic syndrome (HUS) in humans. Ruminants, dairy cows in particular, have been recognized as an important reservoir of O26 strains. Typically, O26 isolated from human cases are different from bovine strains with regards to their virulence profile and to the presence of the shiga-toxin (stx-1) genes. In this study, 18 isolates from bovine faecal matter sampled from different Waikato dairy farms and 2 reference strains had their genomes sequenced using next generation sequencing and their genotypes called. This talk will detail the bioinformatics and clustering methodology used to investigate genetic relatedness of the isolates in the context of shigatoxin, intimin adherence protein and other pathogenicity related genes.

1345 – 1400

Aidan O'Brien
CSIRO, Australia

Aidan O'Brien graduated from the University of Queensland with a Bachelor of Biotechnology (1st class honours) in 2013. With Dr. Timothy Bailey as his honours supervisor, he developed GT-Scan, a CRISPR target predictor (http://gt-scan.csiro.au). Aidan is now based at CSIRO with the transformational bioinformatics team, where he developed VariantSpark, which applies BigData machine learning algorithms to genomic data. Aidan has 4 journal publications (3 first author) with 38 citations (h-index 2). He received the "Best student and postdoc" award at CSIRO in 2015 and attracted $180K in funding to date as AI. Next year he will be commencing a PhD in the field of genome editing.

BREAKING THE CURSE OF DIMENSIONALITY FOR MACHINE LEARNING ON GENOMIC DATA

Aidan O'Brien[1], Piotr Szul[2], Robert Dunne[3], Denis C. Bauer[1]

1. CSIRO, Health and Biosecurity, North Ryde, Sydney
2. CSIRO, Data 61, Dutton Park, Brisbane
3. CSIRO, Data 61, North Ryde, Sydney

Genomic data is becoming increasingly prevalent in the field of medical research with ever larger cohorts of samples analysed jointly. However, traditional software and compute hardware are unable to cater for large-cohort analysis, with the finite resource of computer memory becoming the bottleneck. MapReduce and MapReduce-like systems, such as Apache Hadoop and Spark, aim to overcome these obstacles.

Leveraging the power of Spark and its machine learning libraries (Spark ML), we built VariantSpark: A flexible framework for analysing genomic data. We previously demonstrated VariantSpark's ability on the 1000 Genomes Project (phase 3) data by clustering 2,500 individuals with 80 million genomic variants each into their super-population groups achieving an ARI=0.82 (1 perfect and -1 random clustering). Aiming to improve this performance and distinguish between the American and European populations, we sought to apply a supervised machine learning approach using Spark ML.

However, Spark ML's sophisticated machine learning methods suffer from "the curse of dimensionality". That is, they scale well with samples (n) but not with features (p), exceeding memory limits. This is due to it being developed for web-analytics data, which has different properties to genomics data. Hence performing sophisticated supervised machine learning tasks, such as random forest, was not possible on the 80 million variants of the 1000 genomes.

We hence extended VariantSpark to include CursedForest, an alternate random forests algorithm able to scale not only with samples, but also with features. We successfully trained a random forest on the 1000 genomes project data and achieved a cross-validated accuracy of ARI=0.96. Our implementation delivers this improved accuracy 80% faster than the unsupervised clustering approach (7 vs 30 hours) with low memory footprint (8 vs 24 GB).

illumina®

The MiniSeq System. Explore the possibilitilties.

Discover demonstrated NGS workflows for molecular biology applications.

Visit Illumina at AGTA booth **#15** to discover more about the MiniSeq.

Complete a feedback form to go into the draw to

WIN AN IPAD*

Prize drawn at lunch on final day.

POSTER PRESENTATIONS

#	POSTER TITLE	PRESENTING AUTHOR
1	AUTOSOMAL GENETIC CONTROL OF GENE EXPRESSION DOES NOT DIFFER ACROSS THE SEXES	Mr Irfahan Kassam
2	THE RECYCLED GENOME – RNA CAPTURESEQ REVEALS WIDESPREAD EXPRESSION OF HUMAN PSEUDOGENES AS CHIMERIC GENE ISOFORMS	Dr Daniel Thomson
3	PROPR: A SOFTWARE PACKAGE FOR IDENTIFYING PROPORTIONALLY EXPRESSED GENES USING COMPOSITIONAL DATA ANALYSIS	Dr Thom Quinn
4	FROM GENOMIC SEQUENCING TO PATIENT INFORMATION PROVISION: WHAT RESOURCES DO PATIENTS NEED?	Dr Dana Bradford
5	HIGH TEMPORAL RESOLUTION OF GENE EXPRESSION DYNAMICS IN DEVELOPING MOUSE ESC'S	Mr Brian Gloss
6	ESTABLISHING CELL COMPOSITION VIA DECONVOLUTION	Dr Saskia Freytag
7	THE MINION: NEW DATA; NEW ANALYSES	Dr David Eccles
8	IDENTIFYING HUMAN PAPILLOMAVIRUS IN HEAD AND NECK SQUAMOUS CELL CARCINOMAS	Dr Alexandra Garnham
9	SUPERTRANSCRIPT: CONSTUCTING AND VISUALISING A REFRENCE-LIKE TRANSCRIPTOME FOR NON-MODEL AND SEMI-ANNOTATED ORGANISMS	Dr Anthony Hawkins
10	DESIGNING A NEXT GENERATION GENOMIC DATA PLATFORM	Dr Velimir Gayevskiy
11	WHEN TO USE GENOMIC DATA: THE DEVELOPMENT OF CLINICIAN DECISION SUPPORT	Dr Dana Bradford
12	GROWING KIDNEYS IN A DISH: VERIFYING THE ROBUSTNESS AND REPRODUCIBILITY OF KIDNEY ORGANOID DIFFERENTIATION	Dr Belinda Phipson
13	IN-SILICO HLA GENOTYPING FROM GENOMIC SEQUENCING DATA	Dr Denis Bauer
14	ZIPF'S LAW AND BEYOND: TOWARD NEW AND USEFUL 10,000FT VIEWS OF YOUR NGS SEQUENCE DATA	Mr Alan McCulloch
15	PHARMACOGENETIC IMPLICATIONS OF VARIANTS OF RAAS GENES AMONGST PUNJABI POPULATION	Miss Jaspreet Kaur
16	SULPHUR-NITROGEN FERTILIZATION STRATEGY AND MECHANISM UNDERLYING THE INTERACTION OF SULPHUR AND NITROGEN METABOLIC PATHWAY	Miss Zitong Yu

AUTHORS: Student Early Career Researcher

MONDAY 10 OCTOBER 2016

#	POSTER TITLE	PRESENTING AUTHOR
17	UNDERSTANDING THE REGULATORY ROLE OF INTERGENIC VARIANTS: GIVING MEANING TO GWAS META-ANALYSES OF RHEUMATOID ARTHRITIS	Mr William Schierding
18	RNA-SEQ ANALYSIS OF GENE EXPRESSION CHANGES CAUSED BY MOOD STABILIZER DRUGS IN A SEROTONERGIC CELL LINE	Mrs Priyanka Sinha
19	RNA EDITING IN AN EDITING DEFICIENT ADAR1 MOUSE MODEL	Dr Alistair Chalk
20	GENOME SEQUENCING OF RICKETTSIA-LIKE ORGANISM ISOLATED FROM FARMED CHINOOK SALMON IN NEW ZEALAND	Dr Edna Gias
21	CHALLENGES OF WHOLE GENOME BISULPHITE SEQUENCING ON THE HISEQ X TEN	Dr Shalima Nair

TUESDAY 11 OCTOBER 2016

#	POSTER TITLE	PRESENTING AUTHOR
22	DEVELOPMENT OF ROBUST, LOW DNA INPUT GENOTYPING BY SEQUENCING (GBS)	Dr Gai McMichael
23	PREDICTING EXON SPLICING CHANGES TRIGGERED BY EPIGENETIC PROFILES	Dr Denis Bauer
24	THE RECYCLED HUMAN GENOME: DEEP TRANSCRIPTIONAL ANALYSIS OF PSEUODGENES REVEALS THE SCALE OF LINE1 MEDIATED EXON-SHUFFLING FORMING CHIMERIC GENE ISOFORMS	Dr Daniel Thomson
25	ENHANCER LANDSCAPE OF TUBERCULOSIS-INFECTED MURINE MACROPHAGES	Ms Elena Denisenko
26	DECIPHERING THE REGULATORY NETWORK OF MICRORNAS IN TUBERCULOSIS INFECTED MURINE MACROPHAGES	Mr Daniel Ho
27	THE INVOLVEMENT OF KIDNEY DNA METHYLATION IN BLOOD PRESSURE REGULATION	Miss Ingrid A Wise
28	USING DNA METHYLATION TO IDENTIFY THE FETAL CELL OF ORIGIN FOR CHILDHOOD ACUTE LYMPHOBLASTIC LEUKAEMIA	Mr Abdulmonem Alsaleh
29	EVALUATION OF RAPID LOW-INPUT PCR-FREE LIBRARY PREPARATION FOR HIGH-THROUGHPUT WHOLE GENOME SEQUENCING	Dr Jiang Tao

AUTHORS: Student Early Career Researcher

#	POSTER TITLE	PRESENTING AUTHOR
30	ANTISENSE PNA AND ANTIBIOTICS, SYNERGISTS AGAINST GRAM-NEGATIVE BACTERIA	Sr Jaime Iván Castillo Silva
31	OPTIMUSPRIME: MULTIPLEX PRIMER DESIGN TOOL FOR HI-PLEX SEQUENCING	Mr Edmund Lau
32	IDENTIFYING GENETIC POLYMORPHISMS IN PHARMACOGENES USING NANOPORE SEQUENCING	Dr Simone Cree
33	INVESTIGATING THE ROLE OF HISTONE VARIANT H2A.B1 IN THE REGULATION OF CANCER/TESTIS (CT) ANTIGENS IN HODGKIN LYMPHOMA.	Dr Sebastian Kurscheid
34	GENOTYPING-BY-SEQUENCING FOR DIVERSE APPLICATIONS IN BIOLOGY	Dr Jeanne Jacobs
35	ICE-T: INTEGRATED CELL EXPRESSION TOOLKIT FOR LOW INPUT TRANSCRIPT PROFILING	Mr Chris Harris
36	DIVERSIFICATION OF INNATE IMMUNE RESPONSES BY TRANSCRIPTIONAL MECHANISMS	Miss Suzanne Butcher
37	THE GENOMIC LANDSCAPE OF OESOPHAGEAL ADENOCARCINOMA AND ITS PRECURSOR BARRETT'S OESOPHAGUS	Dr Felicity Newell
38	WHOLE TRANSCRIPTOME ANALYSIS OF CONSTITUTIVE PD-L1 EXPRESSION IN MELANOMA CELL LINES	Mr Antonio Ahn
39	GENOMICS CHALLENGES AND OPPORTUNITIES FOR BIOPESTICIDE DEVELOPMENT	Mr Mikhail Fokin
40	ESTABLISHING THE INFRASTRUCTURE REQUIRED TO SUPPORT A HIGH-THROUGHPUT WHOLE GENOME SEQUENCING FACILITY	Dr Joice Kuroiwa-Trzmielina
41	GETTING THE MOST FROM AN RNA-SEQ EXPERIMENT: SOME DESIGN CONSIDERATIONS	Dr Susan Corley
42	INVESTIGATION INTO THE CAUSE OF ENCEPHALITIS IN KAKÎ CHICKS	Dr Barbara Binney
43	ZIPF'S LAW AND BEYOND: TOWARD NEW AND USEFUL 10,000FT VIEWS OF YOUR NGS SEQUENCE DATA	Mr Alan McCulloch

AUTHORS: Student Early Career Researcher

POSTER 1

AUTOSOMAL GENETIC CONTROL OF GENE EXPRESSION DOES NOT DIFFER ACROSS THE SEXES

Mr Irfahan Kassam[1], Dr Luke Lloyd-Jones[1], Mr Alexander Holloway[1], Consortium of the Architecture of Gene Expression (CAGE), Associate Professor Jian Yang[1], Professor Peter M. Visscher[1,2], Dr Allan F. McRae[1]

1. The Queensland Brain Institute, The University of Queensland
2. University of Queensland Diamantina Institute, Translational Research Institute, The University of Queensland

Abstract

Despite their nearly identical genomes, males and females differ in risk, incidence, prevalence, severity, and age-at-onset of many diseases. Sexual dimorphism is also seen in human autosomal gene-expression, and has largely been explored by examining the contribution of genotype-by-sex interactions to variation in gene expression, with weak evidence for sex-specific eQTLs. In this study, we use data from a mixture of pedigree and unrelated individuals with verified European ancestry to investigate sex-specific genetic architecture of gene expression measured in whole blood across n = 1048 males and n = 1005 females by treating gene expression intensities in the sexes as two distinct traits and estimating the genetic correlation (rG) between them. These correlations measure the similarity of the combined additive genetic effects of all causal loci across the autosomal chromosomes, and thus the level of common genetic control of gene expression across the sexes. Genetic correlations were estimated across the sexes for the expression levels of 12,528 autosomal genes using bivariate REML, and tested for differences in autosomal genetic control of gene expression across the sexes. Overall, no deviation of the distribution of test statistics was observed from that expected under the null hypothesis of a common autosomal genetic architecture for gene expression across the sexes. This indicates that males and females share the same common genetic control of gene expression.

POSTER 2

THE RECYCLED GENOME – RNA CAPTURESEQ REVEALS WIDESPREAD EXPRESSION OF HUMAN PSEUDOGENES AS CHIMERIC GENE ISOFORMS

Dr Daniel W Thomson[1] and A/Prof Marcel E Dinger[1]

1. Garvan Institute of Medical Research, Sydney, Australia.

Abstract

Mutated gene copies called pseudogenes are ubiquitous in mammalian genomes, but it is unclear why such high numbers are tolerated in higher eukaryotes. Several pseudogenes have been characterized as functional long noncoding RNAs, however a broad functional role of most pseudogenes is obfuscated by a lack of comprehensive transcriptomic annotation.

We present a global transcriptomic investigation of human pseudogenes using RNA CaptureSeq, capable of measuring transcription from 30% of annotated pseudogenes (Genecode v19). We show pseudogene expression is highly tissue specific and regularly at odds to the expression of the parent gene from which it was derived, perpetuating the idea that pseudogenes show function that is divergent from its parent gene.

Using *de novo* RNA assembly and chimeric RNA analysis we demonstrate an overwhelming trend of the reuse of pseudogene sequences as functional exons, and UTRs (untranslated regions) of unrelated protein coding genes, building apon known protein coding genes to expand the number and complexity of protein coding and non coding isoforms

This provides evidence of a broader functional role of all pseudogenes. Pseudogenes are an intermediate product of a process by which higher eukaryotes recycle functional genetic components creating chimeric gene isoforms with a greater complexity of function.

POSTER 3

PROPR: A SOFTWARE PACKAGE FOR IDENTIFYING PROPORTIONALLY EXPRESSED GENES USING COMPOSITIONAL DATA ANALYSIS

Dr. Thom Quinn[1], Dr. Mark Richardson[1], Dr. David Lovell[2], Dr. Tamsyn Crowley[1]

1. Bioinformatics Core Research Facility, Deakin University, Waurn Ponds, Victoria, Australia
2. Queensland University of Technology, Brisbane, Australia

Abstract

The identification of differentially expressed genes is becoming a cornerstone in the analysis of biological data. Advances in the biological methods used to produce differential gene expression datasets have led to a rapid increase in the amount of data generated; however, analysis methods have not kept up. Often, analysis of these data relies on the use of correlation-based statistics to prioritize a subset of genes for subsequent analysis. Yet, correlation is not a valid measure of association for compositional data (i.e., data that carry only relative information). These data typify some of the most frequently studied biological data, including data produced by microarray assays or high-throughput RNA-sequencing. Here, we present the topic of compositional data analysis, discuss its relevance to biological research, and detail a computational framework for dealing with these kinds of data. Specifically, we show how proportionality, an adjusted measure of the co-variation between log-ratio transformed feature vectors, can avoid the pitfall of correlation misuse that sometimes yields spurious results. Finally, we introduce our free and open source software, the *propr* package for the R programming environment, which provides a user-friendly interface for calculating proportionality as a statistically valid alternative to correlation. Although we focus its application here to the analysis of gene expression data, the methods bundled in this software apply equally well to all count data, including data produced by chromatin immunoprecipitation (ChIP), ChIP-sequencing, Methyl-Capture sequencing, or other techniques.

