AMATA 2013

13-16 October 2013
Marriott Surfers Paradise
Queensland Australia

HANDBOOK & PROGRAM

Sample to insight

Ready for advanced genetic analysis

The QIAGEN GeneReader™ sample to insight NGS workflow

- **Automated and standardized** — streamlined from sample prep to data analysis
- **Flexible and scalable** — multiple, independent sequencing runs processed in parallel
- **Disease-focused content** — target enrichment using comprehensive gene panels
- **Deep biological insight** — integrated bioinformatics and biology-based data interpretation

Visit www.qiagen.com/goto/NGS to find out more about the QIAGEN GeneReader sample to insight next-generation sequencing ecosystem.

Australia ■ Orders: 1800 243 800 ■ Fax: 03 9840 9888 ■ Technical: 1800 243 066

QIAGEN®

Sample & Assay Technologies

Contents

Committees

AMATA
Executive Committee

Assoc Prof Daniel Catchpoole (President)
TUMOUR BANK, THE CHILDREN'S HOSPITAL AT WESTMEAD, SYDNEY

Prof Erik Thompson
ST VINCENT'S INSTITUTE, MELBOURNE

Mark van der Hoek
ADELAIDE MICROARRAY CENTRE, HANSON INSTITUTE, ADELAIDE

Dr Ruby C Y Lin
RAMACIOTTI CENTRE, UNIVERSITY OF NEW SOUTH WALES, SYDNEY

Vikki Marshall
VICE-PRESIDENT & AMATA 2010 CONVENOR

Dr Carsten Kulheim
ACRF BIOMOLECULAR RESOURCE FACILITY, ANU, CANBERRA

Assoc Prof Christine Wells
UNIVERSITY OF QUEENSLAND NODE, BRISBANE

Dr Lutz Krause
QIMR, HERSTON, BRISBANE

Dr Richard Tothill
PETER MACCALLUM CANCER CENTRE, MELBOURNE

Dr Alicia Oshlack
MURDOCH CHILDRENS RESEARCH INSTITUTE, MELBOURNE

Dr Jac Charlesworth
MURDOCH CHILDRENS RESEARCH INSTITUTE, MELBOURNE

Prof Ryan Lister
THE UNIVERSITY OF WESTERN AUSTRALIA, PERTH

Prof Greg Goodall
CENTRE FOR CANCER BIOLOGY, ADELAIDE

Prof Ian Paulsen
DPT. OF CHEMISTRY & BIOMOLECULAR SCIENCES, MACQUARIE UNIVERSITY

Liam Williams
CTR FOR GENOMICS, PROTEOMICS & METABOLOMICS, AUCKLAND UNIVERSITY

AMATA2013
Organising Committee

Assoc Prof Christine Wells (Conference Convenor)
(AIBN), THE UNIVERSITY OF QUEENSLAND, BRISBANE, AND INSTITUTE OF INFECTION, IMMUNITY AND INFLAMMATION, UNIVERSITY OF GLASGOW, SCOTLAND

Dr Denis Bauer
CSIRO, SYDNEY

Dr Nicole Cloonan
GENOMIC BIOLOGY LABORATORY, QUEENSLAND INSTITUTE OF MEDICAL RESEARCH (QIMR), BRISBANE

Dr Carsten Kulheim
ACRF BIOMOLECULAR RESOURCE FACILITY, ANU, CANBERRA

Dr Andreas Scherer
AUSTRALIAN GENOME RESEARCH FACILITY (AGRF), AUSTRALIA/ SPHEROMICS, KONTIOLAHTI, FINLAND

Dr David Lovell
AUSTRALIAN BIOINFORMATICS NETWORK/CSIRO, CANBERRA

Dr Lutz Krause
GENOMIC BIOLOGY LABORATORY QUEENSLAND INSTITUTE OF MEDICAL RESEARCH (QIMR)

Dr Mark Crowe
QFAB, THE UNIVERSITY OF QUEENSLAND, BRISBANE

Conference Managers

113 Harrington Street, Hobart TAS 7000
Phone: 03 6234 7844 Fax: 03 6234 5958
Mobile: 0408 625 033
Email: brigitte@leishman-associates.com.au
www.leishman-associates.com.au

Leishman
it's all about people

Welcome

Dear Delegates and Invited Guests,

Welcome to the 13th Annual meeting for the Australasian Genomics and Associated Technologies Association, and to the beautiful Queensland Gold Coast. This meeting marks the transition from AMATA to AGTA, which in turn reflects the changing face of genomics in Australia. AGTA members include those who generate data, and those who analyse data, technology innovators and biology enthusiasts. We recognise that the boundaries between wet and dry labs, research and clinic, industry and academia are becoming increasingly merged. This year is an exciting one for our society, and sees the launch of new AGTA initiatives, including funding schemes and training. We trust that AGTA meetings will continue to facilitate networking and collaboration across the 'omics sectors. Whether you are a long-time member, or a newcomer to the society, welcome!

The conference organising committee would particularly like to welcome our international plenary speakers to Australia, and thank them for their contributions throughout the meeting. They are headlining an exciting program. We are very appreciative of the outstanding line-up of keynote speakers and national chairs. The AMATA2013 committee is very proud to be showcasing early career researchers, as well as our well-established innovators, in the line-up of keynote speakers, and in the talks and posters drawn from abstracts. Australia continues to capitalise on the rapid developments in integrated technologies to tackle a diverse array of medical, agricultural and environmental questions. We have focused on three main themes – clinical sequencing, data integration and the emergence of single cell technologies. We would like to thank each one of you for contributing to this exciting line-up, and hope that you enjoy the program ahead.

We are delighted to be partnered with our Platinum sponsor, QIAGEN, and their choice of AMATA2013 to showcase their new Genomic platforms. QIAGEN's sponsorship has provided many of the networking opportunities at the conference, and we thank them for this. Thank you also to our silver sponsors Illumina, Roche and Millennium Science. This year QIAGEN and our silver sponsors participate in a moderated industry forum as part of the program – this is your opportunity to look beyond the tradehall and ask questions about the challenges of applying new technologies, merging platforms or working up new methods in your own laboratories. We recognise the importance that technology companies play in adoption of genomics in Australia, and hope you'll take up the opportunities to interact with these lead companies in a new way.

The conference logo was designed by Denis Bauer, who has also played a critical role in the organisation of the conference program. It includes the black-breasted button quail, *Turnix melanogaster*, which is only found in SE QLD and northern NSW. Although it was on the 'red-list' for endangered species for many years due to fragmentation of its population, the quail population is regaining strength in numbers. Its use in the AMATA2013 conference is a timely reminder of the value in nurturing the 'rare bird' and the strength in coming together as a community.

It's my pleasure to welcome you to AMATA2013, on behalf of the AGTA executive and the AMATA2013 conference organising committee.

Christine Wells

Convenor AMATA2013

Conference Information

Registration Desk

On Sunday afternoon the Registration Desk will be located in the Lobby Foyer, ground floor of the Surfers Paradise Marriott Resort.

From Monday through to Wednesday the Registration Desk will be located in the Garden Terrace Gallery on level 1 of the Surfers Paradise Marriott Resort.

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

Registration Desk Opening Times

Sunday 13 October 2013	**2.00pm – 6.00pm**
Monday 14 October 2013	**7.30am – 5.30pm**
Tuesday 15 October 2013	**7.30am – 5.30pm**
Wednesday 16 October 2013	**7.30am – 5.30pm**

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.

ATM's

An ATM is located in the Lobby of the Surfers Paradise Marriott Resort.

Conference 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.

Conference Proceedings

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

Dress

Dress throughout the day is smart casual or informal business.

Emergency Medical Care

For any medical emergency please telephone 000. The staff at your hotel will have information if you require contact details for a doctor, dentist or other health professional.

Internet

Complimentary wireless internet will be available on the events level of the Surfers Paradise Marriott Resort for the duration of the conference. Please use the Marriott@public wiFi connection, you will need to refresh the connection every 24 hours. No user name/ password is required. If you have trouble connecting please see the staff at the Registration Desk.

Complimentary wiFi is available in all guest rooms at the Surfers Paradise Marriott Resort. Please note that guests will be required to join Marriott Rewards, and use their surname and room number for connection. Please see the staff at the Resort Reception for further information.

Mobile Phones

As a courtesy to other delegates, please ensure that all mobile phones are turned off or in a silent mode during all sessions and social functions.

Parents Room

A private room will be available at the conference venue for nursing mothers and others with sensitive personal health needs. Please note that this room will not be staffed. Attendees are not permitted to utilise this room for babysitting services. Please see the Leishman Associates staff at the registration desk for further information.

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 and conference organiser.

Smoking

Surfers Paradise Marriott Resort and all other conference venues are non-smoking. Guests are requested to restrict smoking to designated areas.

Speakers and Speaker's Preparation Room

All speakers should present themselves to the Speaker's Preparation Room, located in Terrace Room Two 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 Registration Desk.

A technician will be present in the Speaker's Preparation Room during registration hours. There will be facility to test and modify your presentation as required.

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 3 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 Garden Terrace of the Surfers Paradise Marriott Resort for the duration of the conference. There will be a poster session on Monday 14 October 2013 from 2.00pm to 3.30pm.