POSTER 4

FROM GENOMIC SEQUENCING TO PATIENT INFORMATION PROVISION: WHAT RESOURCES DO PATIENTS NEED?

Dr Dana Bradford[1], Dr Melissa Martyn[2], Michelle Rodriguez[1], Dr Clara Gaff[2,3]

1. Commonwealth Scientific and Industrial Research Organisation, Australia
2. Melbourne Genomics Health Alliance, Australia
3. University of Melbourne, Australia

Abstract

Australia's health services and medical research organisations are currently undertaking a number of projects to prepare Australia for genomic medicine. These initiatives are already demonstrating how genomics can be used effectively in clinical medicine to improve patient outcomes and reduce health costs. An integral part of the integration of genomics into everyday care will be the provision of accurate, reputable information and resources for clinicians, patients and researchers. In this study, we sought to determine the views of patients and parents of patients of a patient portal, information leaflets and a website. Here we describe the feedback on those resources and the challenge of tailoring resources for a use by a diverse population throughout their genomic journey.

POSTER 5

HIGH TEMPORAL RESOLUTION OF GENE EXPRESSION DYNAMICS IN DEVELOPING MOUSE ESC'S

Mr Brian Gloss[1], Ms Bethany Signal[1], Mr Dominik Kaczorowski[1], Prof Andrew Perkins[2], A/Prof Marcel Dinger[1]

1. Garvan Institute of Medical Research, Sydney, Australia
2. Mater–UQ Research Institute, The University of Queensland, Translational Research Institute, Brisbane, Australia

Abstract

Investigations of transcriptional responses during developmental transitions typically use time courses with intervals that are not commensurate with the timescales of known biological processes. Moreover, such experiments typically focus on protein-coding transcripts, ignoring the important impact of long noncoding RNAs. We evaluated coding and noncoding expression dynamics at high temporal resolution (6-hourly) in differentiating mouse embryonic stem cells and report the effects of increased temporal resolution on the characterization of the underlying molecular processes. We present an improved resolution of transcriptional alterations, including regulatory network interactions, coding and noncoding gene expression changes as well as alternative splicing events, many of which cannot be resolved by existing coarse developmental time-courses. We describe novel short lived and cycling patterns of gene expression and temporally dissect ordered gene expression at bidirectional promoters and responses to regulatory lncRNAs & transcription factors. These findings have important implications upon the investigation of regulatory dynamics of any biological transition.

POSTER 6

ESTABLISHING CELL COMPOSITION VIA DECONVOLUTION

Dr Saskia Freytag[1], Ass Prof Johann Gagnon-Bartsch[2], Prof Terry Speed[1], Prof Melanie Bahlo[1]

1. Walter And Eliza Hall Institute
2. University of Michigan

Abstract

Humans are comprised of a diverse collection of cell types that carry out a variety of specialized functions. Thus loss or increase of one or more cell types can result in malfunction and potentially result in disease. For example, some patients with Huntington's disease or schizophrenia experience loss of neuronal cells, which has been hypothesized to cause memory loss. Cell composition, of course, also effects measurements of gene expression, as such measurements are typically extracted from samples of cell mixtures.

Computational biologists are able to establish relative proportions of cell types directly from gene expression data with the help of deconvolution. There exist many different deconvolution methods, some using marker genes and others requiring expression profiles measured from pure samples. Here, we present the first comprehensive comparison of different deconvolution methods, including our own tool. Unlike other studies, we compare the performance of these tools on a variety of microarray and RNA-seq test datasets with true cell proportions established either by fluorescence-activated cell sorting or by experimental design.

POSTER 7

THE MINION: NEW DATA; NEW ANALYSES

Dr David Eccles[1]

1. Gringene Bioinformatics, Wellington, New Zealand

Abstract

The Oxford Nanopore MinION is unique among DNA sequencing technology, in that it is the only technology that doesn't require synthesis in order to derive a DNA sequence. Instead, sensors detects subtle changes in the electrical properties of bases, observing single strands of DNA as they pass through a nanopore.

The novelty of signal-level data produced by the MinION presents new opportunities for data display and analysis. Over the course of the last two years, a number of different methods were developed to represent both signal-level and base-called MinION data. Three of these methods will be demonstrated:

- Correcting assemblies with noisy reads
- Aligning raw sensor data to sequenced bases
- Turning signal squiggles into sound

The disruptive nature of this new sequencing technology means that there is plenty of scope for new research and analysis. It is hoped that by demonstrating a few things that can be done with the technology, others can be inspired to discover other novel uses for the device.

POSTER 8

IDENTIFYING HUMAN PAPILLOMAVIRUS IN HEAD AND NECK SQUAMOUS CELL CARCINOMAS

Dr Alexandra Garnham[1,2], Dr Kendrick Koo[1], Associate Professor Oliver Sieber[1,2], Professor Gordon Smyth[1,3]

1. The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, Victoria, Australia
2. Department of Medical Biology, The University of Melbourne, Parkville, Victoria, Australia
3. Department of Mathematics and Statistics, The University of Melbourne, Parkville, Victoria, Australia

Abstract

Human Papillomavirus (HPV) is comprised of a large group of related viruses, with over 100 subtypes identified. It has been well established that many of the HPV subtypes cause cancers of the genitals and anus, as well as in some regions of the head and neck. HPV is detected in approximately 25% of all Head and Neck Squamous Cell Carcinomas (HNSC) with the majority of cases occurring in the oropharynx. While there is strong evidence connecting HPV as a major cause of oropharyngeal cancers, it has remained unclear if this same link exists between HPV and cancers of the oral cavity. To explore this possibility, we acquired raw RNA sequencing data from HNSC samples generated by The Cancer Genome Atlas. With these data, we have developed a pipeline utilizing the Subjunc aligner that enables us to deconvolve viral RNA from human at the subtype level. This allows us to detect the presence of viruses in the tumour samples. Using this pipeline we are now investigating genetic differences between HPV positive and negative samples in both oropharyngeal and oral cavity cancers. Furthermore, we are studying which viral genes are active in positive samples.

POSTER 9

SUPERTRANSCRIPT: CONSTUCTING AND VISUALISING A REFRENCE-LIKE TRANSCRIPTOME FOR NON-MODEL AND SEMI-ANNOTATED ORGANISMS

Dr Anthony Hawkins[1], Dr Nadia Davidson[1], Associate Professor Alicia Oshlack[1]

1. Murdoch Children's Research Institute, Victoria, Australia

Abstract

For organisms without a reliable genome or genome annotation, de novo assembly of RNA-seq reads can be performed to reconstruct and analyse expressed transcripts. However, analysing and visualising the splicing structure of genes, built through assembly, is problematic because each gene is represented by a set of transcripts, and not by a single complete gene sequence. Here we present the concept of a SuperTranscript, a single linear representation for each gene. Each SuperTranscript contains all the unique exonic sequence and ordering from its original constituent transcripts. A SuperTranscript enables not only the visualisation of read coverage and exon-like structure but provides the basis for doing Differential Exon Usage (DEU) for the first time in non-model organisms. Finally it can be used to infer the existence of novel genetic sequence not in current annotation, because it allows transcriptome from a variety of sources to be combined: De novo assembly, Genome Guided Assembly and Annotated transcriptome. A software package Ribbon has been developed to produce SuperTranscripts.

POSTER 10

DESIGNING A NEXT GENERATION GENOMIC DATA PLATFORM

Dr Velimir Gayevskiy[1], Dr Tony Roscioli[1,2,3], Dr Marcel E Dinger[1,2,3], Dr Mark J Cowley[1,3]

1. Kinghorn Centre for Clinical Genomics, Garvan Institute of Medical Research, Sydney, Australia
2. Genome.one, Sydney, Australia
3. St Vincent's Clinical School, University of New South Wales, Sydney, Australia

Abstract

We are no longer content to generate genomic data, answer a specific question and put it aside. Increasingly, we are interested in combining data from disparate technologies and storing these data long term in a searchable manner with rich up-to-date annotations. To support this, we have developed an in-house clinical genomics variant filtration platform, *Seave* (http://www.seave.bio). It enables the storage and filtration of short variants, from targeted or whole genome-scale data, and allows searching for variants matching inheritance models. Here we share details for how we have improved this platform in relation to variant annotation, and in representing Copy Number and Structural Variants.

Annotations are crucial to contextualising observed genomic variation within the current state of knowledge. We have found that while aggregated variant annotation databases such as dbNSFP are appealing, their low frequency of updates, set annotation sources and lack of entire classes of variation (e.g. indels), can significantly impair variant filtering and prioritisation. Instead, creating the infrastructure to store and update individual databases on a weekly basis has seen the number of pathogenic variants annotated increase from 23,104 in dbNSFP v2.9, to 127,814 from just the ClinVar database. Furthermore, OMIM generates an average 49 new entries and 614 updates per month, and after creating our own OMIM database, we are able to keep up with these updates rather than being reliant on stale data.

Continued next page

This has significant benefits prioritising known clinically relevant variants, resulting in improved efficiency in making clinical diagnoses.

For genomic variation to be comparable between different technologies, analysis pipelines and cohorts, variant information and annotations should be stored with maximal compatibility. We have found that splitting multi-allelic variants into separate entries (decomposing) and left aligning (normalising) both annotations and variants reduces the number of clinically relevant filtered variants by ~25%. The excluded variants were now correctly matched to annotations and excluded by the filtering process, reducing the time to diagnosis and allowing better comparison against future data.

Finally, as we move from inferring and storing only short variants to also inferring CNVs, SVs and regions of homozygosity, we have developed an architecture for representing, and querying both short, and long variants, detected from any analysis method. The integration of short and long variants is compatible with traditional inheritance model filtering, and substantially filters out large swathes of the genome that do not match the expected inheritance pattern, again leading to less putative variants to interpret.

We will present the impact that these improvements have had upon various clinical genomic investigations.

POSTER 11

WHEN TO USE GENOMIC DATA: THE DEVELOPMENT OF CLINICIAN DECISION SUPPORT

Dr Dana Bradford[1], Dr Shlomo Berkovsky[1], Dr Melissa Martyn[2], Tim Baker[2], Michelle Rodriguez[1], Dr Denis Bauer[1], Dr Clara Gaff[2,3]

1. Commonwealth Scientific and Industrial Research Organisation, Australia
2. Melbourne Genomics Health Alliance, Australia
3. University of Melbourne, Australia

Abstract

Healthcare is currently being transformed by the introduction of genomic sequencing – a major advancement in personalised medicine. This advent provides new opportunities for clinicians to use genomic data in decision making about patient diagnosis and treatment, but this can only be achieved through access to data and support in its use. Engaging with clinicians in the development of decision support tools will optimize their relevance and the adoption of genomic sequencing in healthcare. In this study, existing data from clinician workshops and interviews together with horizon scanning of relevant technologies were used to define clinician portal specifications. We describe a preliminary design of a decision support tool for use by clinicians and the manner in which the technology may be evaluated.

POSTER 12

GROWING KIDNEYS IN A DISH: VERIFYING THE ROBUSTNESS AND REPRODUCIBILITY OF KIDNEY ORGANOID DIFFERENTIATION

Dr Belinda Phipson[1], Dr Minoru Takasato[2], Pei X Er[1], Professor Melissa H Little[1,3], Associate Professor Alicia Oshlack[1,4]

1. Murdoch Childrens Research Institute, Melbourne, Australia
2. RIKEN Center for Developmental Biology, Kobe, Japan
3. Department of Pediatrics, The University of Melbourne, Melbourne, Australia
4. School of BioSciences, The University of Melbourne, Melbourne, Australia

Abstract

The ability to make three dimensional organoids from human pluripotent stem cells through directed differentiation opens up the possibilities of personalised drug testing, disease modelling and regeneration, as well as enhancing our knowledge of organ development. Recent studies have shown spontaneous patterning of a number of tissue organoids, including brain, optic cup, stomach and intestine. Recently we published a protocol for the differentiation of human pluripotent stem cells into kidney organoids containing all the major components of the kidney: nephrons, collecting duct, vasculature and surrounding interstitium (Takasato et al. Nature 2015). However, successfully using kidney organoids for drug screening or disease modelling will rely on the reproducibility and robustness of the protocol. In order to understand the sources of variability that arise during the step-wise differentiation process, we designed an experiment to analyse the transcriptomes of multiple organoids. Our design investigates major contributions to variation between organoids derived from distinct differentiations separated in time, as well as organoids grown concurrently from the same starting cells in separate vials. While transcriptional correlations were high between organoids (>0.9 between batches, >0.95 within batches), there were many genes showing a great deal of variability across organoids. We used sophisticated statistical methods to partition variability from different sources for each gene.

Combining this data with a complete RNA-Seq time course of differentiation from induced pluripotent stem cells to kidney organoid revealed the relative maturity of the organoids was the major source of variation between batches. Additionally, we provide a framework for using these highly variable, maturity-related genes for removing sources of unwanted, but known, variation in independent kidney organoid experiments.

POSTER 13

IN-SILICO HLA GENOTYPING FROM GENOMIC SEQUENCING DATA

Dr Denis Bauer[1], Ms Armella Zadoorian[1,2], Dr Laurence Wilson[1], Melbourne Genomics Health Alliance, Dr. Natalie Thorne[3,4,5,6]

1. CSIRO, Health and Biosecurity, North Ryde, Sydney
2. School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, 2033, Australia
3. Murdoch Childrens Research Institute, Royal Children's Hospital, Parkville, 3052, Australia
4. Department of Medical Biology, The University of Melbourne, Parkville, 3052, Australia
5. Melbourne Genomics Health Alliance
6. Walter and Eliza Hall Institute, Parkville, 3052, Australia.

Abstract

Despite being essential for numerous clinical and research applications, high resolution human leukocyte antigen (HLA) typing remains challenging. With next generation sequencing (NGS) data becoming widely accessible, on-demand *in silico* HLA typing offers an economical and efficient alternative.

In this study we evaluate the HLA typing accuracy and efficiency of five computational HLA typing methods by comparing their predictions against a curated set of over 1000 published polymerase chain reaction (PCR) derived HLA genotypes on three different datasets (whole genome sequencing, whole exome sequencing and transcriptomic sequencing data). Despite reporting to be highly accurate on smaller datasets, the highest accuracy at clinically relevant resolution (4-digit) we observe is 81% on RNAseq data by PHLAT and 99% accuracy by OptiType when limiting to Class I genes only. We also observed variability between the tools for resource consumption, with runtime ranging from an average of 5 hours (HLAminer) to 7 minutes (seq2HLA) and memory from 12.8 GB (HLA-VBSeq) to 0.46 GB (HLAminer) per sample.

While a minimal coverage is required, other factors also determine prediction accuracy and the results between tools do not correlate well. Therefore, by combining tools, there is the potential to develop a highly accurate ensemble method that is able to deliver fast, economical HLA typing from existing sequencing data.

POSTER 14

ZIPF'S LAW AND BEYOND: TOWARD NEW AND USEFUL 10,000FT VIEWS OF YOUR NGS SEQUENCE DATA

Mr Alan McCulloch[1]

1. Agresearch New Zealand

Abstract

The frequencies of words in most languages have been found to follow a power law relationship relative to their rank in the frequency table, an observation popularised by George Zipf and known as Zipf's law. Power law behaviour has also previously been noted in the occurrence of various kinds of genomic properties including DNA words. We have noted power law or near power law behaviour in the occurrence of short DNA words in a wide variety of NGS sequencing contexts, including re-sequencing, amplicon sequencing and genotyping by sequencing, and have exploited this by incorporating a "zipfian analysis" of NGS datasets generated by our own lab and others into a number of pipelines. We show examples of how this interesting power-law behaviour can be used to quickly diagnose common problems such as incompletely trimmed sequence data, sample labelling error and unusual biases that can otherwise be difficult to see. We show an example of how Zipfian analysis also has a potential scientific application via hypothesis generation and confirmation concerning relative levels of biological repeat content. We briefly describe a broader "entropy-reduction" rationale and context for this kind of analysis to try to help understand why it works, which could provide a basis for the development of novel genomic data investigation methods. We refer readers interested in more details of this class of methods to our paper currently available in pre-print, and to our open-source python and R software libraries.

POSTER 15

PHARMACOGENETIC IMPLICATIONS OF VARIANTS OF RAAS GENES AMONGST PUNJABI POPULATION

Miss Jaspreet Kaur[1], Dr Praveen P. Balgir[1]

1. Department of Biotechnology, Punjabi University, Patiala, India

Abstract

Renin angiotensin aldosterone system (RAAS) genes encoding renin (REN), angiotensinogen (AGT), angiotensin-converting enzyme (ACE), angiotensin type 1 receptor (AGTR1) and aldosterone synthase gene (CYP11B2) play the most important role in regulation of blood pressure. Genes encoding components of this system are the main drug targets for developing anti-hypertensive therapeutics. Most common protein targets among currently available marketed antihypertensive drugs are Angiotensin converting enzyme (ACE) inhibitors and Angiotensin receptor blockers (ARBs). Two well studied and two less studied but potentially harmful polymorphisms predicted by the different *In-silico* structure activity analysis tools were selected. In present study Insertion/deletion (I/D) polymorphism (rs1799752) and rs148193919 of ACE gene and rs5186 & rs12721226 of AGTR1 gene were investigated. In case of rs148193919 (Asp353Asn), SIFT categorized the mutation as leading to significant structural change i.e. Deleterious (with 0 SIFT score), PolyPhen as probably damaging (with Polyphen score 0.999). In case of rs12721226 (Ala163Thr) of AGTR1 is present in drug (Losartan) binding site. The analysis included 300 clinically diagnosed hypertensive patients with/without associated metabolic disorders and 100 normal healthy subjects with their informed consent. Study was approved by Institutional clinical ethical committee vide No. ICEC/4/2011. Genotyping was carried out using PCR- RFLP. In ACE gene polymorphisms , the frequencies of the I and D alleles of rs1799752 were 0.41 and 0.59 in hypertensive and 0.54 and 0.47 in normotensive and the frequencies of the G and A alleles of rs148193919 were 0.78 and 0.22 in hypertensive & 0.87 and 0.13 in normotensive.

In AGTR1 gene polymorphisms, the frequencies of G and A allele of rs12721226 were 0.96 and 0.04 in hypertensive whereas, frequency of G allele is 1 and frequency of A allele is 0 in normotensive whereas in rs5186, frequency of A allele is 1 and frequency of C allele is 0; that is they are monomorphic at this locus in hypertensive as well as in normotensive. Out of these 4 polymorphism, I/D rs1799752 (p=0.0393) and rs148193919 (p=0.0135) of ACE gene show statistically significant differences in distribution of variants amongst hypertensive and normotensive Punjabi population.

POSTER 16

SULPHUR-NITROGEN FERTILIZATION STRATEGY AND MECHANISM UNDERLYING THE INTERACTION OF SULPHUR AND NITROGEN METABOLIC PATHWAY

Miss Zitong Yu[1], Dr Angéla Juhász[1], Dr Penghao Wang[1], Dr Wujun Ma[1]

1. School of Veterinary and Life Science

Abstract

Wheat (*Triticum aestivum L.*) is receiving increasing attention on its quality improvement. To achieve high wheat quality, sulphur fertilizer application is essential, whilst adequate amount of nitrogen fertilizer can trigger the positive effects of sulphur. Important agronomic traits and protein parameters such as NUE (nitrogen use efficiency), harvest index, protein percentage, and protein yield are targeted. HPLC analysis results demonstrated that with the increasing ratio of sulphur fertilizer, the percentage of UPP and the ratio of glutenins to gliadins are both increased. For protein compositions, the HMW-GS to LMW-GS ratio is reduced, the percentage of α/α-gliadins is decreased, whereas the percentage of α-gliadins is increased. Based on above results, two comparative treatments (S0N50 and S30N50) in three grain-filling stages (7DPA, 14DPA and 21DPA) are selected for RNA-seq experimental design in order to screen differential expressed genes (DEGs). The survey sequence from IWGCS is used as reference for alignment and mapping in bioinformatics analysis. The preliminary results revealed that there are 1 up-regulated DEG and 63 down-regulated DEGs within low and high sulphur treatment in 7DPA. Gene ontology (GO) enrichment analysis illustrated that these DEGs involved 1,251 nodes in various biological process, and 1,262 involvements in different molecular functions as well as 632 in various kinds of cellular components. Afterwards, the KEGG pathway enrichment analysis identified the interactions of above DEGs are significantly enriched in nitrogen and sulphur metabolism pathway, glycolysis and TCA cycle pathway, fatty acid metabolism pathway, etc.

At last, through blastx the translated target gene sequence to Bdi (*Brachypodium distachyon*) protein sequence in STRING database, the predicted protein-protein interactions network (PPI) is constructed. Further work includes post-translation modifications (PTMs) targeting identification of phosphorylation sites in motif and metabolomic study targeting amino acids involved in the interaction of sulphur and nitrogen metabolic pathways is being conducted at present.

POSTER 17

UNDERSTANDING THE REGULATORY ROLE OF INTERGENIC VARIANTS: GIVING MEANING TO GWAS META-ANALYSES OF RHEUMATOID ARTHRITIS

Mr William Schierding[1,2], Dr. Ralph Grand[1,6], Mr. Cameron Eckblad[1], Professor Wayne Cutfield[1,5], Professor Jörg Langowski[3], Dr. Ada L. Olins[4], Dr. Donald E. Olins[4], Dr. Justin O'Sullivan[1,5]

1. Liggins Institute, University of Auckland, Auckland, New Zealand
2. Bioinformatics Institute, University of Auckland, Auckland, New Zealand
3. Division Biophysics of Macromolecules, German Cancer Research Center (DKFZ), Heidelberg, Germany
4. Department of Pharmaceutical Sciences, University of New England, College of Pharmacy, Portland, ME, USA
5. Gravida: National Centre for Growth and Development, University of Auckland, Auckland, New Zealand
6. Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland

Continued next page

Abstract

GWAS, meta-analysis, spatial chromatin, DNA folding, rheumatoid arthritis

Objectives: Meta-analyses of GWAS have highlighted genetic variation in many disorders. Previously, our lab has identified allele-specific SNP enhancer activity in intergenic variants, associating this activity to putative functions in diabetes and human growth[1,2]. In autoimmune diseases such as rheumatoid arthritis (RA), GWAS SNPs are largely found in intergenic regions, thus leaving their function as undefined non-coding roles, affecting close- and long-distance regulatory activity.

Methods: We performed a novel analysis of hundreds of underlying genetic cohort studies comprising 489 GWAS variants for RA disease or markers of RA drug efficacy. These variants were analysed for: 1) 3D spatial (physical) connections (as captured by proximity ligation); 2) chromatin markers of regulatory functions (DNAse Hypersensitivity Sites, transcription factor binding site motifs, luciferase allele-specific SNP enhancer activity); and 3) linked gene expression (Expression quantitative trait locus, eQTL analysis).

Results: eQTL analysis reveals that many of the GWAS SNPs have more significant gene regulatory roles with distant spatially-related genes than they do with any gene in the local gene neighbourhood. For example, rs3184504, rs653178, and rs11545078 show evidence of differential long distance spatial associations in HL60 cells in undifferentiated versus differentiated states.

rs3184504/rs653178 (SH2B3 locus) has a spatially-reinforced eQTL with the BAZ2A gene almost 60Mb away. BAZ2A expression has a role in platelet formation. In RA, activated platelets interact with neutrophils to boost the neutrophil's inflammatory potential and platelet thrombi directly affect the integrity of the lumen of synovial vessels, contributing to synovial inflammation. Thus, as these SNPs are associated with autoimmune disease onset, BAZ2A could play a role in the platelet-specific roles of RA onset and progression.

rs11545078 (GGH locus) has an inter-chromosomal spatially-reinforced eQTL with the MGEA5 locus.rs11545078 is associated with methotrexate (RA drug treatment) toxicity. While rs11545078 has an effect on local GGH expression, GGH transcript levels are low in most autoimmune-relevant GTEx tissues. However, the MGEA5 gene is overexpressed in peripheral blood mononuclear cells. Thus, as rs11545078 is associated with methotrexate toxicity in RA treatment, this identifies MGEA5 as a novel gene target in RA drug treatment.

Conclusions: Our results provide evidence that spatial connections associate intergenic GWAS SNPs to distant loci with roles in autoimmune disease through regulatory relationships. We have found that regulatory SNPs in RA have a strong relationship with open chromatin, ENCODE enhancer activity in blood cell lines, and long-distance gene regulatory effects on gene expression (i.e. identified the link between spatial associations, enhancer activity, and eQTLs). Crucially, the regulatory relationships of these GWAS loci identify spatial associations with loci that haven't been previously discovered by RA GWAS, potentially expanding our understanding of the role of these genes in autoimmune disease. This supports the hypothesis that intergenic GWAS findings identify loci with roles in gene regulation through 3D genomics. Finally, spatial relationships here have shown a possible avenue for identifying genes regulated via spatial-associations as putative druggable targets, leveraging drug development pathways into re-purposing of existing drugs for designing better treatments for disease.

References:

1. Schierding W, Antony J, Cutfield WS, Horsfield JA, O'Sullivan JM. Intergenic GWAS SNPs are Key Components of the Spatial and Regulatory Network for Human Growth. Hum Mol Genet. 2016; doi:10.1093/hmg/ddw165
2. Schierding W, O'Sullivan JM. Connecting snps in diabetes: A spatial analysis of meta-GWAS loci. Front Endocrinol (Lausanne). 2015;6. doi:10.3389/fendo.2015.00102

POSTER 18

RNA-SEQ ANALYSIS OF GENE EXPRESSION CHANGES CAUSED BY MOOD STABILIZER DRUGS IN A SEROTONERGIC CELL LINE

Mrs Priyanka Sinha[1], Professor Martin Kennedy[1], Dr John Pearson[1]

1. University Of Otago Christchurch

Abstract

Sodium valproate and lithium are drugs of very different chemical classes that are widely prescribed in the treatment of bipolar disorder and other conditions. Molecular and pharmacological studies have revealed some relevant properties and targets of these drugs but their precise modes of action are not yet understood.

Our primary hypotheses are that genes displaying specific regulation by either or both of these drugs are relevant to the mechanism of action, and that genes which are co-regulated by both drugs will highlight common pathways of action which may be of the greatest interest. To test these hypotheses we are using a serotonergic cell line called RN46A, which is an immortalized serotonergic precursor cell line derived from rat raphe nucleus. This cell line represents a relevant model for the neurological effects of valproate and lithium, as most mood disorders appear to involve the serotonergic system of the central nervous system. The cell line is exposed to therapeutically relevant levels of each drug for 72 hour, then RNA is extracted for RNA-Seq analysis. RNA-Seq analysis yielded evidence for 145 genes (log2 fold change >1.5 and false discovery rate of 0.05) being significantly changed after exposure to valproate, and none with Lithium.

A subset of these 145 genes were subsequently chosen for validation with nanostring in a separate experiment where RN46A cell line was exposed to valproate along with lamotrigine, a mood stabiliser; Trichostatin-A (TSA), a histone deacetylase (HDAC) inhibitor; Valpromide, a valproate analog minus HDAC inhibition activity and lithium. 11 genes were validated with nanostring.

In particular, a gene called *Mmp13* (Matrix Metalloproteinase 13) known to be involved in degradation of extracellular matrix, was significantly upregulated by both valproate and TSA, establishing HDAC inhibition upregulation of Mmp13. *Cacna1b, Shank3 and Snap91* were upregulated by both valproate and valpromide, suggesting non-HDAC inhibition regulation. *Cdkn1c, Erbb3, Lingo1, Mpp3, Ngfr, Pak3 and Zcchc12* were upregulated only by valproate. None of these 11 genes were changed either by lithium or lamotrigine. These results provide important insight into the understanding of mechanism of action of valproate and need further investigation to yield a detailed picture.

POSTER 19

RNA EDITING IN AN EDITING DEFICIENT ADAR1 MOUSE MODEL

Dr Alistair Chalk[1,2], Joshua White[1], Brian Liddicoat[1], Jacki Heraud-Farlow[1], Ankita Goradia, Assistant Professor Jin Li[3], Associate Professor Carl Walkley[1,2]

1. St. Vincent's Institute, Fitzroy, 3065 VIC, Australia
2. Department of Medicine, St. Vincent's Hospital, University of Melbourne, Fitzroy, 3065 VIC, Australia
3. Department of Genetics, Stanford University, Stanford, CA 94305, USA

Abstract

The conversion of genomically encoded adenosine to inosine in dsRNA is termed as A-to-I RNA editing. This process is catalysed by two of the three mammalian Adar proteins which have essential functions for normal organismal homeostasis. Several recent studies have identified that a feature of absence or reductions of Adar1 activity, conserved across human and mouse models, is a profound activation of interferon-stimulated gene signatures and innate immune responses.

Using an Adar1 editing deficient mouse, we mapped A-to-I editing events in fetal liver and identified clusters of hyper-edited adenosines located in long 3'UTRs of erythroid-specific transcripts, and that these are Adar1 specific editing events.

Further analysis of this observation has lead to the conclusion that editing by Adar1 is required to prevent activation of the cytosolic innate immune system, primarily focused on the dsRNA sensor MDA5 (Ifih1). The delineation of this mechanism places Adar1 at the interface between the cells ability to differentiate self- from non-self dsRNA. Based on MDA5 dsRNA recognition requisites, the mechanism indicates that the type of dsRNA must fulfil a particular structural characteristic, rather than a sequence-specific requirement.

Here we present the identification of RNA editing sites for an Adar1 editing deficient mouse across two tissues expressing differing levels of dsRNA response, suggesting that the extent of dsRNA response is dependent on the amount of (unedited) dsRNA and the requirement of MDA5 expression. We further quantify known RNA editing sites using mmPCR-Seq.

While additional studies are required to molecularly verify the genetic model, the observations to date collectively identify A-to-I editing by ADAR1 as a key modifier of the cellular response to endogenous dsRNA.

POSTER 20

GENOME SEQUENCING OF RICKETTSIA-LIKE ORGANISM ISOLATED FROM FARMED CHINOOK SALMON IN NEW ZEALAND

Dr Edna Gias[1], Dr Jenny Draper[2], Ms Cara Brosnahan[1], Mrs Della Orr[1], Dr Andrew McFadden[1], Dr Brian Jones[1]

1. Investigation and Diagnostic Centres & Response, Ministry for Primary Industries
2. Institute of Environmental Science and Research Limited

Abstract

Rickettsia-like organism (RLO), a gram-negative bacterium, has been increasingly recognised as an important fish pathogen in the salmonid aquaculture industries. *Piscirickettsia salmonis* was the first described RLO affecting fish and causes a systemic infection known as piscirickettsiosis. The bacterium was first isolated in farmed coho salmon, *Oncoryhnchus kisutch* (Walbaum), during a high mortality epizootic in Chile in 1989 which resulted in huge economic losses. Occurrences of rickettsial septicaemia have been reported in farmed salmonids in many countries and in various non-salmonids species. New Zealand rickettsia-like organism (NZ-RLO) was first isolated during a 2015 high mortality event occurred among Chinook salmon (*Onchorhyncus tshawytscha*) farmed in the Marlborough Sounds. This is the first report of RLO isolated from farmed salmonids in New Zealand. Pyrosequencing was performed to further characterise NZ-RLO and to allow phylogenetic comparison of the isolate with other RLOs and P. *salmonis*.

POSTER 21

CHALLENGES OF WHOLE GENOME BISULPHITE SEQUENCING ON THE HISEQ X TEN

Dr Shalima Nair[1,2], Dr Phuc-Loi Luu[1], Ms Wenjia Qu[1], Prof. Susan Clark[1,2]

1. Genomics and Epigenetics Division, Garvan Institute of Medical Research, Darlinghurst, New South Wales 2010, Australia
2. St Vincent's Clinical School, Faculty of Medicine, University of New South Wales, Darlinghurst, New South Wales 2010, Australia

Abstract

DNA methylation profiling is known to play an important role in both prognostic and diagnostic studies. Among the different DNA methylation profiling technologies, whole genome bisulphite sequencing is considered the gold standard for assaying genome-wide DNA methylation at single base resolution. However, the high sequencing costs for the depth of coverage obtained limits its application in both basic and clinical research. To achieve 30X coverage of the human bisulphite genome, at least four lanes of HiSeq2500 sequencing needs to be performed. The advent of the HiSeq X Ten opens up the possibilities of achieving 30X coverage of the human bisulphite genome at a much lower cost. However, due to the low diversity in base composition of the bisulphite libraries and the smaller fragment size obtained during library prep, achieving 30X coverage per lane of the HiSeq X Ten is challenging. Here, we compared two different library prep methods and different levels of spike in control DNA on the HiSeq X and we now successfully achieve 16 to 22X coverage per lane with a Q30 score > 90 and cluster passing filter ranging between 75-82%. The percentage of mapped reads that we achieve is between 80 – 90%, with the duplication read ranging between 10 - 20%. We observed that the quality of sequencing output is sensitive to the library prep method adopted, the insert size of the library, the accuracy in estimating the library quantity, the absence of adaptor contamination and the spiking concentration of a genome with balanced base composition. To process the big data output from the HiSeq X Ten, we also developed a bioinformatics pipeline package, called WGBS10xmap that can support long bisulphite reads up to 1000bp.