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.

Twitter

The AMATA2013 organising committee encourages all delegates to tweet about the conference. Please use #AMATA. If any speakers do not wish delegates to tweet during their presentation please advise the chairperson prior to your session commencing.

Disclaimer

The 2013 AMATA 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.

Pre-Conference Workshop

AMATA Early Career Researcher Workshop

Dates
Saturday 12 and Sunday 13 Oct 2013

Location
Parklands Drive, Griffith University,
Gold Coast Campus, Science,
Engineering & Architecture Building

Room
G39_1.17 Lecture Theatre

Workshop Fee
$200.00 (inclusive of Sat and Sun)

Transport
Return bus transfers will be provided on Saturday and Sunday from the Surfers Paradise Marriott Resort to Griffith University. A bus will depart from the main entrance of the Surfers Paradise Marriott Resort at 8.20am sharp. **Delegates are requested to be on time for the departure.**

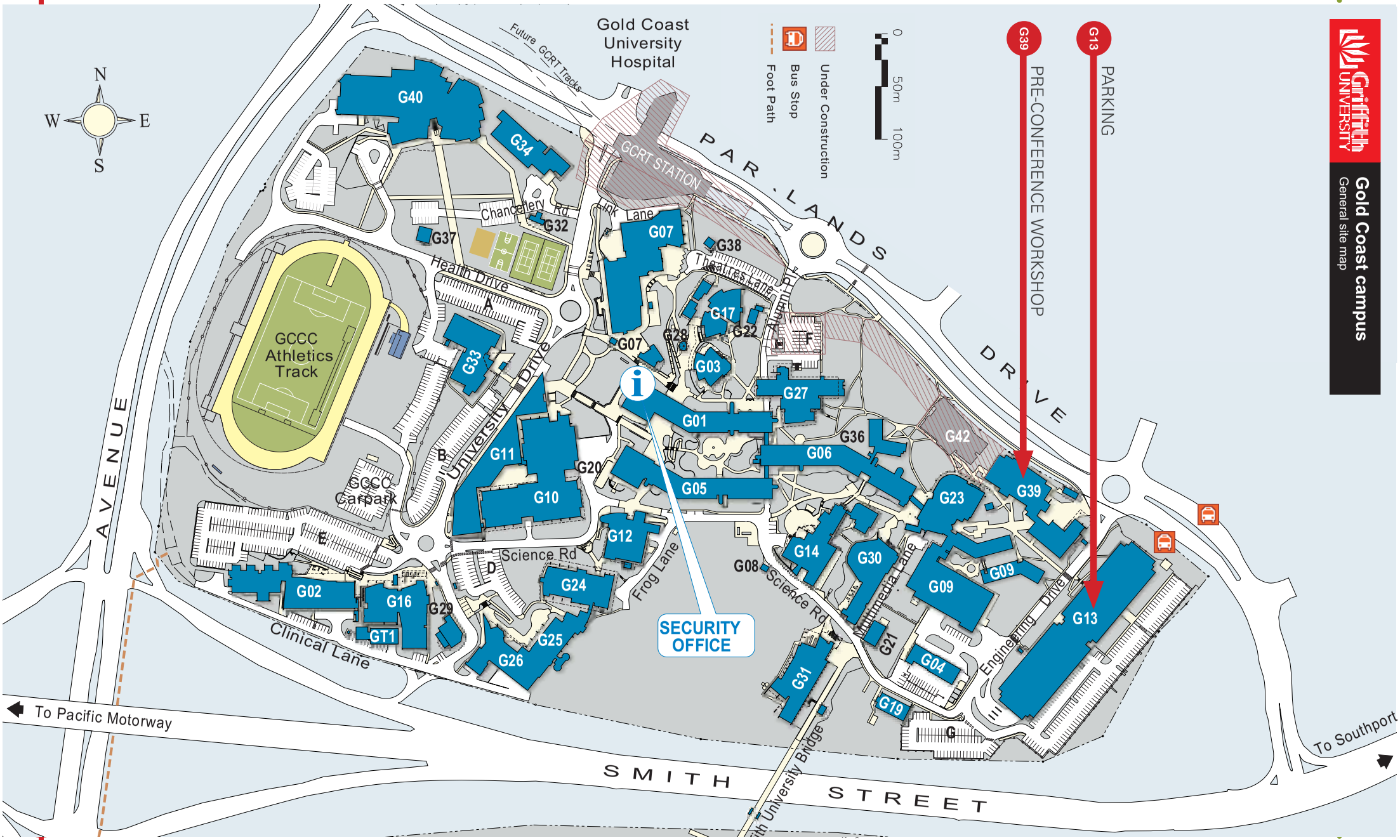
Parking
For those wanting to make their own way to and from the pre-conference workshop, free parking is available in Car Park G. Please refer to the below map for the location.

The Pre-Conference Workshop is not included in any registration type. Registration can be purchased at $200.00 per person.

The AMATA early career researcher workshop will offer seminars and hands on training in genomics, bioinformatics and career development skills, delivered by ten national and international genomics and bioinformatics experts.

Day one opens with sessions covering best-practice in designing, carrying out, and analysing genomics experiments. Participants will then spend the rest of the day in hands-on bioinformatics analysis practical sessions using the user-friendly Galaxy bioinformatics workflow platform. These will introduce techniques for RNA-Seq analysis, followed by options of either an introduction to bacterial genome assembly or more advanced RNA-Seq work.

The second (half) day will be devoted to career development advice. It will start with an open discussion and networking session, at which participants will be encouraged to raise technical and professional issues for input from their peers. This will be followed by presentations and question and answer sessions on how to manage projects and collaborations, before the workshop concludes with a final session about ways to improve a CV and increase the success rate of applications for jobs, grants and fellowships.


Gold Coast campus
General site map
Griffith UNIVERSITY
Gold Coast University Hospital
Future GCRT Tracks
GCRT STATION
PAR-LANDS DRIVE
PRE-CONFERENCE WORKSHOP
PARKING
G39
G13
N
W
E
S
0
50m
100m
Under Construction
Bus Stop
Foot Path
G40
G34
Chancellery Rd
G32
G37
Health Drive
Link Lane
G07
G38
Theatres Lane
Alumni Pl
G17
G22
G28
G07
G03
G27
G01
G36
G42
G06
G05
G23
G39
G11
G10
G20
GCCC Athletics Track
GCCC Carpark
University Drive
G12
Science Rd
G14
G30
G09
G09
G08
Science Rd
Multimedia Lane
Engineering Drive
G13
G02
G16
G29
GT1
G24
G25
G26
Frog Lane
SECURITY OFFICE
Clinical Lane
G21
G04
G31
G19
G
AVENUE
To Pacific Motorway
SMITH STREET
To Southport

Pre-Conference Program

SATURDAY 12 OCTOBER

GRIFFITH UNIVERSITY, GOLD COAST CAMPUS, G39_1.17 LECTURE THEATRE

0820	BUS DEPARTS SURFERS PARADISE MARRIOTT RESORT **Delegates are requested to be on time for the departure.**	
0830	REGISTRATION OPEN – ARRIVAL TEA & COFFEE	
0845	COACH ARRIVES GRIFFITH UNIVERSITY	
0900	INTRODUCTION TO NEXT-GENERATION SEQUENCING TECHNOLOGIES David Miller Queensland Centre for Medical Genomics, The University of Queensland	
0930	EXPRESSION ANALYSIS STANDARDISATION AND QUALITY CONTROL Prof Leming Shi, Fudan University, Shanghi	
1000	EXPERIMENTAL DESIGN AND EXPRESSION ANALYSIS Dr Kim-Ahn Le Cao, QFAB, The University of Queensland	
1030	MORNING REFRESHMENTS	
1100	BIOINFORMATICS WORKSHOPS – INTRODUCTION TO GALAXY AND BASIC RNA-SEQ ANALYSIS Dr Mark Crowe, QFAB Dr Annette McGrath, CSIRO	
1300	LUNCH	
1400	PARALLEL BIOINFORMATICS WORKSHOP ADVANCED RNA-SEQ Dr Mark Crowe, QFAB	PARALLEL BIOINFORMATICS WORKSHOP DE NOVO GENOME ASSEMBLY Dr Annette McGrath, CSIRO
1530	AFTERNOON REFRESHMENTS	
1600	BIOINFORMATICS WORKSHOPS continued	BIOINFORMATICS WORKSHOPS continued
1800	WORKSHOP CLOSE	
1830	BUS DEPARTS GRIFFITH UNIVERSITY FOR SURFERS PARADISE MARRIOTT RESORT	