Our pipeline gives the general quality control metrics of the sequencing output, MDS plots of sample clustering, SNP calling and can give bigwig files that can be used for visualization in IGV or UCSC browser.

POSTER 22

DEVELOPMENT OF ROBUST, LOW DNA INPUT GENOTYPING BY SEQUENCING (GBS)

Dr Gai McMichael[1], Jafar Jabbari[1], Rust Turakulov[1], Lavinia Gordon[1], John Stephen[1], Kirby Siemering[1], Matthew Tinning[1]

1. Australian Genome Research Facility Ltd.

Abstract

Background: Genotyping by sequencing (GBS) is a restriction enzyme based method that produces a reduced representation of the non-repetitive (low copy) regions of the genome. GBS can reproducibly generate a large number of SNPs in complex genomes that lack reference sequences. This approach is not only cost-effective but highly suitable to multiplexing, thus making it ideal for high-throughput sequencing. GBS has largely been applied as a genetic screening tool assisting in plant breeding and crop improvement.1,2 It is also an important genetic tool for examining genetic diversity to further understand the evolutionary relationship between groups of organisms. The most significant GBS technical issue is input DNA quality and quantity. DNA is not always obtained from optimal sources. For many individual samples e.g. herbarium and museum specimens, the yield of DNA is a limiting factor. In samples that are unique and irreplaceable, preserving the specimen is a high priority, thus only a small part will be used for DNA extraction i.e. legs of insects. Low DNA yield can also be a limiting factor in conservation and ecological specimens i.e. native plants, marine invertebrates.3. In response, AGRF assessed the development of a low input DNA GBS service version, suitable for input of 50 ng (5 ng/µl) from a complex plant genome.

Materials and methods: DNA samples from 47 tedera plants including two parents and 45 progenies were normalised to 200, 100 and 50 ng input reactions, respectively. Samples were sequenced as a multiplex pool on an Illumina HiSeq V4 for 100 cycles. Sequencing reads were quality filtered, de-multiplexed and data were analysed according to our current GBS pipeline using stacks software4.

Analysed data compared reproducibility between the three different DNA input amounts in terms of the number of detected and shared tags, sequencing depth (reads per tag) and loss of heterozygosity.

Results: Similar fragment size distribution was seen for the 200, 100 and 50 ng input levels. Total number of tags was 24,505, 27,631 and 26,680 respectively. The proportion of common tags ranged from 22% for 200 ng, 16% for 100 ng and 17% for 50 ng. There was no apparent loss of heterozygosity between the different inputs. Overall concordance among the different input amounts was very high indicating robustness of the protocol for different input quantities. Based on these results minor modifications were made to our initial GBS protocol to provide a low input GBS service of 50 ng. The low input protocol has been validated and implemented in our laboratory and is now a routine GBS service. The substantial reduction in DNA input from 200 to 50 ng opens many new avenues and opportunities, thus providing a cost-effective genotyping solution for researchers where previously DNA has been the limiting factor.

References:

1. Poland J.A., Brown P.J., Sorrells M.E. and Jannink J., 2012 Development of high- density genetic maps for barley and wheat using a novel two-enzyme genotyping by sequencing approach. *PLoS One* 7:e32253.
2. Elshire R.J., Glaubitz J.C., Sun Q. et al. 2011 A robust, simple genotyping by sequencing (GBS) approach for high diversity species. *PLoS One* 6: e19379.
3. Ali O.A., O'Rouke S.M., Amish S.J. et al. 2016 RAD Capture (Rapture): Flexible and efficient sequence-based genotyping. *Genetics* 202: 389-400.
4. Catchen J., Hohenlohe P.A., Bassham S. et al. 2013 Stacks: an analysis tool set for population genomics. *Molecular Ecology* 22: 3124-3140. (http://creskolab.uoregon.edu/stacks/).

POSTER 23

PREDICTING EXON SPLICING CHANGES TRIGGERED BY EPIGENETIC PROFILES

Dr Denis Bauer[1], Mr Florian Heyl[1,2], Dr Robert Dunne[1], Dr Laurence Wilson[1], Dr. Jason Ross[1],

1. CSIRO, Health and Biosecurity, North Ryde, Sydney
2. Freie Universitaet Berlin, Berlin, Germany

Abstract

Splicing transcripts is an essential and highly precise regulated process, which elevates the complexity in the genome but also increases the susceptibility to malfunction, such as cancer. While it has been well documented how sequence motifs and individual variations influence splicing events, less focus has been given to the direct role of DNA methylation despite evidence of its involvement. Recent studies showed that the transcriptional repressor CTCF, the multifunctional protein MeCP2 and the heterochromatin protein HP1 may have an influence on the splicing machinery and that their binding activity in turn is mediated by local DNA methylation (Lev Maor et al, Trends in Genetics, 2015).

We therefore investigate how the methylation profile of the genomic environment surrounding the exon affects splicing by using matching whole genome bisulphite and whole transcriptome sequencing data. We first identified exons that are spliced out of actively expressed genes using a probabilistic model developed by Xiong et al. (Science, 2015). We observe that on average spliced out exons have a 2-fold increase in methylation rate for both the up- and downstream intron compared to the retained exons (150bp region surrounding the exon, p-value < 1.7 x 10-197), while the methylation of the exon itself remains stable (p-value=0.456). However, for the extreme ends of the methylation spectrum (top and bottom quantile), both intron and exon methylation influences splicing (p-value < 2.5 x 10-31, and < 6.35 x 10-76, respectively). This is not due to the whole area being methylated as, for example, low methylated exons have a smaller methylation rate than the surrounding introns, even taking the cytosine content into account (~2.5-fold decrease between exons and introns).

This indicates that gene activity and splicing in particular is regulated by a very complex interconnected local methylation profile. We therefore train a deep-learning approach to build a model on these interacting features over thousands of base pairs to predict whether an individual exon is excluded or retained. We achieve a very high prediction accuracy, which demonstrates the direct impact of local methylation profile on splicing events.

POSTER 24

THE RECYCLED HUMAN GENOME: DEEP TRANSCRIPTIONAL ANALYSIS OF PSEUODGENES REVEALS THE SCALE OF LINE1 MEDIATED EXON-SHUFFLING FORMING CHIMERIC GENE ISOFORMS

Dr Daniel Thomson[1], A/Prof Marcel Dinger[1]

1. Garvan Institute Of Medical Research

Abstract

Mutated gene copies called pseudogenes are ubiquitous in mammalian genomes, the majority of which are products of ancestral retrotransposition. Evidence of pseudogene transcription has ignited interest in the potential for function as noncoding RNA, however characterization is obfuscated by a lack of comprehensive transcriptomic annotation.

For this reason we applied a pseudogene RNA CaptureSeq methodology, revealing that >30% of annotated pseuodgenes (Gencode v19) are transcribed in human brain. This enabled us to identify an unprecedented number of pseudogenes that are incorporated into adjacent genes as novel isoforms, resonant of a mechanism referred to as 'exon shuffling'. Our transcriptional analysis reveals the scale by which exon shuffling has shaped the human genome, revealing an expanded protein coding genome (exome) stitched together from pseudogenes which were previously thought to be non-functional DNA relics, or discretely expressed noncoding RNA.

Continued next page

We interrogate the mechanism of pseudogene exon shuffling highlighting the role of L1 mediated retrotransposition. We show L1 transcription in somatic tissue is linked to pseudogenes, providing the first evidence of 3′ transduction of pseudogene sequences in somatic tissue, observed as chimeric L1/pseudogene transcripts.

This work highlights that pseudogenes are integral to an ongoing, active process by which higher eukaryotes recycle functional genetic components through retrotransposition, creating chimeric gene isoforms contributing to an evolving complexity of function.

POSTER 25

ENHANCER LANDSCAPE OF TUBERCULOSIS-INFECTED MURINE MACROPHAGES

Ms Elena Denisenko[1], Prof. Frank Brombacher[2], Dr. Sebastian Schmeier[1]

1. Massey University, Institute of Natural and Mathematical Sciences, 0630 Albany, Auckland, New Zealand
2. International Centre for Genetic Engineering and Biotechnology, University of Cape Town, South Africa

The area of mammalian transcriptional regulation progressed remarkably with recent advances in high-throughput technologies. Enhancers emerged as crucial regulatory DNA elements capable of activating transcription of target genes at distance in an orientation-independent manner. Enhancer mediated transcriptional control involves direct interaction between enhancers and promoters of target genes via DNA looping and exerts itself in a highly tissue-specific manner [1]. A recent discovery revealed that enhancers are bi-directionally transcribed to generate divergent enhancer RNAs (eRNAs) [2]. Moreover, bi-directional transcription was shown to positively correlate with transcription of the corresponding target genes. These findings gave a rise to a novel method for identification of enhancers based on the co-occurrence of closely located divergent eRNAs [3]. However, the specific roles of enhancers in different cells and tissues remain largely unexplored.

In this study, we analysed the transcribed enhancer and gene promoter usage in murine macrophages on a genome-wide scale using data from cap analysis of gene expression (CAGE) followed by high-throughput sequencing [4]. We focused on enhancers controlling the response of murine macrophages upon infection by *Mycobacterium tuberculosis (Mtb)*, the causative agent of tuberculosis. Integrating high-throughput conformation capture and histone modification data, we identified a set of robust enhancers active in murine macrophages and link to their corresponding target genes. Our analysis revealed a set of transcribed enhancers with an increased eRNA production upon *Mtb* infection. These eRNA changes correlate with an increased expression of proximal immunologically relevant genes. Furthermore, we find that active enhancers contain binding sites for transcription factors with established roles in inflammatory response. In summary, our study provides insights into novel mechanisms of genome-wide enhancer-mediated control of transcription of *Mtb*-responsive genes.

References:

1. Shlyueva, D., G. Stampfel, and A. Stark, *Transcriptional enhancers: from properties to genome-wide predictions*. Nat Rev Genet, 2014. **15**(4): p. 272-86.
2. Kim, T.K., et al., *Widespread transcription at neuronal activity-regulated enhancers*. Nature, 2010. **465**(7295): p. 182-7.
3. Andersson, R., et al., *An atlas of active enhancers across human cell types and tissues*. Nature, 2014. **507**(7493): p. 455-461.
4. Kanamori-Katayama, M., et al., *Unamplified cap analysis of gene expression on a single-molecule sequencer*. Genome Research, 2011. **21**(7): p. 1150-1159.

POSTER 26

DECIPHERING THE REGULATORY NETWORK OF MICRORNAS IN TUBERCULOSIS INFECTED MURINE MACROPHAGES

Mr Daniel Ho[1], Miss Elena Denisenko[1], Dr Frank Brombacher[2], Dr Sebastian Schmeier[1]

1. Massey University, Institute of Natural and Mathematical Sciences, 0630 Albany, Auckland, New Zealand
2. International Centre for Genetic Engineering and Biotechnology, University of Cape Town, South Africa

Abstract

Tuberculosis (TB) is an infectious disease still causing millions of deaths worldwide. The disease is caused by *Mycobacterium tuberculosis* (*Mtb*), an intracellular pathogen that uses macrophages as a host for replication. MicroRNAs (miRNAs) are small regulatory RNAs that silence gene functions post-transcriptionally. Recent studies indicate miRNAs may have prominent roles in cellular host-pathogen interactions [1]. During TB infection, macrophages engulf Mtb and try to kill it by phagocytosis and secretion of antimicrobial molecules [1]. The outcome of the disease depends highly on *Mtb's* strategies to subvert the immune responses of macrophage hosts. The aim of this study is to advance our understanding of the regulatory network that controls key miRNAs in the macrophage host during *Mtb* infection by means of Cap Analysis of Gene Expression (CAGE) and next-generation sequencing (NGS).

A murine miRNA regulatory network was constructed by combining a network of miRNA-controlling transcription factors (TFs) with a miRNA target gene network. We established the upstream and downstream regulatory networks of miRNAs by integrating public experimental data sets, high-throughput ChIP-seq data for TF binding and miRNA–target interaction data from relevant databases, respectively. The final network places miRNAs at the centre of a comprehensive regulatory network of TFs, miRNAs and their target genes. This network represents a useful resource for investigating miRNA functions and control.

Subsequently, we populated the network with CAGE-derived expression data either for normal or Mtb-infected macrophages. We used network analysis to determine key miRNAs and their transcriptional regulators during the infection process. As a result, we identified a core set of TFs, including Maff, Fosl1, Nfkb1, Rel, and Egr1 [2], which exert strong control over many TB immune-related miRNAs in our network. Our results also implicate two highly expressed miRNAs, miR-149 and miR-449, to work in unison with miR-155, a miRNA known to be stimulated during *Mtb* infection [3]. Our data suggests that these miRNAs co-regulate a set of immune response related genes in order to confer survival benefits to early stage *Mtb* infection. More generally, our study highlights the use of large-scale NGS data sets to derive a complex regulatory network model and its application to immunological research.

References:

1. Guler, Reto and Brombacher, *Frank, Host directed drug therapy for tuberculosis*. Nat Chem Biol, 2015. 11(10): p. 748-751.
2. Roy, Sugata, Schmeier, Sebastian, Arner, Erik, Alam, Tanvir, Parihar, Suraj P, Ozturk, Mumin, Tamgue, Ousman, Kawaji, Hideya, de Hoon, Michiel JL, and Itoh, Masayoshi, *Redefining the transcriptional regulatory dynamics of classically and alternatively activated macrophages by deep CAGE transcriptomics*. Nucleic acids research, 2015. 43(14): p. 6969-6982.
3. Kumar, Ranjeet, Halder, Priyanka, Sahu, Sanjaya K., Kumar, Manish, Kumari, Mandavi, Jana, Kuladip, Ghosh, Zhumur, Sharma, Pawan, Kundu, Manikuntala, and Basu, Joyoti, *Identification of a novel role of ESAT-6-dependent miR-155 induction during infection of macrophages with Mycobacterium tuberculosis*. Cellular Microbiology, 2012. 14(10): p. 1620-1631.

POSTER 27

THE INVOLVEMENT OF KIDNEY DNA METHYLATION IN BLOOD PRESSURE REGULATION

Miss Ingrid A Wise[1], Dr Priscilla Prestes[1], Professor Maciej D Tomaszewski[2], Professor Fadi J Charchar[1]

1. Faculty of Science & Technology, Federation University Australia.
2. Faculty of Science & Technology, Federation University Australia.
3. Faculty of Human & Medical Sciences, University of Manchester.
4. Faculty of Science & Technology, Federation University Australia; Department of Physiology, University of Melbourne.

Abstract

Background and aims: Increasing evidence suggests that epigenetic modifications such as DNA methylation (5mC) is important to the development of essential hypertension, and that changes in DNA methylation of blood cells is associated to blood pressure (BP). So far there has been no studies of epigenetic changes in the kidney - an important effector organ in BP regulation. The aim of this study was to compare the global and gene specific methylation status in the kidney between normal and hypertensive subjects.

Methods and results: We used 96 human renal tissue samples from the TRANScriptome of RenaL HumAN TissueE (TRANSLATE) Study to measure DNA methylation. TRANSLATE consists of carefully characterised collections of "apparently healthy" specimens of human kidneys. DNA was extracted from peripheral blood leukocytes and kidney tissue using the DNeasy blood and tissue Qiagen kit. Global methylation was measured by ELISA assay to determine the percentage of 5mC and loci specific methylation status was determined using Infinium HumanMethylation 450K array (Illumina®, Australia).

A significant negative relationship was found in the renal samples between 5mC% and systolic (SBP) and diastolic (DBP) blood pressure readings (SBP r=-0.25, P=<0.05), DBP r=-0.32, P=<0.01). This correlation was also evident when BP is adjusted for hypertensive medication effects (adjusted SBP P=<0.05, adjusted DBP P=<0.01). There was no significant relationship in DNA extracted from peripheral blood leukocytes between 5mC% and BP reading. We found 275 loci differentially methylated between hypertensive and normotensive individuals.

Conclusions: DNA methylation is an important molecular mechanism for BP and hypertension in humans

POSTER 28

USING DNA METHYLATION TO IDENTIFY THE FETAL CELL OF ORIGIN FOR CHILDHOOD ACUTE LYMPHOBLASTIC LEUKAEMIA

Mr Abdulmonem Alsaleh[1], Dr Robert Weeks[1], Professor Ian Morison[1]

1. Department of Pathology, University Of Otago

Abstract

Childhood acute lymphoblastic leukaemia (ALL) is a disease that originates before birth from the proliferation of neoplastic lymphoid precursor cells in the bone marrow. Previously, we identified dense, biallelic methylation of the TES promoter as the most common molecular event in ALL being present in over 90% of B-ALL and 70% of T-ALL cases1. Exploring publicly available human DNA methylation array datasets, we found about 200 genes that are differentially methylated in ALL suggesting a common epigenetic signature in ALL. The large number and consistency of the epigenetically modified genes found in our analysis raise the possibility that methylation-induced gene silencing in ALL may not be an acquired cancer-related phenomenon. We propose the presence of a distinct population of normal fetal lymphocytes that carry an epigenetic profile similar to that reported in ALL. In addition, we propose that if this fetal lymphocyte population does not regress after birth, there is a risk of transformation to B or T ALL.

We developed an assay to detect methylated TES alleles using a deep targeted sequencing approach (MiSeq; Illumina). To date, we have examined blood samples from four neonates from Dunedin Hospital, and 10 cord blood samples from Southland Hospital. Remarkably, we detected the ALL-like methylation profile, i.e., TES methylated alleles, in one of the premature babies (3.2% of the CD19+ B cells of a four week-old, 28 week-gestation baby), and in three of 10 sequenced cord blood samples (1.3%, 0.16%, and 1.2% of CD19 negative cells). Importantly, methylation of TES has never been observed in normal adult blood. The next stage of the project is to purify stem cells, T cells and B cells and their subpopulations using fluorescence-activated cell sorting and then search for enrichment of the methylated TES alleles (and other methylated genes) in each population of cells.

References

1. Weeks RJ, Kees UR, Song S, Morison IM. Silencing of TESTIN by dense biallelic promoter methylation is the most common molecular event in childhood acute lymphoblastic leukaemia. Molecular Cancer 2010;9:163.

POSTER 29

EVALUATION OF RAPID LOW-INPUT PCR-FREE LIBRARY PREPARATION FOR HIGH-THROUGHPUT WHOLE GENOME SEQUENCING

Dr Jiang Tao[1,2,3], Mr Ben Lundie[1,3], Dr Marcel Dinger[1,2,3]

1. Kinghorn Centre for Clinical Genomics, Garvan Institute of Medical Research
2. St Vincent's Clinical School, Faculty of Medicine, UNSW
3. Genome.One

Abstract

Genome.One provides clinical Whole Genome Sequencing (WGS) through the Kinghorn Centre for Clinical Genomics (KCCG) to enhance the lives of patients, families and communities across the world. Our sequencing service provides a minimum mean whole genome at 30X coverage on the Illumina HiSeq X Ten system using SeqLab solution, which is designed to provide the necessary tools for laboratory efficiencies and scalability. SeqLab includes two workflows: TruSeq Nano and TruSeq PCR-Free. When sufficient amounts of genomic DNA are available for library preparation, the PCR-Free method improves coverage and uniformity by reducing bias and gaps, and therefore significantly improves the sensitivity and specificity when calling insertion-deletions (indels) and copy number variants (CNVs). The demand for PCR-free WGS is growing rapidly, however some of users are not able to provide 1 ug high quality genomic DNA as required in the Illumina SeqLab PCR-free workflow, or expect faster turn around time to receive sequencing data.

Continued next page

Thus, an alternative PCR-free workflow is needed to meet user requirements. We prepared libraries for HiSeq X Ten sequencing with the NA12878 control DNA using KAPA Hyper DNA library preparation kit and compared results to the SeqLab TruSeq PCR-Free method. In comparison to the Illumina PCR-free method, the KAPA Hyper Prep workflow has reduced turn-around time and hands-on time by combining end repair and dA tailing into one step and minimizing cleanup steps. The analysis of sequencing data showed that KAPA Hyper prep provides better specificity and comparable sensitivity in detecting variants. The faster library preparation, coupled with the reduced input DNA requirements (500 ng), provides the advantages of PCR-free methods for processing both research and clinical samples for WGS on HiSeq X Ten system.

POSTER 30

ANTISENSE PNA AND ANTIBIOTICS, SYNERGISTS AGAINST GRAM-NEGATIVE BACTERIA

Jaime Iván Castillo Silva[1,2], MSc Marcin Równicki[1,3], PhD Monika Wojciechowska[1], Prof. Joanna Trylska[1]

1. Biomolecular Machines Laboratory, Centre of New Technologies, S. Banacha 2c, 02-097 Warsaw.
2. Colegio "La Paz" A.C. La Paz S/N, 22020 Tijuana, B.C., México.
3. College of Inter-Faculty Individual Studies in Mathematics and Natural Sciences, University of Warsaw, Żwirki i Wigury 93, 02-089 Warsaw.

The rapid growth of antimicrobial resistance and the insufficient progress achieved to fight it, has made the development of novel strategies to attack bacteria an urging matter [1]. The combination of agents seems a promising strategy to enhance therapy [2], which can optimistically lead to antimicrobial synergy. In the present study, the combination effects between antisense peptide nucleic acids (PNA) and conventional antibiotics were investigated against gram-negative bacterium *Escherichia coli*. To improve uptake, PNAs were conjugated to a cell-penetrating peptide (KFF)3K.

PNAs were targeting mRNA transcripts encoding bacterial growth-essential proteins: the acyl carrier protein (gene *acpP*) and the chromosomal replication initiator protein (gene *dnaA*). For this study, antibiotics with diverse inhibition mechanisms were employed: aminoglycosides, tetracyclines and polypeptide antibiotics. To confirm antibacterial susceptibility and synergistic effects the minimal inhibitory concentration and fractional inhibitory concentration index were determined. The results revealed a synergistic interaction between rifampicin and PNA anti-*acpP*, postulating cell wall targeted antibiotics as an attractive agent to improve the uptake of PNA and implying the possibility to assess synergy between mRNA targeted PNA and antibiotics targeting different pathways.

References:

1. D. M. Livermore, "The need for new antibiotics," Clin. Microbiol. Infect., vol. 10, pp. 1–9, 2004.
2. J. I. M. O. Neill, "Tackling drug-resistant infections globally: final report and recommendations. The review on antimicrobial resistance.," no. May, 2016.

POSTER 31

OPTIMUSPRIME: MULTIPLEX PRIMER DESIGN TOOL FOR HI-PLEX SEQUENCING

Mr Edmund Lau[1], Dr Benard Pope[1,3], Dr Daniel Park[1,2]

1. Victorian Life Science and Computation Initiative, The University of Melbourne
2. Department of Pathology, The University of Melbourne
3. Computer Science Department, The University of Melbourne

Abstract

In this work we will present Hi-Plex, a novel Massively Parallel Sequencing (MPS) platform, and a multiplex PCR primer design software that meets the challenges the platform presents.

Hi-Plex Sequencing
The advent of MPS has dramatically reduced the cost and increased the throughput of DNA sequencing. With cost-efficiency in mind, a number of methods have been developed for targeted MPS, in which only specified genomic regions are represented. However, these, variously, are compromised by issues of relative expense, accuracy, requirement for specialist equipment and the cumbersome nature of protocols.[1] Hi-Plex (www.hiplex.org) is a novel platform using a highly multiplexed PCR-based approach for targeted MPS that employs 'tagged' gene-specific primers to 'seed' and universal primers to 'drive' the majority of amplification, minimising bias caused by differing gene-specific primer efficiencies.[1,2,3] Hi-Plex had been employed in several large-scale genetic screening studies for breast and ovarian cancer[4,5,6] and its applicability has been demonstrated for dried blood spot-derived DNA which is an abundant source of DNA that presents challenges due to relative fragmentation for example[7].

In addition to offering low cost, a simple operating protocol and compatibility with the highest fidelity thermostable DNA polymerases available throughout, Hi-Plex chemistry allows the target insert size to be defined within a small and tight range thus enabling complete overlap of sequencing read pairs. Bundled with Hi-Plex is read-pair-overlap-considerate variant calling software which enables high statistical confidence in variant call, which translates to high sensitivity and a low false discovery rate.[8,9] Furthermore, tight amplicon size definition contributes to uniformity of amplification efficiency and allows a size selection step which greatly reduces off-target amplification artefacts.

Primer Design Software for Hi-Plex
Primer design is critical to the success of a multiplex PCR assay. Given a panel of DNA regions to be PCR amplified, the central concern for primer design software is to find a set of primers whose amplicons completely and uniformly cover the specified regions while minimising off-target artefacts.
As a computational problem, we need 1) accurate and discriminative scoring that recognises good solutions and 2) efficient searching of the design space which grows exponentially in the size of the tiled regions.
We devised a statistical scoring method that measures the priming effectiveness of a given primer compared to the mean of the population of all possible primers. Priming effectiveness is considerate of several factors, including the primer's melting temperature and its propensity towards off-target binding and dimerisation. The algorithm then uses dynamic programming to exhaustively visit an exponentially large solution space in linear time and return the best scoring combination of primers. This automated primer design software confers the advantage of having a well-defined optimal solution and being exhaustive in its search for potential solutions.

References

1. Nguyen-Dumont, T. et al. Anal. Biochem. 442, 127–129 (2013).
2. Nguyen-Dumont, T., Pope, B.J., Hammet, F., Southey, M.C. & Park, D.J. Biotechniques 55, 69–74 (2013).
3. Nguyen-Dumont, T. et al. Biotechniques 58, 33–36 (2015).
4. Hasmad, H.N. et al. Gynecol. Oncol. 141, 318–322 (2016).
5. Nguyen-Dumont, T. et al. Breast Cancer Res. Treat. 149, 547–554 (2015).
6. Nguyen-Dumont, T. et al. BMC Med. Genomics 6, 48 (2013).
7. Nguyen-Dumont, T. et al. Anal. Biochem. 470, 48–51 (2015).
8. Pope, B.J., Nguyen-Dumont, T., Hammet, F. & Park, D.J. Source Code Biol. Med.
9. Park, D.J. et al. BMC Bioinformatics 17, 1–7 (2016).

POSTER 32

IDENTIFYING GENETIC POLYMORPHISMS IN PHARMACOGENES USING NANOPORE SEQUENCING

Dr Simone Cree[1], Dr. Simran Maggo[1], Ayorinde Adehin[2], Professor Martin Kennedy[1]

1. University Of Otago
2. Department of Pharmaceutical Chemistry, Obafemi Awolowo University

Abstract

We are interested in exploring the potential of the MinION for rapid and accurate analysis of pharmacogenes in multiple clinical samples. The liver enzymes cytochrome P450s are involved in metabolising many important drugs and xenobiotics. The cytochrome P450 2D6 involved in metabolism of clinically used drugs has been well studied and displays considerable genetic variability. On the other hand, genetic polymorphisms in the cytochrome P450 1A2 gene are less studied whereas cytochrome P450 2A6 is found to be highly polymorphic. We are investigating the ability of the MinION nanopore sequencer to detect sequence variation in long PCR amplicons (7-8 Kb) which encompass the full *CYP2D6, CYP1A2* and *CYP2A6* genes. Sequence reads are extracted and analysed using various tools for comparison. As kit chemistry has improved we have seen corresponding increases in mean length of aligned segments with 85% accuracy. Our most recent runs have successfully used oligonucleotide indexing to carry out multiplexing of samples.

POSTER 33

INVESTIGATING THE ROLE OF HISTONE VARIANT H2A.B1 IN THE REGULATION OF CANCER/TESTIS (CT) ANTIGENS IN HODGKIN LYMPHOMA.

Dr Sebastian Kurscheid[1], Dr Tatiana Soboleva[1], Prof David Tremethick[1]

1. Department of Genome Science, The John Curtin School of Medical Research - The Australian National University, Canberra, Australia

Abstract

Histone variants play an important role in the regulation of chromatin structure. Our group previously demonstrated that the mouse homologue of human histone variant H2A. B1 (H2A.Bbd), H2A.Lap1, is required for the activation of transcriptional programs during spermatogenesis [1]. The expression of CT antigens is normally highly regulated and restricted to male germ cells in the testis. When gene regulation is disrupted, as is the case in cancer, CT antigens can be expressed in tumours [2]. Due to their high tissue specificity, CT antigens are promising candidates for targeted cancer therapeutics, and have also been suggested as biomarkers in e.g. ovarian cancer [3], breast cancer [4] and in Hodgkin lymphoma [5]. The emerging role of histone variants in cancer [6] has prompted us to investigate the role H2A.B1 in cancer biology. Following a bioinformatics analysis of published gene expression datasets, we found that the Hodgkin lymphoma derived cell line L1236 showed the highest expression levels of the H2A.B1-encoding gene H2AFB3. This led us to study the association between H2A.B1 and the expression of CT antigens in L1236.

Continued next page

We performed ChIP-Seq using antibodies raised against H2A.B1 and another histone variant, H2A.Z, in L1236 cells and complemented this data by RNA-Seq from the same cell line. In the first stage of our analysis we studied the occupation pattern of H2A.B1 across the promoters of all genes and found that k-means clustering of the coverage maps identified four groups. Integration of gene expression data showed that H2A.B1 clusters 1-4 are associated with decreasing mean expression levels. We then used publicly available data to identify which CT antigens are associated with which H2A.B1 clusters. This analysis revealed that the vast majority of CT antigens expressed in L1236 were found in H2A.B1 clusters 1 and 2 thus implicating H2A.B1 in the control of CT antigen expression.

References:

1. Soboleva, T.A., et al., *A unique H2A histone variant occupies the transcriptional start site of active genes.* Nat Struct Mol Biol, 2012. 19(1): p. 25-30.
2. Scanlan, M.J., et al., Cancer/testis antigens: *an expanding family of targets for cancer immunotherapy.* Immunol Rev, 2002. 188: p. 22-32.
3. Mirandola, L., et al., *Cancer testis antigens: novel biomarkers and targetable proteins for ovarian cancer.* Int Rev Immunol, 2011. 30(2-3): p. 127-37.
4. Yao, J., et al., *Tumor subtype-specific cancer-testis antigens as potential biomarkers and immunotherapeutic targets for cancers.* Cancer Immunol Res, 2014. 2(4): p. 371-9.
5. Chen, Y.T., et al., E*xpression of cancer testis antigen CT45 in classical Hodgkin lymphoma and other B-cell lymphomas.* Proc Natl Acad Sci U S A, 2010. 107(7): p. 3093-8.
6. Vardabasso, C., et al., *Histone variants: emerging players in cancer biology.* Cell Mol Life Sci, 2014. 71(3): p. 379-404.