SUNDAY 13 OCTOBER 2013

GRIFFITH UNIVERSITY, GOLD COAST CAMPUS, G39_1.17 LECTURE THEATRE

0820	BUS DEPARTS SURFERS PARADISE MARRIOTT RESORT **Delegates are requested to be on time for the departure.**
0830	REGISTRATION OPEN – ARRIVAL TEA & COFFEE
0845	COACH ARRIVES GRIFFITH UNIVERSITY
0900	DISCUSSION FORUM Dr David Lovell, Australian Bioinformatics Network & CSIRO
1000	PROJECT MANAGEMENT Dr Dominique Gorse, QFAB
1030	MANAGING RESEARCH COLLABORATIONS Assoc Prof Marcel Dinger, Garvan Institute of Medical Research
1100	MORNING REFRESHMENTS
1130	IMPROVING YOUR CV AND JOB/GRANT/FELLOWSHIP APPLICATIONS Prof John Quackenbush, Dana Farber Cancer Institute and Harvard School of Public Health, USA Assoc Prof Christine Wells, Australian Institute for Bioengineering and Nanotechnology (AIBN), The University of Queensland & Institute of Infection, Immunity and Inflammation, University of Glasgow
1230	WORKSHOP CLOSE
1235	BUS DEPARTS GRIFFITH UNIVERSITY FOR SURFERS PARADISE MARRIOTT RESORT

AMATA-13 Conference Program

SUNDAY 13 OCTOBER

SURFERS PARADISE MARRIOTT RESORT

1400 – 1800	REGISTRATION OPEN	LOBBY LOUNGE FOYER
	OPENING ORATION Chair: Assoc Prof Dan Catchpoole	LOBBY LOUNGE
1700 – 1800	WHITHER GENOMICS Guest Speaker – Emeritus Prof Ian Dawes School of Biotechnology and Biomolecular Sciences, University of New South Wales, Australia	
1800 – 2000	WELCOME RECEPTION – SURFERS PARADISE MARRIOTT RESORT	POOL PERGODA

MONDAY 14 OCTOBER

SURFERS PARADISE MARRIOTT RESORT

0730 – 1735	REGISTRATION DESK OPEN & WELCOME REFRESHMENTS	GARDEN TERRACE GALLERY
0800 – 1730	TRADE EXHIBITION OPEN	GARDEN TERRACE
0815 – 0830	OFFICIAL WELCOME AND CONFERENCE OPENING	
SESSION 1: CANCER GENOMICS Chair: Dr Richard Tothill and Dr Nick Hayward		ELSTON ROOM
0830 – 0910	THE NETWORK EFFECT: INTEGRATIVE SYSTEMS APPROACHES TO MODELING BIOLOGICAL PROCESSES Plenary Speaker – Prof John Quackenbush Dana Farber Cancer Institute and Harvard School of Public Health, USA	
0910 – 0935	PERVASIVE EVOLUTIONARY POTENTIAL AMONG TUMOURIGENIC MELANOMA CELLS Keynote Speaker – Dr Mark Shackleton Peter MacCallum Cancer Centre, Australia	
0935 – 1000	WHOLE GENOME SEQUENCING OF PANCREATIC CANCER: MUTATIONAL SIGNATURES AND NEW THERAPEUTIC OPTIONS Keynote Speaker – Dr Nic Waddell Institute for Molecular Bioscience, The University of Queensland, Australia	
1000 – 1015	A FUNCTIONAL GENOMICS APPROACH TOWARDS IDENTIFYING TUMOURIGENICITY ASSOCIATED GENES Contributed Paper – Dr Mark Waltham St Vincent's Institute of Medical Research, Australia	
1015 – 1030	GENOMIC LANDSCAPE OF HUMAN BRAIN METASTASIS Contributed Paper – Dr Michael Quinn The University of Queensland, Australia	
1030 – 1100	MORNING REFRESHMENTS	GARDEN TERRACE
SESSION 2: ECOGENOMICS Chair: Dr Paraic O'Cuiv		ELSTON ROOM
1100 – 1140	SPATIAL INSIGHTS ON SPECIATION GENOMICS IN EARLY DIVERGENCE Keynote Speaker – Dr Rose Andrew Department of Botany, University of British Columbia, Canada	
1140 – 1200	30,000 YEARS BY 1.2 BILLION BASE PAIRS: SIZING UP THE GENOMIC HISTORY OF CLIMATE ADAPTATION IN THE ADÉLIE PENGUIN Keynote Speaker – Dr Matthew Parks Science, Environment, Engineering and Technology, Griffith University, Australia	
1200 – 1220	GENETIC CONTROL OF QUANTITATIVE AND QUALITATIVE VARIATION OF PLANT SECONDARY METABOLITES IN AUSTRALIAN MYRTACEAE Keynote Speaker – Dr Carsten Kulheim Research School of Biology, Australian National University, Australia	

SESSION 3: NEW TECHNOLOGIES MODERATED DISCUSSION Chair: Dr David Lovell		ELSTON ROOM
1220 – 1250	Dr Vikram Devgan Director Functional Genomics/Biomarker Development, QIAGEN Dr Jian-Bing Fan Senior Director, Scientific Research, Illumina Dr Paul Lacaze Product Manager / Applications Scientist, Millennium Science Pty Ltd Dr Jeff Jeddeloh Roche Diagnostics	
1250 - 1400	LUNCH AMONGST THE TRADE EXHIBITORS	GARDEN TERRACE
1300 – 1330	QIAGEN GENEREADER: SAMPLE TO INSIGHT Optional Lunchtime Presentation – Dr Vikram Devgan *Sponsored by QIAGEN*	ELSTON ROOM
1400 – 1530	POSTER SESSION	GARDEN TERRACE GALLERY
1500 – 1530	AFTERNOON REFRESHMENTS	GARDEN TERRACE
SESSION 4: CLINICAL GENOMICS Chair: Prof Erik Thompson and Prof Lyn Griffith		ELSTON ROOM
1530 – 1535	THE AMATA SMALL GRANT SCHEME ANNOUNCEMENT	
1535 – 1615	BIOLOGY AND BIOMARKERS IN ORGAN FAILURE Plenary Speaker – Prof Paul Keown University of British Columbia, Canada	
1615 – 1640	MOVING GENOMICS FROM BENCH TO BEDSIDE: APPLICATION OF HIGH-THROUGHPUT SEQUENCING TO PEDIATRIC PATIENTS WITH UNRESOLVED DIAGNOSES Keynote Speaker – Dr Ryan Taft Institute for Molecular Bioscience, The University of Queensland, Australia	
1640 – 1705	GENE DISCOVERIES IN EPILEPSY AND ITS CO-MORBIDITIES USING EXOME SEQUENCING Keynote Speaker – Assoc Prof Leanne Dibbens School of Pharmacy and Medical Sciences, University of South Australia, Australia	
1705 – 1720	GENETIC VARIATION NEAR GRB2 AND KCNB2 IDENTIFIED BY A GENOME-WIDE ASSOCIATION STUDY ARE REPRODUCIBLY ASSOCIATED WITH DIABETIC RETINOPATHY Contributed Paper – Dr Kathryn Burdon Flinders University, Australia	
1720 – 1735	LIONS AND TIGERS AND SPERM TAILS? OH MY! MAKING SENSE OF VARIANT DIVERSITY FROM WHOLE-GENOME SEQUENCING Contributed Paper – Dr Jac Charlesworth Menzies Research Institute, University of Tasmania, Australia	

TUESDAY 15 OCTOBER 2013

SURFERS PARADISE MARRIOTT RESORT

0730 – 1730	REGISTRATION DESK OPEN & WELCOME REFRESHMENTS	GARDEN TERRACE GALLERY
0800 – 1730	TRADE EXHIBITION OPEN	GARDEN TERRACE
0825 – 0830	WELCOME TO DAY TWO	
SESSION 5: RNA/NONCODING Chair: Dr Ruby Lin and Dr Mark Cowley		ELSTON ROOM
0830 – 0910	THE SECRET LIVES OF STEM CELLS Plenary Speaker – Prof Jeanne Loring The Scripps Research Institute, USA *Sponsored by StemCells Australia*	
0910 – 0935	POST-TRANSCRIPTIONAL REGULATION OF MITOCHONDRIAL GENE EXPRESSION Keynote Speaker – Prof Aleksandra Filipovska Western Australian Institute for Medical Research, Australia	
0935 – 1000	FUNCTIONAL ROLE OF RNA-PROTEIN INTERACTIONS IN CELLS Keynote Speaker – Dr Minna-Liisa Änkö Walter and Eliza Hall Institute of Medical Research, Australia	
1000 – 1015	ASSESSING THE TARGET SPECIFICITY OF TRIPLEX FORMING OLIGONUCLEOTIDES USING DNA MICROARRAYS Contributed Paper – Dr Denis Bauer CSIRO, Australia	
1015 – 1030	LOC134466: A NOVEL TRANSCRIPT REGULATING CELL POLARITY THAT IS FREQUENTLY HYPERMETHYLATED IN TYPE II EPITHELIAL OVARIAN CANCER Contributed Paper – Dr Brian Gloss Garvan Institute, Australia	
1030 – 1100	MORNING REFRESHMENTS	GARDEN TERRACE
SESSION 6: PROTEOMICS Chair: Prof Kirill Alexandrov		ELSTON ROOM
1100 – 1120	IN VITRO RECONSTITUTION AND ANALYSIS PROTEIN INTERACTION NETWORKS Keynote Speaker – Prof Kirill Alexandrov Institute for Molecular Bioscience, The University of Queensland, Australia	
1120 – 1140	PROTEOMICS IN NON-MODEL ORGANISMS: FROM PLANTS TO MAMMALS AND BACK AGAIN Keynote Speaker – Dr Michelle Colgrave CSIRO Animal, Food and Health Sciences, Australia	
1140 – 1200	GETTING GENOMICS AND PROTEOMICS DATA TO WORK TOGETHER – WHAT CAN THEY TELL US ABOUT EACH OTHER? Keynote Speaker – Dr Jason Wong Lowy Cancer Research Centre, University of New South Wales, Australia	
1200 – 1215	USING MULTI-LEVEL OMICS DATA TO INFER CAUSAL RELATIONSHIPS BETWEEN CORRELATED TRANSCRIPTS AND METABOLITES Contributed Paper – Anita Goldinger The University of Queensland, Australia	