POSTER 34

GENOTYPING-BY-SEQUENCING FOR DIVERSE APPLICATIONS IN BIOLOGY

Dr Jeanne Jacobs[1], Shannon Clarke[2], Marty Faville[3], Andrew Griffiths[3], Mingshu Cao3, Rachel Tan[3], Tracey van Stijn[2], Rayna Anderson[2], Rachael Ashby[2], Rudiger Brauning[2], Alan McCulloch[2], Ken Dodds[2], John McEwan[2]

1. AgResearch Ltd, Lincoln Science Centre, Private Bag 4749, Christchurch 8140, New Zealand
2. AgResearch Ltd, Invermay Agricultural Centre, Private Bag 50034, Mosgiel 9053, New Zealand
3. AgResearch Ltd, Grasslands Research Centre, Private Bag 11008, Palmerston North 4442, New Zealand

Abstract

Genotyping-by-Sequencing (GBS) is a method used to develop rapid and cost-effective high-density genetic marker data for diverse applications in biology. We describe the Genomics for Production and Security programme in which we developed the infrastructure and skill base required to apply GBS for different purposes across a wide range of species. Thus far, we have optimised GBS in 25 different species encompassing plants, mammals, shellfish, fish, birds, and insects. The methodology has been scaled from small sample sizes (< 100) for diversity studies up to many thousands in animal and plant breeding programmes where GBS underpins genomic selection. GBS is also being extended into environmental metagenomic, population genetics and conservation genetics studies. Continuous improvements in wet-lab methods have enabled increased quality and quantity of data generated. Furthermore, data analysis has been enhanced through improved bioinformatic pipelines tailored to each species, including a novel statistical method (KGD) designed specifically for utilising GBS data to develop genomic relationship matrices (Dodds et al., BMC Genomics (2015) 16:1047). Components of the data analysis pipeline are made available in a public Github repository (https://github.com/Agresearch). Many studies are collaborations with New Zealand Crown Research Institutes, New Zealand universities & research organisations, New Zealand Genomics Ltd, as well as commercial entities, both nationally and internationally.

POSTER 35

ICE-T: INTEGRATED CELL EXPRESSION TOOLKIT FOR LOW INPUT TRANSCRIPT PROFILING

Mr Chris Harris[1], Ms Tanis Godwin[1], Dr Peter Stockwell[1], Mr Anthony Shaw[1], Dr Mik Black[1], Prof Parry Guilford[1], Dr Robert Day[1]

1. Department of Biochemistry, University of Otago, Dunedin, New Zealand.

Abstract

Transcriptome profiling at high cellular resolution has many applications in biology but is hampered by a need to boost the representation of RNA to guard against losses during processing and enable efficient use of high throughput technologies.

The Integrated Cell Expression Toolkit provides a core set of processes that enable the researcher to choose from established PCR and IVT amplification strategies or combine the two together. Here we implement different strategies and demonstrate the flexibility and capability of the combined PCR-IVT approach by sequencing a range of cell mixtures and individual cells. This novel approach detected 80% of the top 50 differentially expressed genes of a discrepant single cell in a background cell population by low pass sequencing. Different implementations of ICE-T enable either full length or 3′ end RNA profiling and provides the researcher with the opportunity to exploit synergies achieved by using distinct library types on the same sequencing run. Optional workflows also enable molecular counting and increased multiplexing through the combination of tagmentation and inline barcoding.

POSTER 36

DIVERSIFICATION OF INNATE IMMUNE RESPONSES BY TRANSCRIPTIONAL MECHANISMS

Suzanne Butcher[1], Edward Huang[1], Ashley Waardenberg[2], Kelly Hitchens[3], Othmar Korn[3], Rowland Mosbergen[1], Tyrone Chen[1], Dipti Vijayan[4], Anthony Beckhouse[5], Mark Walker[6], Tim Barnett[6], James Fraser[6], Antje Blumenthal[7], Stuart Kellie[6], Mikael Boden[6], Masayoshi Itoh[8,9,10], Timo Lassman[8,9], Hideya Kawaji[8,9,10], FANTOM Consortium, Yoshihide Hayashizaki[8,10], Piero Carninci[8,9], Alistair R.R. Forrest[8,9], Christine Wells[1]

1. Department of Anatomy and Neuroscience, University of Melbourne, Parkville, VIC, Australia
2. Victor Chang Cardiac Research Institute, Darlinghurst, NSW, Australia
3. Australian Institute for Bioengineering and Nanotechnology, The University of QLD, QLD, Australia
4. QIMR Berghofer Medical Research Institute, Herston, QLD, Australia
5. Bio-Rad Laboratories Pty Ltd, Gladesville, NSW, Australia
6. School of Chemistry and Molecular Biosciences, The University of QLD, QLD, Australia
7. Diamantina Institute, The University of QLD, QLD, Australia
8. Riken Omics Science Center (OSC), Yokohama, Japan
9. Division of Genomic Technologies, RIKEN Center for Life Science Technologies, Yokohama, Japan
10. RIKEN Preventative Medicine and Diagnosis Innovation Program, Yokohama, Japan

Abstract

It has long been established that innate immune cells elicit a broad inflammatory program as a first-line defence against pathogens. Less well understood are pathogen-specific responses, which vary by amplitude, the timing of inflammatory activity and the type of molecules involved. As the spectrum of innate cell phenotypes produced by these inflammatory programs drives long term adaptive immunity, it is important that the factors underlying their diversity are understood.

Continued next page

Current models of innate immune signalling do not reflect this diversity of response, or provide a mechanism for how divergent responses arise from convergent signalling pathways. This convergence is driven from evolutionary conservation of the genome, however, transcriptomic diversification mechanisms may contribute to diversity in innate immune responses.

We have used Cap Analysis Gene Expression (CAGE) to characterise the monocyte transcriptomes elicited by ten different pathogenic challenges. CAGE maps the location and expression of distinct transcription initiation events. Engagement of multiple transcription start sites (TSSs) by genes in our dataset was prevalent. Half of the 10470 genes expressed by treated monocytes were expressed from multiple TSSs. Multi-TSS genes constituted a majority of differentially expressed genes (66%) and were enriched for biological functions related to the innate immune response, including cytokine and pattern recognition receptor mediated signalling, and apoptosis. Conversely, single-TSS genes were enriched for housekeeping functions, encompassing metabolic processes, and nucleic acid processing and repair. This implicates multi-TSS genes in adapting to extracellular stimuli and adjusting homeostasis during inflammatory responses to infection.

To measure the diversity introduced by engagement of multiple TSSs, we conservatively predicted protein coding outcomes for 5476 TSSs which were differentially expressed, or consistently highly expressed in our dataset. Stringent filtering to remove low-evidence, ambiguously mapped, and non-coding transcript identifiers left us with 262 genes for which we predict multiple protein outcomes. For 33% of this high confidence gene set, we predict expression of isoforms with different domains or sequence features capable of altering protein location, function, or interactions. Our findings implicate multi-TSS engagement as a mechanism with potential to diversify evolutionarily conserved gene networks, presenting an exciting regulatory layer to be considered when deciphering phenotypic diversity within innate immune responses.

POSTER 37

THE GENOMIC LANDSCAPE OF OESOPHAGEAL ADENOCARCINOMA AND ITS PRECURSOR BARRETT'S OESOPHAGUS

Dr Felicity Newell[1], Dr Katia Nones[1], Mrs Kalpana Patel[4], Mr Michael Gartside[4], Associate Professor Lutz Krause[2], Dr Kelly A Loffler[4], Dr Stephen Kazakoff[1], Dr Ann-Marie Patch[1], Dr Luke Hourigan[5], Dr Bradley Kendall[1,5,6], Professor David C Whiteman[1], Mr John V Pearson[1], Dr Nicola Waddell[1], Associate Professor Andrew P Barbour[3,4]

1. QIMR Berghofer Medical Research Institute, 300 Herston Road, Herston 4006, Brisbane, QLD Australia
2. The University of Queensland, Diamantina Institute, Translational Research Institute, Woolloongabba 4102, Brisbane, QLD, Australia
3. Department of Surgery, School of Medicine, The University of Queensland, Princess Alexandra Hospital, Woolloongabba 4102, Brisbane, QLD, Australia
4. Surgical Oncology Group, School of Medicine, The University of Queensland, Translational Research Institute at the Princess Alexandra Hospital, Woolloongabba 4102, Brisbane, QLD, Australia
5. Department of Gastroenterology and Hepatology, Princess Alexandra Hospital, Woolloongabba 4102, Brisbane, QLD, Australia
6. School of Medicine, The University of Queensland, Princess Alexandra Hospital, Woolloongabba 4102, Brisbane, QLD, Australia

Continued next page

Abstract

Oesophageal adenocarcinoma (EAC) is associated with poor survival, with less than 20% of patients surviving for 5 years. The precursor to EAC is metaplastic Barrett's oesophagus, however the molecular characteristics that are associated with the progression from Barrett's to EAC are not yet fully understood. In this study we used whole genome sequencing to examine the genomic landscape of 57 EAC and 22 Barrett's oesophagus samples. Barrett's oesophagus samples included samples from patients who did not progress to EAC, as well as non-dysplastic and high grade dysplastic Barrett's samples. The mutational burden of somatic SNP and indel mutations was similar between high grade dysplastic Barrett's (median of 5.8 mutations per megabase) and EAC samples (median of 5.3 mutations per megabase) but lower in non-progressor and non-dysplastic Barrett's samples. We observed the presence of five mutational signatures in EAC, including the T>G at TT sites signature that is common in EAC and is thought to be a result of exposure to bile and gastric acids. This signature was also found to be present in Barrett's samples. Significantly mutated genes in EAC included TP53, CDKN2A and SMAD4, with mutations for TP53 and CDKN2A also present in Barrett's samples. Copy number and structural variations were frequently observed in EAC samples, including recurrent amplifications across the genome. Copy number variations were low in Barrett's samples, particularly in terms of amplification events. These observations agree with previous studies that many of the molecular events associated with EAC are already present in the Barrett's oesophagus precursor phase.

POSTER 38

WHOLE TRANSCRIPTOME ANALYSIS OF CONSTITUTIVE PD-L1 EXPRESSION IN MELANOMA CELL LINES

Mr Antonio Ahn[1], Dr Aniruddha Chatterjee[1,2], Dr Peter Stockwell[1], Professor Mike Eccles[1,2]

1. Department of Pathology, University of Otago, Dunedin, NZ
2. Maurice Wilkins Centre, NZ
3. Department of Biochemistry, University of Otago, Dunedin, NZ

Abstract

Melanomas commonly express the Programmed Death Ligand-1 (PD-L1) to suppress immune attack. Two monoclonal antibody drugs that inhibit the PD-1/PD-L1 interaction have now been clinically approved in New Zealand due to their ability to achieve durable responses in patients. However there is still a lack of understanding of how PD-L1 is constitutively expressed on melanoma cells, which remains a major barrier to identifying predictive biomarkers. Here we performed whole transcriptome RNA-sequencing on 4 melanoma cell lines without PD-L1 expression and 4 cell lines that constitutively express PD-L1 on its cell surface. Differential gene expression analysis was performed using the edgeR pipeline (quasi-likelihood F-test). There were 56 genes that identified as differentially expressed (FDR corrected $p \leq 0.20$). Gene set analysis using the FRY method identified cytokine regulation as one of the most important biological processes involved in the constitutive PD-L1 group.

POSTER 39

GENOMICS CHALLENGES AND OPPORTUNITIES FOR BIOPESTICIDE DEVELOPMENT

Mr Mikhail Fokin[1], Dr Damien Fleetwood[2], Dr Bevan Weir[3], Dr Silas Villas-Boas[1]

1. School of Biological Sciences, University of Auckland, Auckland, New Zealand
2. Biotelliga Limited, Auckland, New Zealand
3. Landcare Research, Auckland, New Zealand

Abstract

The biological pesticide (BP) market has grown rapidly over the last decade, due to increasing consumer demand for environmentally-friendly technologies and changes in agricultural policies of many countries. BPs provide an alternative to cheaper but often more toxic and persistent synthetic chemical pesticides, but there are significant challenges in equating BP efficacy with their chemical competitors. BPs can include a variety of living organisms, small molecule natural products or both. However, the biological nature of all BPs means that the whole range of "omics" methods can be applied to facilitate their development.

The early stage of BPs product development is the identification of an active substance - usually a naturally occurring strain of microorganism, which is often a rival for the unwanted species Metagenomics and microbial community ecology can be used to identify candidates via negative correlations between potential BPs and pathogens. Through whole genome sequencing of environmental samples it is also possible to access the repertoire of effectors and secondary metabolites, which facilitate the activity of a particular biological agent and are sometimes active in isolation. At the same time whole genome analysis can help with safety assessment by analysing pathogenicity and toxin genes. Growing databases of microbial secondary metabolites provide a good foundation for function prediction.

Having the active biological identified, the next step of BP development involves mass production, which is often difficult or not possible in the original wild type organism. At this stage transcriptomic and metabolomic analysis provides a tool for functional characterisation of the strain/compound and also for genetic modification and heterologous expression to increase production. Among numerous benefits of BPs the complex and multitargeted mode of action provides a great opportunity for resistance management and IPM programmes. Transcriptome wide analysis of target species can be very useful for assessment of effects of BPs on the cell metabolism.

As a case study for the approaches discussed above we present the natural bio-fungicide development derived from the filamentous fungus *Epicoccum purpurascens* (Didymellaceae, Ascomycota) by using a range of "omics" technologies.

POSTER 40

ESTABLISHING THE INFRASTRUCTURE REQUIRED TO SUPPORT A HIGH-THROUGHPUT WHOLE GENOME SEQUENCING FACILITY

Dr Joice Kuroiwa-Trzmielina[1], Justin Stockmyer[1], Dr Maria Lubka-Pathak[1], Dahlia Saroufim[1], Senel Idrisoglu[1], Dr Jiang Tao[1], Aaron Statham[1], Joe Copty[1], Kevin Ying[1], Dr Warren Kaplan[1,2], Dr Marcel Dinger[1,2]

1. Kinghorn Centre for Clinical Genomics, Garvan Institute of Medical Research, Sydney, New South Wales, Australia
2. Faculty of Medicine, St Vincent's Clinical School, University of New South Wales, Sydney, New South Wales, Australia

Continued next page

Abstract

The rapid evolution of whole genome sequencing (WGS) technology culminated in the release of Illumina's HiSeq X Ten sequencing platform, which is capable of generating over 15,000 high quality genomes per year. Since receiving this technology in May 2014, Garvan Institute's Kinghorn Centre for Clinical Genomics (KCCG) has sought to maximise the output potential of this platform through the development of a robust infrastructure capable of supporting high-throughput research sequencing.

DNA samples are accessioned to the KCCG laboratory information management system (LIMS), assigned a receipt date and a storage location; and proceed to initial QC. Initial QC includes assessment of purity, quality, quantity and volume. LIMS allows for the automatic uploading and parsing of QC data. Samples that pass initial QC are assigned to the predefined library preparation workflow. High-throughput library preparation is performed using the Hamilton STAR liquid handling system followed by QC assessment (purity, quality and quantity). LIMS allows for automatic QC Pass/Fail assignment. Libraries that pass QC are scheduled for clustering. Sequencing is performed on the HiSeq X using standard run parameters (Paired-End 151bp). Post-sequencing run QC is determined by assessing instrument metrics captured and collated in LIMS during each run. Lanes of sequence flagged as "Pass" progress to the relevant production informatics pipeline. Raw data (FASTQs) that pass QC (Yield >100Gb, %Q30 bases >75, mean coverage >30x, human:human contamination < 3%) are aligned to the human reference genome (b37d5) using bwa-mem, and duplicate reads marked with novosort. Gold-standard alignment software GATK and variant calling software BWA are used to generate best practice BAM and VCF files. Genomic data are shared with Garvan's collaborators via DNAnexus (cloud-based) or through a dedicated metadata management system at NCI (National Computation Infrastructure, Canberra).

The KCCG team is utilising this unique infrastructure, which has recently received NATA clinical accreditation (ISO 15189), to drive several flagship, population-scale genomic sequencing initiatives, many the first of their kind worldwide.

POSTER 41

GETTING THE MOST FROM AN RNA-SEQ EXPERIMENT: SOME DESIGN CONSIDERATIONS

Dr Susan Corley[1], Marc Wilkins[1]

1. The New South Wales Systems Biology Initiative, School of Biotechnology and Biomolecular Sciences, University of New South Wales

Abstract

RNA-Seq technology is capable of detecting all forms of RNA transcribed from the genome (including mRNA coding for proteins, microRNA, snoRNA, lincRNA, and mtRNA) and is now widely used by research groups as a valuable research tool. A number of decisions need to made prior to sequencing, including the appropriate number of replicates to sequence, blocking strategies, how to filter the RNA prior to sequencing and then whether to use a stranded library preparation kit and whether to employ single-end or paired-end sequencing. The decisions taken at this point will affect the cost of the RNA-Seq experiment and the utility of the data produced. Here we compare the outcomes of read mapping, feature counting and differential expression analysis using different experimental protocols involving single-end or paired-end sequencing and a stranded or non-stranded library preparation protocol. We explore these issues in four experiments involving human or mouse samples. Library preparation and sequencing was performed at the Ramaciotti Centre for Genomics. Reads were mapped to the human or mouse reference (Ensembl GRCh38 or Ensemble GRCm38) using Tophat2and then to features using Subread. Differential expression analysis was performed using edgeR and limma (voom).