1215 – 1330	LUNCH AMONGST THE TRADE EXHIBITORS	GARDEN TERRACE
1245 – 1315	SEQCAP EPI: ULTRA-HIGH COMPLEXITY PROBE POOLS FOR TARGETED BISULFITE SEQUENCING Optional Lunchtime Presentation – Dr Jeff Jeddeloh Roche Applied Science, USA *Sponsored by Roche Diagnostics Australia Pty Ltd*	ELSTON ROOM
SESSION 7: POPULATION VARIATION Chair: Dr Carsten Kulheim		ELSTON ROOM
1330 – 1350	ANALYSING GENOMIC NETWORKS IN CANCER Keynote Speaker – Assoc Prof Cristin Print New Zealand Bioinformatics Institute, University of Auckland, New Zealand	
1350 – 1410	MAKING SENSE OF NEXT GENERATION SEQUENCE DATA WITH GENERATIVE MODELS Keynote Speaker – Dr Lachlan Coin Institute for Molecular Bioscience, The University of Queensland, Australia	
1410 – 1430	SINGLE CELL GENE EXPRESSION ANALYSIS TO REFINE CELL POPULATIONS IN MAMMARY EPITHELIUM HIERARCHY Keynote Speaker – Dr Bhupinder Pal The Walter and Eliza Hall Institute of Medical Research, Australia *Sponsored by Millennium Science and Fluidigm*	
1430 – 1445	IDENTIFICATION OF SINGLE NUCLEOTIDE POLYMORPHISMS (SNPS) INVOLVED IN THE DETERMINATION OF HUMAN PHYSICAL APPEARANCE Contributed Paper – Mark Barash Bond University, Australia	
1445 – 1500	DETECTION AND REPLICATION OF EPISTASIS INFLUENCING TRANSCRIPTION IN HUMANS Contributed Paper – Dr Joseph Powell The University of Queensland, Australia	
1500 – 1530	AFTERNOON REFRESHMENTS	GARDEN TERRACE

Visit the Trade Exhibitors now!

Don't forget to get your Exhibitor Passport stamped to go into the draw to win some great prizes!

SESSION 8: COMPUTATIONAL BIOLOGY Chair: Dr Denis Bauer and Dr Jac Charlesworth	ELSTON ROOM
1530 – 1610	BRIDGING THE GAP: ENABLING TOP RESEARCH IN TRANSLATIONAL RESEARCH Plenary Speaker – Prof Knut Reinert Freie Universität Berlin, Germany
1610 – 1635	RNA-SEQ: FROM READS TO GENES TO PATHWAYS Keynote Speaker – Prof Gordon Smyth Walter and Eliza Hall Institute of Medical Research, Australia
1635 – 1700	DATA INTEGRATION OF HIGHLY DIMENSIONAL BIOLOGICAL DATA SETS WITH MULTIVARIATE ANALYSIS Keynote Speaker – Dr Kim-Anh Lê Cao Queensland Facility for Advanced Bioinformatics, The University of Queensland, Australia
1700 – 1715	BLUE: FAST, ACCURATE ERROR CORRECTION USING K-MER CONSENSUS AND CONTEXT Contributed Paper – Paul Greenfield CSIRO Computational Informatics, Australia
1715 – 1730	BUILDING TISSUE-SPECIFIC PROMOTER-ENHANCER REGULATORY MAPS Contributed Paper – Timmothy O'Connor Institute For Molecular Biosciences, The University of Queensland, Australia
1900 – 2300	CONFERENCE DINNER — WAIANBAH ROOM

Have you booked a ticket to the Conference Dinner? There may still be tickets available!

Please see the staff at the Conference Registration Desk for more information.

WEDNESDAY 16 OCTOBER

SURFERS PARADISE MARRIOTT RESORT

0730 – 1730	REGISTRATION DESK OPEN & WELCOME REFRESHMENTS	GARDEN TERRACE GALLERY
0800 – 1530	TRADE EXHIBITION OPEN	GARDEN TERRACE
0900 – 0910	WELCOME TO DAY THREE	
SESSION 9: APPLIED GENOMICS Chair: Dr Brooke Gardiner and Liam Williams		ELSTON ROOM
0910 – 0950	HIGH RESOLUTION TRANSCRIPTOME ANALYSIS – ONE CELL AT A TIME Plenary Speaker – Dr Jian-Bing Fan Senior Director of Scientific Research, Illumina *Sponsored by Illumina*	illumina®
0950 – 1015	CLINICAL IMPLEMENTATION OF NEXT-GENERATION SEQUENCING FOR DIAGNOSTICS Keynote Speaker – Dr Karin Kassahn Genetics and Molecular Pathology, SA Pathology, Australia	
1015 – 1030	DETECTION OF MARKERS THAT DISTINGUISH HUMAN FETAL AND MATERNAL DERIVED MESENCHYMAL STEM/STROMAL CELLS (MSCS) Contributed Paper – Dr Celena Heazlewood Australian Institute For Bioengineering And Nanotechnology, Australia	
1030 – 1100	MORNING REFRESHMENTS	GARDEN TERRACE
SESSION 10: EPIGENETICS Chairs: Dr Stacey Edwards and Dr Lutz Krause		ELSTON ROOM
1100 – 1140	ENHANCER RNA BIOGENESIS AND AN ATLAS OF TRANSCRIBED ENHANCERS ACROSS THE HUMAN BODY Plenary Speaker – Assoc Prof Albin Sandelin The Bioinformatics Centre, Denmark	
1140 – 1200	THE ROLE OF MYST LYSINE ACETYLTRANSFERASES IN TRANSCRIPTIONAL REGULATION AND CHROMATIN CONFORMATION Keynote Speaker – Dr Anne Voss Walter and Eliza Hall Institute of Medical Research, Australia	
1200 – 1220	MOTIF-BASED ANALYSIS OF CHIP-SEQ DATA Keynote Speaker – Assoc Prof Tim Bailey Institute for Molecular Bioscience, The University of Queensland, Australia	
1220 – 1235	EPIGENETIC REGULATION OF RIBOSOMAL GENE TRANSCRIPTION DURING MALIGNANT TRANSFORMATION Contributed Paper – Dr Jeannine Diesch Peter MacCallum Cancer Centre, Australia	
1235 – 1330	AGM AND LUNCH SERVED DURING	ELSTON ROOM / BALLROOM GALLERY
1235 – 1330	STUDENT AND PLENARY SPEAKER LUNCH	VERANDAH ROOM

SESSION 11: RNA/TRANSCRIPTOMICS Chair: Assoc Prof Marcel Dinger and Dr Mark Waltham		ELSTON ROOM
1330 – 1410	POWER AND LIMITATIONS OF RNA-SEQ: A PROGRESS REPORT FROM THE SEQC PROJECT Plenary Speaker – Dr Leming Shi School of Pharmacy, Fudan University, China	
1410 – 1430	DECODING MIRNA REGULATED GENETIC CIRCUITS Keynote Speaker – Dr Nicole Cloonan Queensland Institute of Medical Research, Australia	
1430 – 1450	GENOMIC DARK MATTER: THE COMPLEXITY OF LONG NON-CODING RNAS FROM MECHANISM TO THERAPEUTIC Keynote Speaker – Prof Kevin Morris School of Biotechnology and Biomedical Sciences, University of New South Wales, Australia	
1450 – 1505	INTRON RETENTION COUPLED WITH NONSENSE MEDIATED DECAY: A PREVIOUSLY OVERLOOKED MECHANISM OF GENE EXPRESSION CONTROL IN GRANULOPOIESIS Contributed Paper – Dr Justin Wong Centenary Institute, Australia	
1505 – 1530	AFTERNOON REFRESHMENTS	GARDEN TERRACE
SESSION 12: PLATFORM GENOMICS Chair: Dr Nicole Cloonan and Assoc Prof Christine Wells		ELSTON ROOM
1530 – 1610	BIOINFORMATICS SERVICES IN AUSTRALIA – A COLLABORATION WITH THE EUROPEAN BIOINFORMATICS INSTITUTE Keynote Speaker – Graham Cameron Bioinformatics Resource Australia – EMBL, Australia	
1610 – 1630	HIGH THROUGHPUT FUNCTIONAL GENOMICS APPROACHES TO IDENTIFYING NOVEL DRIVERS OF MELANOMA Keynote Speaker – Assoc Prof Brian Gabrielli Diamantina Institute, The University of Queensland, Australia	
1630 – 1700	AWARDING OF PRIZES, 2014 CONFERENCE LAUNCH AND CONFERENCE CLOSE	

Conference Social Program

Welcome Reception

Date
Sunday 13 October 2013

Venue
Pool Pergoda
Surfers Paradise Marriott Resort

Time
6.00pm – 8.00pm

Dress
Business or Smart Casual

Join us for the official Welcome Reception for the AMATA 2013 Conference at the Lagoon Poolside, Surfers Paradise Marriott Resort. Enjoy networking with old and new acquaintances whilst enjoying drinks and canapes along the poolside. The Welcome Reception is included in a full registration only. Additional tickets can be purchased at $70.00 per person.

Conference Dinner

Proudly sponsored by

QIAGEN®

Date
Tuesday 15 October 2013

Venue
Waianbah Room
Surfers Paradise Marriott Resort

Time
7.00pm – 11.00pm

Dress
Business or Smart Casual

Come and join us on Tuesday evening for another chance to network and meet with colleagues, whilst enjoying a great night of food and wine at the Surfers Paradise Marriott Resort.

The conference dinner is not included in any registration type. Tickets can be purchased at $130.00 per person.

Important Information

Seating and table allocation for the AMATA Dinner will be by way of sticker allocation. All delegates registered to attend the AMATA Dinner will receive a sticker to be placed on the table sheets near the Registration Desk. These sheets will be available from Monday 14 October and will be taken down at the end of morning refreshments on Tuesday 15 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.

Student/Speaker Lunch

Date
Wednesday 16 October 2013

Venue
The Verandah Room
Surfers Paradise Marriott Resort

Time
12.35pm – 1.30pm

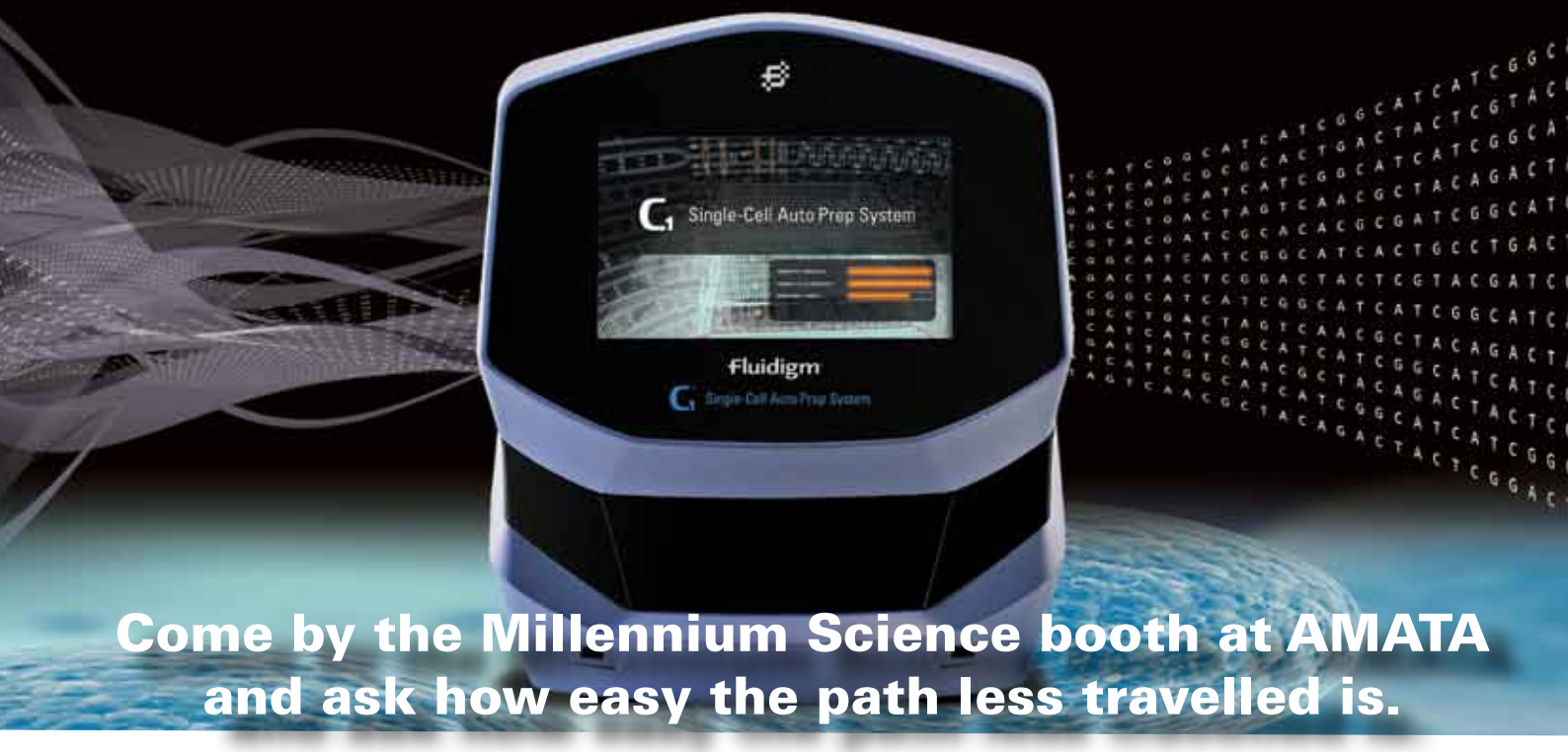
Students and 2013 plenary & keynote presenters are invited to attend a networking lunch on the last day of the program. This is a great opportunity for AMATA students to meet and interact with our guest speakers.

Fluidigm

THE PATH LESS TRAVELLED JUST GOT EASIER

C1™ SINGLE-CELL AUTO PREP SYSTEM FOR mRNA SEQUENCING OF INDIVIDUAL CELLS

- **Complete**
  End-to-end workflow for single-cell transcriptome analysis
- **Easier**
  Streamlined sample prep from cells, 3 hours hands-on time, no fragmentation or purification
- **Highest throughput**
  Unprecedented parallel processing of 96 cells per run
- **Affordable**
  One-seventh the cost of other library prep systems



Come by the Millennium Science booth at AMATA and ask how easy the path less travelled is.

millennium science

www.mscience.com.au | sales@mscience.com.au
Australia 1800 678 242 | New Zealand 0800 507 896

DAY 1 SUNDAY 13 OCTOBER

OPENING ORATION

CHAIR: ASSOC PROF DAN CATCHPOOLE

1700 – 1800

WHITHER GENOMICS

Guest Speaker – Emeritus Prof Ian Dawes

SCHOOL OF BIOTECHNOLOGY AND BIOMOLECULAR SCIENCES,
UNIVERSITY OF NEW SOUTH WALES, AUSTRALIA

BIOGRAPHY

Ian W. Dawes gained his BSc (UNSW) in 1965. He was the 1966 NSW Rhodes scholar obtaining his DPhil in Biochemistry from Oxford in 1969. He is currently Emeritus Professor at the University of New South Wales and was Director of the Ramaciotti Centre for Gene Function Analysis at UNSW from its inception in 2000 until 2010. He was Chair of the International Committee of Yeast Genetics and Molecular Biology and has been President of ASBMB and the Society for Free Radical Research Australasia. The powerful modern techniques of functional genomics have made possible very rapid progress in determining how cells regulate sets of pathways in the cell, and how the control systems interact with each other and he has applied these techniques to his research interests in the area of regulation of gene expression in the yeast Saccharomyces cerevisiae during cell development, redox homeostasis and in response to metabolites and stress.

DAY 2

MONDAY 14 OCTOBER

SESSION 1
CANCER GENOMICS

CHAIR: DR RICHARD TOTHILL AND DR NICK HAYWARD

0830 – 0910

THE NETWORK EFFECT: INTEGRATIVE SYSTEMS APPROACHES TO MODELING BIOLOGICAL PROCESSES

Plenary Speaker – Prof John Quackenbush

DANA FARBER CANCER INSTITUTE
AND HARVARD SCHOOL OF PUBLIC HEALTH, USA

BIOGRAPHY

John Quackenbush earned his PhD in theoretical particle physics from UCLA in 1990 and then completed a postdoctoral fellowship in experimental high energy physics. After receiving a fellowship from the National Center for Human Genome Research, he worked with Glen Evans on the physical mapping of human chromosome 11, and later with Richard Myers and David Cox on large-scale DNA sequencing of chromosomes 21 and 4. In 1998 he joined the faculty at The Institute for Genomic Research (TIGR) where his work focused on the use of genomic and computational methods for the study of human disease. He joined DFCI in 2005 where his work has increasingly focused on the analysis of women's cancers although the methods he and his group develop can be broadly applied. In 2009 he launched the Center for Cancer Computational Biology (CCCB), a Dana-Farber Strategic Plan Center focused on providing computational support more broadly to the DFCI research community.

ABSTRACT

Two trends are driving innovation and discovery in biological sciences: technologies that allow holistic surveys of genes, proteins, and metabolites and the growing realization that analysis and interpretation of the resulting requires an understanding of the complex factors that mediate the link between genotype and phenotype. The growing body of biological and biomedical information, driven by an exponential drop in the cost of generating genomic data, provides an outstanding opportunity for leveraging what we already "know" in a systematic way to understand the problems we are studying. Here, I will provide an overview of some of the methods we are using to investigate the complexities of human phenotypes and to explore how we can use biological data to uncover the cellular networks and pathways that underlie human disease, building predictive models of those networks that may help to direct therapies, with an emphasis on exploring functional pathways in ovarian cancer.