POSTER 42

INVESTIGATION INTO THE CAUSE OF ENCEPHALITIS IN KAKî CHICKS

Dr Barbara Binney[1], Dr Kelly Buckle[1], Dr Wlodek Stanislawek[1], Dr Kate McInnes[2], Liz Brown[2], Dr Maurice Alley[3]

1. Ministry for Primary Industries
2. Department of Conservation
3. Massey University

Abstract

The black stilt (Himantopus novaezelandiae) or kaki, is an endemic New Zealand bird, and it is now critically endangered. Since the 1930's the population has declined reaching a 1981 low of 23 birds, while currently there are 77 breeding adults, they are limited in distribution to the MacKenzie Basin. The decline has been attributed to a variety of causes; predation, habitat loss, human activity and hybridisation with pied stilts. The Department of Conservation (DOC) has a captive breeding programme where the eggs are incubated and the chicks raised until 2-9 months when they are released to the wild. Records since 2001 show chicks suffering from a sometimes fatal neurological condition. Histological evidence show a myocarditis and encephalitis which could be of bacterial, viral or nutritional origin. So far viral and bacterial culture and PCR's have not conclusively identified a pathogen. We are using RNA Sequencing (RNA Seq) of an affected kaki brain to investigate if an infectious pathogen can be identified, and once identified suitable protective actions devised for the chicks.

POSTER 43

ZIPF'S LAW AND BEYOND: TOWARD NEW AND USEFUL 10,000FT VIEWS OF YOUR NGS SEQUENCE DATA

Mr Alan McCulloch[1]

1. Agresearch

Abstract

The frequencies of words in most languages have been found to follow a power law relationship relative to their rank in the frequency table, an observation popularised by George Zipf and known as Zipf's law. Power law behaviour has also previously been noted in the occurrence of various kinds of genomic properties including DNA words. We have noted power law or near power law behaviour in the occurrence of short DNA words in a wide variety of NGS sequencing contexts, including re-sequencing, amplicon sequencing and genotyping by sequencing, and have exploited this by incorporating a "zipfian analysis" of NGS datasets generated by our own lab and others into a number of pipelines. We show examples of how this interesting power-law behaviour can be used to quickly diagnose common problems such as incompletely trimmed sequence data, sample labelling error and unusual biases that can otherwise be difficult to see. We show an example of how Zipfian analysis also has a potential scientific application via hypothesis generation and confirmation concerning relative levels of biological repeat content. We briefly describe a broader "entropy-reduction" rationale and context for this kind of analysis to try to help understand why it works, which could provide a basis for the development of novel genomic data investigation methods. We refer readers interested in more details of this class of methods to our paper currently available in pre-print, and to our open-source python and R software libraries.

geneious

BRINGING LEADING EDGE **NGS** TECHNOLOGIES TO AUSTRALASIA

Distributors of:

- **Fragment Analyzer (nucleic acid quantity & quality)**
- **Denovix Spectrophotometer & Fluorometer**
- **Lexogen (RNA Sequencing)**
- **Swift Biosciences (NGS Library Prep)**

and more ……

dnature

• DETECT • DISCOVER •

www.dnature.co.nz

GeneTarget*Solutions*

www.genetargetsolutions.com.au

OUR SPONSORS & EXHIBITORS

SPONSORS

The organising committee for the 2016 AGTA/NZ NGS Conference would like to take the opportunity to thank the following organisations for their valuable contribution, and ask that all delegates to please consider these organisations for future business relationships.

GOLD SPONSOR

Satchel Sponsor

bio-strategy
delivering technology

Bio-Strategy

Bio-Strategy is an established locally-owned ISO 9001:2008 accredited distributor, supplying laboratory instruments, automated technology solutions and reagents from world famous brands. We serve the clinical diagnostic, life sciences and applied sectors throughout Australia and New Zealand. Our highly qualified and trained team offer the best solutions to our customers, backed by technical support from dedicated Applications Scientists and Service Engineers.

www.bio-strategy.com

SILVER SPONSOR

Name Badge Sponsor

illumina®

Illumina

Unlocking the power of the genome. Illumina is pleased to sponsor AGTA. Our sequencing and array technologies are fueling groundbreaking advancements in life science research, translational and consumer genomics, and molecular diagnostics. NGS is opening new doors in genomics. Please come to our booth to learn more about our latest product news, including the MiniSeq. The MiniSeq System supports a broad range of DNA and RNA applications with both predefined and custom panels.

www.illumina.com

SILVER SPONSOR

Barista Cart Sponsor

ThermoFisher
SCIENTIFIC

ThermoFisher Scientific

Thermo Fisher Scientific is the world leader in serving science. Our mission is to enable our customers to make the world healthier, cleaner and safer. From Applied BiosystemsTM gold standard Sanger sequencing to simple and scalable Ion Torrent™ next-generation sequencing, we have a platform to meet your needs. Achieve fast, accurate, and affordable results with optimized applications, high-quality consumables, and simplified analysis solutions.

www.thermofisher.com

SPONSORS

BRONZE SPONSOR

Day Refreshment Break Sponsor

Agilent Technologies

Demand Precision – Because finding the right answer is everything. Agilent Technologies, with our Genomics distributor, Global Science, is pleased to sponsor AGTA. Delivering on the promise of precision medicine takes getting results that bring clarity to the complex. We are dedicated to developing solutions for oncology, human and reproductive genetics and life sciences with precision that outperforms – so you find answers that make a difference. Please visit our booth to learn more about our latest products.

www.agilent.com | www.globalscience.co.nz

BRONZE SPONSOR

Day Refreshment Break Sponsor

dnature

dnature diagnostics & research is a New Zealand company specialising in PCR / real-time PCR and related areas. We offer a range of instruments, reagents and services to help you in the extraction, amplification and cloning of your DNA and RNA targets. Unlike other scientific companies, we use the machines and reagents that we distribute in our own laboratory so we can personally vouch for their efficacy. Featured at the conference will be the Advanced Analytical Fragment Analyzer and NGS library kits.

www.dnature.co.nz

BRONZE SPONSOR

Day Refreshment Break Sponsor

Gene Target Solutions

Gene Target Solutions will be showcasing: (1) The latest generation of NGS Library Preparation Technology from Swift Biosciences, allowing the end-user to develop high complexity libraries (exceptional coverage of AT- and GC- rich regions) from limiting amount of material (PCR-free from 10 ng gDNA) and (2) The Fragment Analyzer from Advanced Analytical Technologies Inc (award winning Technology currently used by Illumina to QC NGS preparations).

www.genetargetsolutions.com

SPONSORS

BRONZE SPONSOR

Day Refreshment Break Sponsor

New Zealand Genomics Limited

Customised genomics solutions for New Zealand scientistsNZGL provides an end-to-end genomics service for New Zealand's genomics researchers – from experimental design and data sequencing, through to bioinformatics support and data storage. The NZGL team provides a personalised service, working alongside researchers to offer customised solutions and on-going support as required. NZGL is a collaboration between The University of Auckland, Massey University and the University of Otago, supported by the New Zealand Government through the Ministry of Business, Innovation and Employment.

www.nzgenomics.co.nz

BRONZE SPONSOR

Early Career Researcher Award

Biomatters

First released in 2005, Geneious is the world's most cited molecular biology platform used by over 2500 universities and institutes, and commercial companies all over the world. Geneious takes you from data to discovery faster with an an easy-to-use interface, industry-leading tools, and intuitive data management - all within one customizable and extendable platform. With Geneious, any researcher working with sequencing data has immediate access to a wide-range of essential features and world-class support.

www.geneious.com

BRONZE SPONSOR

Poster Session

ESR

ESR is a Government owned research institute employing over 400 passionate scientists and support staff whose independent scientific advice and services help protect people and products in New Zealand, and around the world. Our world class knowledge, research and laboratory services lead, support and advise the decisions that safeguard people's health, protect our food-based economy, improve the safety of our freshwater and groundwater resources and provide the justice sector with expert forensic science.

www.esr.cri.nz

SPONSORS

SESSION SPONSOR

Beef + Lamb New Zealand (B+LNZ)

B+LNZ Genetics was established in 2014 to consolidate New Zealand's sheep and beef Genetics research and innovation into a single entity. We are a subsidiary of the farmer-owned industry organisation, Beef + Lamb New Zealand (B+LNZ) and are funded by sheep and beef levy payers and the Ministry of Business, Innovation and Employment. We are supported by additional investment from third parties, including meat processors, breed societies and commercial entities interested in sheep and beef genetics and genomics. B+LNZ Genetics is building on already significant genetic gains in New Zealand farming.

www.blnzgenetics.com

MEDIA PARTNER

Lab + Life Scientist

www.LifeScientist.com.au | www.labonline.com.au

TRADE EXHIBITORS

DON'T FORGET TO HAVE YOUR EXHIBITOR PASSPORT STAMPED TO GO INTO THE DRAW TO WIN SOME GREAT PRIZES!

BOOTH 1

bio-strategy
delivering technology

Bio-Strategy

Bio-Strategy is an established locally-owned ISO 9001:2008 accredited distributor, supplying laboratory instruments, automated technology solutions and reagents from world famous brands. We serve the clinical diagnostic, life sciences and applied sectors throughout Australia and New Zealand. Our highly qualified and trained team offer the best solutions to our customers, backed by technical support from dedicated Applications Scientists and Service Engineers.

www.bio-strategy.com

BOOTH 2

Humanizing Genomics
macrogen
OCEANIA

Macrogen Oceania

Macrogen is the current market leader in biotechnology services based in Korea, providing pivotal research tools to a global client base. The company provides high quality genomic sequencing, microarray, oligo synthesis, knock-out mouse, and bioinformatics services at competitive prices. Using only the most up to date technology available on the market, Macrogen is able to facilitate large scale evolutionary studies, whilst still taking care to provide smaller scale investigations, customised to the client's needs.

www.macrogen.com

BOOTH 3

QIAGEN®

Qiagen

QIAGEN, with its leading expertise in sample preparation technologies, has developed a complete portfolio of dedicated products to optimize your workflow – on any NGS instrument. Reliable nucleic acid purification from any sample material, high-quality, target-enriched samples and streamlined library preparation procedures, together with market-leading data analysis and interpretation tools, ensure focused and accurate results. Overcome bottlenecks in your NGS research and get the most from limited sequencing capacity! Our complete portfolio of NGS solutions has been designed to deliver the most accurate data possible and is fully compatible with other sequencing platforms.

https://www.qiagen.com

TRADE EXHIBITORS

BOOTH 4

Partek®
turning data into discovery®

Partek

Partek provides software for next generation sequencing, microarray, and qPCR data analysis and visualization that empowers biologists to find biological meaning within their data. With an intuitive interface designed for researchers, Partek's complete solution takes raw data to statistical and pathway analysis that allows true multi omics integration.

www.partek.com

BOOTH 6

Ramaciotti Centre for Genomics

Ramaciotti Centre for Genomics

Ramaciotti Centre for Genomics The Ramaciotti Centre for Genomics, located at UNSW, is a not-for-profit provider of genomic services. We enable access to state-of-the-art, cutting edge technologies at affordable prices and deliver data of the highest quality. Our technology suite includes long- and short-read next-generation sequencing, single-cell analysis, microarrays and Sanger sequencing. Our professional team of scientists has many years of experience and provides researchers with a personalized service from design through to downstream analysis.

www.ramaciotti.unsw.edu.au

BOOTH 7

KINGHORN CENTRE FOR CLINICAL GENOMICS | GARVAN INSTITUTE OF MEDICAL RESEARCH

Kinghorn Centre for Clinical Genomics

The Kinghorn Centre for Clinical Genomics (KCCG) is a purpose-built facility for the provision of next-generation sequencing (NGS) and data analysis. It features multiple sequencing platforms, including the state-of-the-art Illumina HiSeq X Ten, the latest laboratory robotics and high performance computing architecture. The laboratory operates with dual purpose; offering NATA approved whole genome sequencing (WGS) for the provision of clinical diagnosis, while also providing best-in-class WGS and other bespoke sequencing formats to the wider research community.

www.garvan.org.au/kccg

TRADE EXHIBITORS

BOOTH 8

Medi'Ray

are suppliers of innovative, leading edge biotechnology products to the Diagnostic, Molecular Biology and Cell Biology market. We offer a variety of NGS solutions from Clontech and Lucigen:

- DNA-Seq AmpFREE Low DNA Library Kit, ChiP Seq Library Prep Kits and kits which offer the longest mate pair reads.
- RNA-Seq products for expression analysis of the entire transcriptome, with high sensitivity from all types of low-input RNA, including single cell and compromised samples.

www.mediray.co.nz

BOOTH 9

QCIP/QFAB Bioinformatics

QFAB Bioinformatics provides rapid, flexible and customized bioinformatics and biostatistics services to life science and clinical researchers. Working closely with researchers, we apply data management, integration, analysis and visualisation to unlock the full value of large-scale biological and clinical datasets. We pride ourselves in delivering robust, high quality and state-of-the-art results. To empower researchers in mastering their data generation and analysis, we also offer a wide range of courses and workshops, covering statistics, computing and bioinformatics.

www.qfab.org

BOOTH 10

BGI

BGI was founded in 1999 with the vision of using genomics to benefit mankind and has since become the largest genomic organization in the world. With a focus on research and applications in the healthcare, agriculture, conservation, and environmental fields, BGI has a proven track record of innovative, high profile research, which has generated over 1,847 publications, many in top-tier journals such as Nature and Science.

http://www.genomics.cn/en/index

TRADE EXHIBITORS

BOOTH 11

Tacan Australia

Tecan is a leading global provider of automated laboratory instruments and solutions. Our systems, components and support help people working in clinical diagnostics, basic, translational research and drug discovery bring their science to life. We develop, produce, market and support automated workflow solutions that empower laboratories to achieve more. Our expandable, reconfigurable solutions cover all your genomic automation needs including DNA extraction, Normalization, PCR set up, DNA shearing, and NGS sample preparation.

www.tecan.com

BOOTH 12

Dnature/Gene Target Solutions

dnature diagnostics & research is a New Zealand company specialising in PCR / real-time PCR and related areas. We offer a range of instruments, reagents and services to help you in the extraction, amplification and cloning of your DNA and RNA targets. Unlike other scientific companies, we use the machines and reagents that we distribute in our own laboratory so we can personally vouch for their efficacy. Featured at the conference will be the Advanced Analytical Fragment Analyzer and NGS library kits.

www.dnature.co.nz

Gene Target Solutions will be showcasing: (1) The latest generation of NGS Library Preparation Technology from Swift Biosciences, allowing the end-user to develop high complexity libraries (exceptional coverage of AT- and GC- rich regions) from limiting amount of material (PCR-free from 10 ng gDNA) and (2) The Fragment Analyzer from Advanced Analytical Technologies Inc (award winning Technology currently used by Illumina to QC NGS preparations).

www.genetargetsolutions.com

TRADE EXHIBITORS

BOOTH 13

Millennium Science New Zealand

Millennium Science New Zealand is a specialist distribution company established in 1999 specifically to source, supply and support high technology solutions to the Australasian Life Science market. We offer scientists long-range sequencing technologies, single-cell genomics technologies and other enabling technologies to advanced genomics research. The manufacturers we represent include Pacific Biosciences, Affymetrix, 10X Genomics, Sage Science, Roche NimbleGen and more.

www.mscience.co.nz

BOOTH 14

AGRF

Australian Genome Research Facility (AGRF) is Australia's largest provider of genomics services and solutions. AGRF has laboratories in Brisbane, Melbourne, Adelaide, Sydney and Perth, each providing a gateway to a national network of state-of-the-art facilities, technology and expertise. From microbes to plants, animals and humans, AGRF provides a full range of services across the entire biological spectrum. We provide services to academia and industry with clients from bioscience, environmental science, biomedicine and agricultural biotechnology.

www.agrf.org.au

BOOTH 15

Ngaio Diagnostics

Ngaio Diagnostics Ltd. is a privately owned supply company based in Nelson, New Zealand. We represent many world leading developers and manufacturers of high quality analytical technologies, and specialise in the distribution of these technologies and provision of associated technical support to various laboratories throughout New Zealand. Zymo Research, Intron Biotechnology and Bioo Scientific are some of the companies that we represent, which manufacture high quality molecular biology research products for a variety of applications.