0910 – 0935

PERVASIVE EVOLUTIONARY POTENTIAL AMONG TUMOURIGENIC MELANOMA CELLS

Keynote Speaker – Dr Mark Shackleton

PETER MACCALLUM CANCER CENTRE, AUSTRALIA

BIOGRAPHY

Dr Mark Shackleton is a Medical Oncologist and Group Leader of the Melanoma Research Laboratory at the Peter MacCallum Cancer Centre. After training in medical oncology in Melbourne and at the Ludwig Institute, Dr. Shackleton did his PhD at the WEHI and post-doctoral work at the University of Michigan, publishing widely in journals such as Nature, Cell and the New England Journal of Medicine. A winner of the 2006 Victorian Premier's Award for Medical Research and a 2010 NHMRC Achievement Award, Dr Shackleton is a current Pfizer Australia Senior Research Fellow and a Fellow of the Victorian Endowment for Science, Knowledge and Innovation and the 2012-13 recipient of the Australian Science Minister's Prize for Life Scientist of the Year. His laboratory at Peter Mac focuses on understanding mechanisms of melanoma initiation and propagation.

ABSTRACT

Recent studies of certain cancers have revealed unexpected levels of genetic heterogeneity. However, it is uncertain what extent genetic heterogeneity affects tumourigenic cells capable of contributing to disease progression and whether these cells are also epigenetically heterogeneous. To address these questions, we have used a highly efficient in vivo xenotransplantation assay, transplanting single cells isolated from patient melanomas and creating

multi-generational clonal tumour families. Tumors in these families have been analysed for copy number alterations (CNAs) and DNA methylation. Extensive CNAs and DNA methylation differences were observed in comparisons of related clones in each family. These divergent changes affected genes that are functionally important in melanoma, DNA repair and cancer therapy resistance. A high proportion of tumourigenic melanoma cells therefore forms genetically and epigenetically distinct progeny. This contrasts with the notion that most tumorigenic cancer cells form tumors that recapitulate their tumour-of-origin.

0935 – 1000

WHOLE GENOME SEQUENCING OF PANCREATIC CANCER: MUTATIONAL SIGNATURES AND NEW THERAPEUTIC OPTIONS

Keynote Speaker – Dr Nic Waddell
INSTITUTE FOR MOLECULAR BIOSCIENCE,
THE UNIVERSITY OF QUEENSLAND, AUSTRALIA

BIOGRAPHY

Dr Waddell is a cancer researcher at the Queensland Centre of Medical Genomics (QCMG) at the IMB in the University of Queensland. She gained her PhD in 2003 from the University of Leicester, UK. She has held post doc positions at QIMR and the University of Queensland. Her research focuses on the analysis of next generation sequence data from cancer genomes. She is a member of the Australian International Cancer Genome Consortium (ICGC) sequencing project (PI Grimmond) and sits on the ICGC sequencing validation subgroup which is attempting to implement common validation methods throughout the ICGC. Recently she has published in journals such as Nature and PNAS.

ABSTRACT

Next generation sequencing of cancer genomes is redefining the mutation landscape of many tumour types. Two large consortia have been established to undertake sequence analysis of thousands of tumours, the TCGA (The Cancer Genome Atlas) and the ICGC (International Cancer Genome Consortium). As part of the International Cancer Genome Consortium we are sequencing 350 pancreatic tumours and matched normal samples. Pancreatic cancer is the fourth cause of cancer death in Australia and during the last 40 years there has been no improvement in survival rates. This is in depth analysis of the genome, transcriptome and epigenome is enabling a better understanding of the disease and identifying mutational signatures in each tumour. Encouragingly for subsets of pancreatic tumours alternative therapeutic opportunities have been identified, some of which involve repurposing existing cancer drugs to treat pancreatic cancer. This talk will summarize some of the key outcomes of whole genome sequencing of different tumours and will highlight the promising clinical potential of these findings.

1000 – 1015

A FUNCTIONAL GENOMICS APPROACH TOWARDS IDENTIFYING TUMOURIGENICITY ASSOCIATED GENES

Contributed Paper – Dr Mark Waltham
ST VINCENT'S INSTITUTE OF MEDICAL RESEARCH, AUSTRALIA

BIOGRAPHY

Mark Waltham (PhD, 1990) has a primary research interest in cancer biology, anticancer therapeutics and bioinformatics. He undertook postdoctoral training at Sloan-Kettering Cancer Institute (New York, 1990-94) and then as a Visiting Fellow at NCI (Bethesda, 1995-98) he worked on bioinformatic and profiling based approaches to anti-cancer drug discovery. He was recruited to St. Vincent's Institute (SVI, Australia) in 2001 and appointed as Unit Head (Pharmacogenomics, 2002), which is one of twelve research units within SVI. He remains today as Unit Head and Senior Faculty within SVI. While his main role at SVI is an academic researcher, his lab also performs contract based research work (drug evaluation) using animal models of cancer progression and metastasis for local and US biopharmaceutical companies.

ABSTRACT

Breast cancer is the leading cause of cancer related deaths among women in the western world. Research mouse models of the disease are used extensively to understand the biology and also to evaluate potential therapeutics before progressing to clinical trial. Xenograft models, where human cell lines or patient samples are grafted into immunocompromised mice, are a well established tool, however not all lines or patient-derived samples successfully graft into animals. Further to this, it is primarily the more dedifferentiated and more aggressive (mesenchymal-like) lines and patient samples that are successfully grafted (i.e. tumourigenic). The current study sought to identify genes or pathways that control the process of tumourgenicity. The poorly-tumourigenic human breast cancer line PMC42-LA was transduced with

a 5,000-member shRNA library targeting genes associated with cell polarity, cancer stem cell state, metastasis and those associated with epithelial mesenchymal plasticity. These transduced cells, both as a total shRNA pool and also as 10 smaller sub-pools, were then inoculated into the mammary fat pad of immunocompromised mice "Dear Delegates and Invited Guests,"and tumour take and/or growth monitored. A markedly increased rate of tumourigenicity was observed in mice injected with the hairpin-transduced cells relative to untransduced or scrambled controls. Moreover, next-generation sequencing of hairpins in the resulting tumours identified a marked enrichment of a select set of hairpins that were consistently identified across tumours isolated from different mice. The biological implications and technical aspects of using next-generation technology in functional genomics screening will be discussed. This work was supported by USDOD Grant #BC084667 and the NBCF.

1015 – 1030

GENOMIC LANDSCAPE OF HUMAN BRAIN METASTASIS

Contributed Paper – Dr Michael Quinn
THE UNIVERSITY OF QUEENSLAND, AUSTRALIA

BIOGRAPHY

Michael Quinn completed his PhD in the Department of Anatomy at Otago University in Dunedin, New Zealand. He undertook a postdoc at McGill University in Montreal, Canada investigating gene expression profiling of ovarian cancer. In 2012 he started at the Queensland Centre of Medical Genomics working in collaboration with the UQ Centre for Clinical Research on a project studying the genomics of brain metastases.

ABSTRACT

Metastasis from the primary cancer to the brain is one of the leading causes of all cancer related deaths, with a prevalence of approximately 40% and a median survival of nine months following treatment. The etiology and underpinning molecular mechanisms of brain metastasis are not well understood, although recent research has implicated the Wnt/TCF4 and HER/EGFR pathways in lung and breast cancer metastases respectively. Very limited next generation sequencing studies of brain metastases have been conducted to date. The present study investigated the genomic landscape of brain metastases from 36 patients (with matched normal samples) that originated from a range of primary tumour sites (including breast, lung and melanoma) utilizing both Illumina HumanOmni 2.5M SNP arrays and Illumina 2500 HiSeq Exome sequencing. Copy number events were determined from the SNP arrays using GAP and GenoCN tools. Somatic substitutions and indels were identified using qSNP and Pindel, and verification was performed using the in-house qVerify tool. Key biological pathways were investigated using pathway analysis tools. Particular attention was paid to the NRG-HER3 axis, which has been shown to be upregulated in breast metastases compared to matched primary breast tumours. It is envisaged that the current work may provide important biological insights into the genomic characterization and evolution of brain metastasis and additionally identify druggable targets of clinical relevance.

SESSION 2
ECOGENOMICS

CHAIR: DR PARAIC O'CUIV

1100 – 1140

SPATIAL INSIGHTS ON SPECIATION GENOMICS IN EARLY DIVERGENCE

Keynote Speaker – Dr Rose Andrew
DEPARTMENT OF BOTANY, UNIVERSITY OF BRITISH COLUMBIA, CANADA

BIOGRAPHY

During my PhD, I studied the genetic basis and chemical ecology of variation in antiherbivore defence compounds of Eucalyptus, which have a critical influence on interactions with both mammalian and insect herbivores. After graduating in 2007, I moved to Canada for a postdoctoral fellowship at the University of British Columbia, where my main focus was on adaptive divergence and speciation in wild sunflowers. I began to use genomic tools and explore their application to ecological and evolutionary questions. I am now a visiting scientist at the Australian National University, studying the landscape genomics of Eucalyptus.