www.ngaio.co.nz

TRADE EXHIBITORS

BOOTH 16

Illumina

Unlocking the power of the genome Illumina is pleased to sponsor AGTA. Our sequencing and array technologies are fueling groundbreaking advancements in life science research, translational and consumer genomics, and molecular diagnostics. NGS is opening new doors in genomics. Please come to our booth to learn more about our latest product news, including the MiniSeq. The MiniSeq System supports a broad range of DNA and RNA applications with both predefined and custom panels.

www.illumina.com

BOOTH 17

Roche Diagnostics

At Roche, we believe in doing what should be done, rather than what has been done. So we follow where the science leads. We focus on today and look beyond tomorrow to tackle the world's most formidable healthcare challenges, by concentrating on medical needs, considering the bigger picture of diagnostic testing and treatments and constantly innovating through valuable collaborations with our partners globally.

www.roche.co.nz

BOOTH 18

Perkin Elmer Inc

PerkinElmer is a global leader focused on improving the health and safety of people and the environment. Our innovative detection, imaging, informatics and service capabilities, combined with deep market knowledge and expertise, help customers gain greater insights into their science to better protect our environment, our food supply and the health of our families. We provide a truly unique experience for our customers as we help them improve human and environmental health. Together we are making a difference for people, for the world, for the better.

www.perkinelmer.com/au

TRADE EXHIBITORS

BOOTH 19

Agilent Technologies

Demand Precision – Because finding the right answer is everything. Agilent Technologies, with our Genomics distributor, Global Science, is pleased to sponsor AGTA. Delivering on the promise of precision medicine takes getting results that bring clarity to the complex. We are dedicated to developing solutions for oncology, human and reproductive genetics and life sciences with precision that outperforms – so you find answers that make a difference. Please visit our booth to learn more about our latest products.

www.agilent.com | www.globalscience.co.nz

BOOTH 20

Genesearch

Genesearch is the Australian and New Zealand distributor for New England Biolabs (NEB) and home of the eFreezer. NEB produces the NEBNext range of next generation sequencing reagents, including library prep kits for Illumina and Ion Torrent instruments, and accessory kits for library quant, microbiome DNA enrichment, rRNA depletion, FFPE DNA purification, and more!

www.genesearch.com.au (Australia)
www.genesearch.co.nz (New Zealand)

BOOTH 21

Biomatters

First released in 2005, Geneious is the world's most cited molecular biology platform used by over 2500 universities and institutes, and commercial companies all over the world. Geneious takes you from data to discovery faster with an an easy-to-use interface, industry-leading tools, and intuitive data management - all within one customizable and extendable platform. With Geneious, any researcher working with sequencing data has immediate access to a wide-range of essential features and world-class support.

www.geneious.com

TRADE EXHIBITORS

BOOTH 22

Agena Bioscience

Agena Bioscience is a San Diego, CA based life sciences and clinical diagnostics company that recently acquired the Bioscience business of Sequenom, Inc. and is now offering the MassARRAY® System. The system is a highly sensitive, quantitative method for nucleic acid detection via mass spectrometry for high-throughput genotyping and mutation profiling for cancer and other disease research, companion diagnostics, pharmacogenomics, molecular blood group typing, epigenetics, clinical genetics, agrigenomics, and molecular sample identification for bio-banking.

www.agenabio.com

BOOTH 23

New Zealand Genomics Ltd

Customised genomics solutions for New Zealand scientists. NZGL provides an end-to-end genomics service for New Zealand's genomics researchers – from experimental design and data sequencing, through to bioinformatics support and data storage. The NZGL team provides a personalised service, working alongside researchers to offer customised solutions and on-going support as required. NZGL is a collaboration between The University of Auckland, Massey University and the University of Otago, supported by the New Zealand Government through the Ministry of Business, Innovation and Employment.

www.nzgenomics.co.nz

BOOTH 24

ThermoFisher Scientific

Thermo Fisher Scientific is the world leader in serving science. Our mission is to enable our customers to make the world healthier, cleaner and safer. From Applied BiosystemsTM gold standard Sanger sequencing to simple and scalable Ion Torrent™ next-generation sequencing, we have a platform to meet your needs. Achieve fast, accurate, and affordable results with optimized applications, high-quality consumables, and simplified analysis solutions.

www.thermofisher.com

TRADE EXHIBITORS

BOOTH 26

Zygem NZ

Nucleic acid extraction made easy! Introducing the new Automated DNA extraction PDQeX 2400 from ZyGEM. The PDQeX is a rapid, robust DNA extraction device that can achieve PCR-ready DNA in five to fifteen minutes, depending on the sample type. The PDQeX2400 can process 24 samples simultaneously and can be used with a variety of substrates like buccal swabs, blood, semen, plant leaves with more substrates to follow.

www.zygem.com

BOOTH 27

Custom Science

Custom Science NZ was established in April 2011 to deliver best in class products to the New Zealand Scientific community. We have since established Custom Science Australia, with an office in Melbourne, Australia and introduced a range of products for nucleic acid extraction and purification, PCR, qPCR, transfection, protein analysis and next generation sequencing. Please feel free to visit us at our booth.

www.customscience.co.nz
www.customscience.com.au

2016 DELEGATE LIST

NAME	ORGINISATION	STATE/ COUNTRY
Antonio Ahn	University of Otago	NZL
Mila Al-Halbouni	Global Science	NZL
Abdulmonem Alsaleh	University of Otago	NZL
Nicola Angel	University of Queensland	QLD
Stuart Archer	Monash Bioinformatics Platform	VIC
Kelly Atkinson	University of Auckland	NZL
Brant Bassam	Bio-Strategy	VIC
Denis Bauer	CSIRO	NSW
Traude Beilharz	Monash	VIC
Dan Belluoccio	Agilent Technologies	VIC
Graeme Benny	ESR	NZL
Miles Benton	Queensland University of Technology	NZL
Tony Bhalla	PerkinElmer	NSW
Louise Bicknell	University of Otago	NZL
Barbara Binney	Ministry for Primary Industries	NZL
Mik Black	University of Otago	NZL
Kristine Boxen	University of Auckland	NZL
Rachel Boyle	Auckland University of Technology	NZL
Luke Boyle	University of Auckland	NZL
Dana Bradford	CSIRO	QLD
Aubrey Brand	Bio-Strategy	NZL
Lara Brian	Plant & Food Research	NZL
Sam Buckberry	University of Western Australia	WA
Alayna Burrett	Millennium Science	NZL
Suzanne Butcher	University of Melbourne	VIC
Michelle Cagney	Bio-Strategy	NZL
Piero Carninci	RIKEN	JPN
Jaime Iván Castillo Silva	University of Warsaw	MEX
Natalie Castrechini	QIAGEN	VIC
Alistair Chalk	St Vincent's Institute	VIC
Amanda Chamberlain	Department of Economic Development, Jobs, Transport and Resources	VIC
Jac Charlesworth	Menzies Institute for Medical Research	TAS
Tanya Hf Chua	Partek	SGP
Jane Clapman	ESR	NZL
Shannon Clarke	AgResearch	NZL

2016 DELEGATE LIST

NAME	ORGINISATION	STATE/ COUNTRY
Nicole Cloonan	QIMR Berghofer Medical Research Institute	QLD
Catherine Collins	University of Otago	NZL
Joe Copty	Kinghorn Centre for Clinical Genomics	NSW
Susan Corley	University of New South Wales	NSW
Angela Cornelius	ESR	NZL
Claudia Carolina Correa Ospina	The Ramaciotti Centre for Genomics	NSW
Simone Cree	University of Otago	NZL
Mallory Crookenden	DairyNZ	NZL
Tamsyn Crowley	Deakin University	VIC
John Davis	Thermo Fisher Scientific	QLD
Robert Day	University of Otago	NZL
Elena Denisenko	Massey University	NZL
Marcel Dinger	Garvan Institute of Medical Research	NSW
Steven Dodd	Thermo Fisher Scientific	NZL
Jenny Draper	ESR	NZL
Nicola Dryden	Thermo Fisher Scientific	NZL
Nicolas Dussex	University of Otago	NZL
David Eccles	Gringene Bioinformatics	NZL
Dave Edwards	University of Western Australia	WA
Bradford Elmore	Tecan Australia	VIC
Yvette Emmanuel	Agena Bioscience	QLD
Vicky Fan	University of Auckland	NZL
Rachel Fleming	ESR	NZL
Mikhail Fokin	University of Auckland	NZL
Saskia Freytag	Walter & Eliza Hall Institute	VIC
Austen Ganley	University of Auckland	NZL
Juan Carlos Garcia	Massey University	NZL
Alexandra Garnham	Walter & Eliza Hall Institute	VIC
Velimir Gayevskiy	Garvan Institute of Medical Research	NSW
Neil Gemmell	University of Otago	NZL
Chunyu Geng	BGI	CHN
Edna Gias	Ministry for Primary Industries	NZL
Greg Gibson	Georgia Institute of Technology	USA
Brian Gloss	Garvan Institute of Medical Research	NSW

2016 DELEGATE LIST

NAME	ORGINISATION	STATE/ COUNTRY
Dominique Gorse	QFAB Bioinformatics	QLD
Natalie Graham	Scion Research	NZL
Sol Green	Bio-Rad	NZL
Richard Green	University of California, Santa Cruz	USA
Ruth Griffin	MPI	NZL
Catherine Grueber	University of Sydney	NSW
Chris Harris	University of Otago	NZL
Anthony Hawkins	MCRI	VIC
Roger Hellens	Queensland University of Technology	QLD
Natasha Henden	LabPLUS	NZL
Syrie Hermans	Auckland University of Technology	NZL
Karl Herron	Agena Bioscience	QLD
Elena Hilario	Plant & Food Research	NZL
Andrew Hill	La Trobe University	VIC
Daniel Ho	Massey University	NZL
Wellcome Ho	MPI	NZL
Roger Huerlimann	James Cook University	QLD
Jung-ah Hwang	National Cancer Center	KOR
Tracey Immanuel	MPI	NZL
Violeta Ivanov	Australian Clinical Laboratories	VIC
Cameron Jack	ANU Bioinformatics Consultancy	ACT
Jeanne Jacobs	AgResearch	NZL
Aaron Jeffs	University of Otago	NZL
Dean Jerry	James Cook University	QLD
Rubina Jibran	Massey University	NZL
Elena Jiyoon	Macrogen Inc	KOR
Margaret Jordan	James Cook University	QLD
Clair Joyce	Tecan Australia	VIC
Irfahan Kassam	The Queensland Brain Institute	QLD
Jaspreet Kaur	Punjabi University	IND
Matt Kearse	Biomatters	NZL
Martin Kennedy	University of Otago	NZL
Subuhi Khan	Ministry for Primary Industries	NZL
Jee In Kim	Keimyung University School of Medicine	KOR
Irene Kourtis	AGRF	VIC

2016 DELEGATE LIST

NAME	ORGINISATION	STATE/ COUNTRY
Irene Kourtis	AGRF	VIC
Jason Koval	The Ramaciotti Centre for Genomics	NSW
Joice Kuroiwa-Trzmielina	Garvan Institute of Medical Research	NSW
Sebastian Kurscheid	Australian National University	ACT
Steve Lasky	Advanced Analytical	USA
Edmund Lau	Victorian Life Science & Computation Initiative	VIC
Rebecca Laurie	NZGL	NZL
Tim Lawrence	University of Auckland	NZL
Aaron Lewis	Australian Genome Research Facility	VIC
Dongmei Li	Ministry for Primary Industries	NZL
Lia Liefting	Ministry for Primary Industries	NZL
Ruby Lin	Asbestos Diseases Research Institute	NSW
Nooriyah Lohani	University of Auckland	NZL
Donia Macartney-Coxson	ESR	NZL
John Mackay	dnature	NZL
Paul Maclean	Agresearch	NZL
Elspeth MacRae	Scion Research	NZL
Vikki Marshall	University of Melbourne	VIC
Christopher Mason	Cornell University	USA
Elizabeth Mason	University of Melbourne	VIC
Lisa Matisoo-Smith	University of Otago	NZL
Abishek Mauralidhar	Zygem NZ Ltd	NZL
Damara McAndrew	Genesearch	QLD
Alan McCulloch	Agresearch	NZL
Ken McGrath	AGRF	QLD
Gai McMichael	AGRF	VIC
Les McNoe	Otago Genomics & Bioinformatics Facility	NZL
Kathryn McRae	Agresearch	NZL
Lance Miller	Wake Forest School of Medicine	USA
Cath Moore	QIAGEN	VIC
Shalima Nair	Garvan Institute of Medical Research	NSW
Mani Narayanaswamy	Bio-Rad	NZL
Felicity Newell	QIMR Berghofer Medical Research Institute	QLD
Helen Norrie	Medi' Ray	NZL
Aidan O'brien	CSIRO	NSW

2016 DELEGATE LIST

NAME	ORGINISATION	STATE/ COUNTRY
Thomas Ohnesorg	Murdoch Childrens Research Institute	VIC
Sal Palagummi	Ngaio Diagnostics LTD	NZL
Kwon Moo Park	Kyungpook National University	KOR
Duckchul Park	Landcare Research	NZL
Carolina Penaloza	Roslin Institute	GBR
Zoila Perez	Ministry for Primary Industries	NZL
Elizabeth Permina	University of Otago	NZL
Belinda Phipson	Murdoch Childrens Research Institute	VIC
Sarah Pilkington	Plant & Food Research	NZL
Katherine Pillman	University of South Australia	SA
Gretchen Poortinga	Peter Maccallum Cancer Centre	VIC
Joseph Powell	University of Queensland	QLD
Cris Print	University of Auckland	NZL
Debbie Prosser	LabPLUS	NZL
Thom Quinn	Deakin University	USA
Christy Rand	University of Otago	NZL
Edwardo Reynolds	University of Auckland	NZL
Anna Santure	University of Auckland	NZL
William Schierding	Bioinformatics Institute	NZL
Sebastian Schmeier	Massey University	NZL
Vicky Schneider	EMBL Australia Bioinformatics Resource	VIC
Andreas Schreiber	ACRF South Australian Cancer Genomics Facility	SA
Beth Shapiro	University of California, Santa Cruz	USA
Kirby Siemering	Australian Genome Research Facility	VIC
Beth Signal	Garvan Institute of Medical Research	NSW
Cas Simons	University of Queensland	QLD
Siddharth Singh	Millennium Science	VIC
Priyanka Sinha	University of Otago	NZL
Peter Smith	Agresearch/University of Otago	NZL
Jo Stanton	University of Otago	NZL
Aaron Statham	Kinghorn Centre for Clinical Genomics	NSW
Peter Stockwell	University of Otago	NZL
Andrew Stone	Kinghorn Centre for Clinical Genomics	NSW
Andrew Szentirmay	Gene Target Solutions	NSW
Jiang Tao	Garvan Institute of Medical Research	NSW

2016 DELEGATE LIST

NAME	ORGINISATION	STATE/ COUNTRY
Daniel Thomson	Garvan Institute of Medical Research	NSW
Anthony Thrush	NZGL	NZL
Peri Tobias	University of Sydney	NSW
Richard Tothill	Peter Maccallum Cancer Centre	VIC
Peter Tsai	University of Auckland	NZL
Mark Van Der Hoek	SAHMRI	SA
Christy Vander Jagt	Agriculture Victoria	VIC
Stella Veerakone	Ministry for Primary Industries	NZL
Tina Visnvoska	Massey University	NZL
Brett Wagner	Auckland University of Technology	NZL
Logan Walker	University of Otago	NZL
Yunjie Wang	BGI	CHN
Christine Wells	University of Melbourne	VIC
Trystan Whang	Macrogen Inc	NSW
Daniel White	Landcare Research	NZL
Phillip Wilcox	University of Otago	NZL
Ingrid Wise	Federation University Australia	VIC
Joseph Youssef	Bio-Strategy	NZL
Zitong Yu	Murdoch University	WA
Monika Zavodna	Otago Genomics & Bioinformatics Facility	NZL
Huaijun Zhou	University of California, Davis	USA

SAVE THE DATE

AGTA
Australasian Genomic Technologies Association

2017 ANNUAL CONFERENCE

HOBART October 29th - November 1st

www.agtagenomics.org.au | www.agtaconference.org

SPONSORS & EXHIBITORS