ABSTRACT

Reports that ecological speciation in the presence of gene flow may involve large genomic regions of divergence have stimulated work attempting to explain how genomes diverge. However, in plants there is little evidence for this, possibly because most

comparisons to date have been at the species level, where long histories of gene flow, the accumulation of genome-wide divergence and population structure might erode or obscure such patterns. For example, the annual *Helianthus* sunflowers have been a model system for ecological speciation, but as yet we have little information on the earliest stages of the process. I will present data exploring ecotypes of the prairie sunflower, *Helianthus petiolaris*, occurring at Great Sand Dunes National Park (Colorado). We have found that genomic regions of elevated divergence are large and that selective sweeps have been mainly restricted to the dune ecotype. Spatial analysis within the genome highlights contrasts between early divergence and older, more diverged species.

The geographic context of speciation is also likely to be still preserved in young incipient species, and departures from isolation by distance (or isolation by landscape resistance) can be used to understand the selective pressures operating at this stage. The genome varies in whether proximity of the dune ecotype to the ancestral form promotes or constrains divergence. I will describe our investigations of a genomic signature reminiscent of reinforcement, where divergence is enhanced by gene flow between populations. These spatial insights have implications for when and how this controversial process may be detected.

1140 – 1200

30,000 YEARS BY 1.2 BILLION BASE PAIRS: SIZING UP THE GENOMIC HISTORY OF CLIMATE ADAPTATION IN THE ADÉLIE PENGUIN

Keynote Speaker – Dr Matthew Parks
SCIENCE, ENVIRONMENT, ENGINEERING AND TECHNOLOGY, GRIFFITH UNIVERSITY, AUSTRALIA

BIOGRAPHY

My research background centers on the application of high-throughput DNA sequence data to phylogenetic and population genetic questions. With formal training as a botanist at the University of Idaho and Oregon State University (USA), much of my previous work involves plant phylogenetics, primarily organellar genomics in pines (genus Pinus) and strawberries (genus Fragaria). In 2012, I joined David Lambert's group at Griffith University in Brisbane as a research fellow, and we are now investigating climate-related adaptation and contemporary and historic population trends in the Adélie penguin (Pygoscelis adeliae). Our work utilizes whole-genome sequencing and analyses of both modern and ancient Adelie samples up to 30,000 years old.

ABSTRACT

Species may adapt to a changing climate through alterations of behavior or genetic composition. In certain cases, a species' characteristics may preclude behavioral changes, largely shifting the burden of adaption to the genome. The Adélie penguin (*Pygoscelis adeliae*), which is known only to breed in ice-free areas of Antarctica, is one such species. Studies of modern population and genetic structure can thus provide significant insight into genomic adaptation in this 'bellwether' species as global temperatures continue to rise. Unlike many species, however, investigations of historical climate adaptation in the Adélie penguin are not limited to modern collections. Abandoned nesting colonies aged up to 30,000 years contain well-preserved bone and tissue samples from which DNA can be extracted. This window of history, dating to before the last glacial maximum, encompasses an estimated 8-10 °C increase in average global temperatures. The Adélie thus presents a unique opportunity to directly interrogate genetic adaptation to long-term climate change on a genomic scale. In this talk, I will present data and analyses from Illumina (HiSeq2000) sequencing of nearly two-dozen 1.2 Gbp complete genomes from modern Adélie colonies around Antarctica, as well as 60 ancient Adélie samples. Our results are providing unprecedented acuity into the demographic history and genetic architecture of this important species. Perhaps more importantly, our data lend insight into the complex nature of genetic adaptation to climate change, and suggest what might be expected of or required for species survival in the coming decades of global temperature increase.

1200 – 1220

GENETIC CONTROL OF QUANTITATIVE AND QUALITATIVE VARIATION OF PLANT SECONDARY METABOLITES IN AUSTRALIAN MYRTACEAE

Keynote Speaker – Dr Carsten Kulheim
RESEARCH SCHOOL OF BIOLOGY, AUSTRALIAN NATIONAL UNIVERSITY, AUSTRALIA

BIOGRAPHY

Carsten Kulheim graduated from the Umea University in Sweden with a major in Plant Molecular Biology in 2000. He completed his PhD in 2005 at Umea University, having worked with short term adaptation mechanisms in photosynthesis to changes in light

environment. Carsten then moved on for a two-year Post Doc at The University of British Columbia in Vancouver, Canada, where he worked on plant-herbivore interactions using poplar as a model system. In 2008 he joined the Australian National University continuing in the field of plant-herbivore interactions with the focus on Australian Myrtaceae. His current work focuses on the genetic basis of quantitative and qualitative variation of plant secondary metabolites in eucalypts and tea tree.

ABSTRACT

Variations in Plant Secondary Metabolites (PSMs) have profound influence in mediating the interactions between plants and animals. The variation within a species can occur on a number of scales including qualitative variation within a single tree canopy (mosaic trees), or between neighboring trees and quantitative variation across the landscape. How this variation arises and is maintained has been of wide interest to ecologists and to date, most research has focused on resource availability via environmental variation. However, the heritability of variation in some PSMs is high (h^2> 0.7) suggesting that genetic control of variation is significant. Here I describe a number of studies that elucidate the genes that underlie some of this variation.

Variation in the qualitative profile of foliar terpenes in both *Eucalyptus* and *Melaleuca* is due to genomic presence/absence patterns and/or transcriptomic control of a large gene family (113 members in *Eucalyptus grandis*) called terpene synthases. These genes encode enzymes that use common intermediates (geranyl pyrophosphate or farnesyl pyrophosphate) to produce a large variety of mono – and sesquiterpenes, which have profound effects on plant-herbivore interactions. Quantitative variation in terpenes in *Eucalyptus globulus*, *E. loxophleba* and *Melaleuca alternifolia* is partly due to single nucleotide polymorphisms in genes of the terpene biosynthetic pathway. Although as expected, each polymorphism explains only a small amount of phenotypic variation. Low levels of linkage disequilibrium in outcrossing forest trees limit the opportunities for genome-wide association studies for now, but identifying transcriptional regulators that control biosynthetic pathways of PSMs remains a high priority. These studies are the first to unravel the molecular basis of quantitative variations in ecologically important traits in Australian Myrtaceae.

SESSION 3
NEW TECHNOLOGIES
MODERATED DISCUSSION

CHAIR: DR DAVID LOVELL

1220 – 1250

New technologies in molecular bioscience are at the very heart of AMATA. In this session you will get to hear perspectives from four companies operating at the technological frontier:

- Dr Vikram Devgan, Director Functional Genomics/ Biomarker Development, QIAGEN
- Dr Jian-Bing Fan, Senior Director, Scientific Research, Illumina
- Dr Paul Lacaze, Product Manager/Applications Scientist, Millennium Science Pty Ltd
- Dr Jeff Jeddeloh, Roche Diagnostics

Will join Dr David Lovell, Australian Bioinformatics Network to explore some of the opportunities and challenges we face in a technology landscape where the only constant is change.

BIOGRAPHIES

Vikram Devgan is a Director of Product Development at QIAGEN. Dr. Devgan received his Ph.D from Indian Institute of Science and did post-doctoral research at Harvard Medical School. Dr. Devgan has more than 10 years experience in life science research industry and involved in technology and product development in Next Generation Sequencing, qPCR and cell based assay arena.

Jian-Bing Fan, Senior Director of Scientific Research, Illumina, Inc., has over 20 years of experience in human genome research and microarray/next-gen sequencing technology development. He played important roles in developing the BeadArray technology and a series of genomics products at Illumina for high throughput genotyping, gene expression and methylation analysis. Currently, Dr. Fan is leading the Illumina Dx research efforts to develop highly sensitive and quantitative assays for use with medical specimens such as archived tissue samples, blood samples, circulating tumor cells and cell-free DNA. He has applied all these technologies in a broad spectrum of clinical applications such as non-invasive prenatal testing (NIPT), preimplantation genetic diagnosis and screening (PGD/ PGS), and cancer profiling and early detection.

Paul Lacaze is a Field Application Scientist and Product Manager at Millennium Science. His role involves delivering high-level technical training and ongoing scientific support to Australia's leading molecular diagnostics, high-throughput genotyping, AgBio and single-cell genomics laboratories. Paul specializes in genomic tools such as Fluidigm's microfluidic PCR systems (C1 Single-Cell Autoprep, BioMark HD) and Pacific Biosciences (PacBio) single-molecule real time sequencing.

Jeff Jeddeloh is currently Director of Technology Innovation at Roche Applied Science, based in Madison, WI. In this role, Jeff identifies and secures innovative, disruptive technologies for the future life science portfolio of Roche. Prior to joining Roche, Jeff was the Science and Technology Director at Orien Genomics where he managed platform development and implementation for bio marker discovery and the epigenetics molecular diagnostics business. Jeff has a PhD in Molecular Genetics from Washington University in St. Louis.

OPTIONAL LUNCHTIME PRESENTATION

1300 – 1330

QIAGEN GENEREADER: SAMPLE TO INSIGHT

Dr Vikram Devgan
DIRECTOR FUNCTIONAL GENOMICS/ BIOMARKER DEVELOPMENT, QIAGEN

Sponsored by

QIAGEN®

BIOGRAPHY

Vikram Devgan is a Director of Product Development at QIAGEN. Dr. Devgan received his Ph.D from Indian Institute of Science and did post-doctoral research at Harvard Medical School. Dr. Devgan has more than 10 years experience in life science research industry and involved in technology and product development in Next Generation Sequencing, qPCR and cell based assay arena.

ABSTRACT

Next generation sequencing (NGS) has revolutionized the way genomic information can be extracted. NGS facilitates rapid advances in the fields of basic and clinical research, as well as molecular diagnostics and drug development. However, certain bottlenecks in the NGS workflow still pose challenges to researchers. QIAGEN is introducing sample to insight workflow, which provides a streamlined and integrated approach to next-generation sequencing (NGS), from sample preparation to the biological interpretation of sequencing data. The majority of the targeted DNA sequencing workflow is automated, ensuring greater standardization and more accurate results. Preanalytic sample preparation can be done with either the QIAcube® or QIAsymphony®, apart from the target enrichment step, which utilizes an easy to use and fast approach for amplification of tens to hundreds of genes of interest in a highly multiplexed PCR. The QIAGEN GeneReader benchtop sequencer is at the core of the workflow, and offers unprecedented scalability and flexibility in an NGS instrument. Highly accurate and cost-effective, proven sequencing-by-synthesis chemistry, along with its unique ability to process up to 20 flow cells in parallel through an innovative turntable design, make the GeneReader a scalable NGS system that can grow with increasing throughput needs. Additionally, the continuous loading mode of the GeneReader allows researchers to add additional flow cells during an ongoing run, significantly decreasing the overall turnaround time for projects that require multiple flow cells. The data analysis of targeted DNA sequencing is powered by of Ingenuity®. QIAGEN is employing market-leading NGS data interpretation by integrating cloud-based Variant Analysis™ software into GeneRead enrichment panels to enable the rapid translation of raw sequencing data into actionable results.

SESSION 4
CLINICAL GENOMICS

CHAIR: PROF ERIK THOMPSON AND PROF LYN GRIFFITH

1535 – 1615

BIOLOGY AND BIOMARKERS IN ORGAN FAILURE

Plenary Speaker – Prof Paul Keown

UNIVERSITY OF BRITISH COLUMBIA, CANADA

BIOGRAPHY

Dr. Paul A. Keown is Professor of Medicine and Director of Immunology at the University of British Columbia, with appointments in Medicine, Pathology and Laboratory Medicine. Dr. Keown graduated in Medicine from the University of Manchester, and pursued postgraduate training in England, France, and Canada. He holds research Doctorates in both Medicine and in Science from the University of Manchester, and an MBA from Simon Fraser University in British Columbia. Dr. Keown's research focuses particularly on the immune response in transplantation and autoimmune disease, and ranges from molecular genetics to healthcare economics.

He has served as Executive Director of the British Columbia Transplant Program, Head of the UBC Division of Nephrology, president of the Canadian Transplant Society, Vice President and member of the Executive Committee of the Transplantation Society, and in numerous other national and international scientific societies and professional organisations. He is a Fellow of the Royal College of Physicians of Canada, the Royal College of Physicians of London, the Royal College of Pathologists, the Royal Society of Chemistry, the Society of Biology, the American College of Physicians, and the American Society of Nephrology. Dr. Keown is the founder and C.E.O. of Syreon Corporation, a global research corporation specializing in the use of advanced information technologies for health sciences research.

ABSTRACT

Vital organ dysfunction is associated with profound changes in gene and protein expression that provide important insights into the cellular biology of these states, highlight potential targets for therapeutic intervention, and offer the opportunity for development of clinical biomarkers for diagnosis and therapy. Over 9000 genes are differentially expressed in kidney failure, of which the majority are down-regulated. Gene set enrichment analysis shows the mRNA processing and transport, protein transport, chaperone functions, the unfolded protein response and other key cellular functions are prominently inhibited while the complement system, liproprotein metabolism and other functions are up-regulated. Organ transplantation causes a rapid and highly dynamic perturbation of gene networks including chemotaxis and cell migration, inflammation and innate immunity and wound and tissue healing. Transcripts for many key cytokines and chemokines, which are respectively enhanced and reduced during renal failure and dialysis consistent with the complex inflammatory nature of this state, gradually normalize during the first months after organ transplantation. Even in healthy and well-functioning graft recipients, however, gene expression differs markedly from normal reflecting residual alterations in cell biology. The occurrence of acute graft rejection engenders a distinct alteration in gene and protein expression that reflects changes in the genes encoding cytoskeletal organisation and biogenesis, signal transduction, immune system processes, cell motility and leukocyte activation. Alterations in the plasma proteome include increases in proteins that encompass processes related to inflammation, complement activation, blood coagulation and wound repair. These may be integrated to create sensitive and specific biomarkers of rejection or quiescence.

1615 – 1640

MOVING GENOMICS FROM BENCH TO BEDSIDE: APPLICATION OF HIGH-THROUGHPUT SEQUENCING TO PEDIATRIC PATIENTS WITH UNRESOLVED DIAGNOSES

Keynote Speaker – Dr Ryan Taft

INSTITUTE FOR MOLECULAR BIOSCIENCE, THE UNIVERSITY OF QUEENSLAND, AUSTRALIA

BIOGRAPHY

Ryan Taft is a Laboratory Head and Senior Research Fellow at the University of Queensland's Institute for Molecular Bioscience. He obtained his Bachelor of Science in Biochemistry and Molecular Biology from the University of California, Davis on a Regent's Scholarship, and his PhD in Genomics and Computational Biology from the University of Queensland on a US National Science Foundation Research Fellowship. He has published in a variety of well-regarded journals – including articles in Nature Genetics, Nature Structural and Molecular Biology, the American Journal of Human Genetics, RNA and others – and has had his work recognised by the Australia

Museum Eureka Prizes and the UQ Foundation Research Excellence Awards. His current research is focused on the cryptic genetics of disease, with an emphasis on rare paediatric disorders and the role of non-protein-coding RNAs in epigenetic processes.

ABSTRACT

Rare diseases, the vast majority of which are genetic in origin, have a particularly detrimental effect on children. More than 50% of rare disease patients are paediatric, and at least 30% of these children will not live to see their first birthday. Additionally, many of these patients will be incorrectly diagnosed (at least 40% according to EURODIS), or will remain without a final diagnosis. For example, it is well established that half of the patients with leukodystrophies, rare central nervous system white-matter disorders, remain a diagnostic mystery despite the fact that all are known to have a genetic basis. Here we present a number of studies showing the successful application of high-throughput genome and exome sequencing to paediatric patients with unresolved diagnoses. By partnering with clinician scientists in Australia, the US, and Europe we have i) identified the mutations in DARS responsible for an Australian boy's leukodystrophy, which subsequently led to the characterisation of a novel disease (HBSL), ii) in less than four months, identified the de novo TUBB4A mutation responsible for a leukodystrophy unknown genetic aetiology, H-ABC, iii) identified a novel de novo mutation in KCNT1, a potassium transporter, in a child with unclassifiable white matter defects and severe epilepsy, which led to treatment with specific channel therapies, and iv) in less than a month, identified the mutation responsible for an unexplainable case of Leigh disease, which led to a change in patient care. In collaboration with the Global Leukodystrophy Initiative (GLIA) and others, we are now working to apply these approaches to large cohorts of paediatric patients with unsolved diagnoses, with the aim of enabling the deployment of genomics technologies into the clinic as quickly as possible.

1640 – 1705

GENE DISCOVERIES IN EPILEPSY AND ITS CO-MORBIDITIES USING EXOME SEQUENCING

Keynote Speaker – Assoc Prof Leanne Dibbens
SCHOOL OF PHARMACY AND MEDICAL SCIENCES, UNIVERSITY OF SOUTH AUSTRALIA, AUSTRALIA

BIOGRAPHY

Associate Professor Leanne Dibbens is currently head of the Epilepsy Research Program at the University of South Australia and the Sansom Institute for Health Research. Her group, together with clinical collaborators in Melbourne, have discovered a number of the major genes which cause genetic forms of epilepsy and its comorbidities, including intellectual disability and autism. Gene discoveries include PROTOCADHERIN 19 in a form of epilepsy and mental retardation limited to females and genes causing progressive myoclonus epilepsies including lysosomal related genes such as SCARB2. Most recent gene discoveries include the potassium channel gene KCNT1 in an autosomal dominant form of focal epilepsy which shows co-morbidity with intellectual disability and psychiatric features. Identification of DEPDC5 mutations as a common cause of familial focal epilepsy has changed the way these forms of epilepsy are considered ie. that genetic factors play an increasing role in their etiology. These gene discoveries have led to the implementation of molecular diagnostic tests for different forms of epilepsy and are important in guiding the prognosis and genetic counselling of patients. Identifying new biological pathways in the genesis of seizures also provides new targets for improved therapies.

Leanne's early training was as a molecular and developmental geneticist. Her Honours and PhD studies in Drosophila were carried out in the departments of Biochemistry and Genetics at the University of Adelaide. She then became interested in human molecular genetics and spent several years at the Womens' and Childrens' Hospital in Adelaide before moving her group to UniSA in 2011.

ABSTRACT

Epilepsy is a common disorder comprising more than 30 different syndromes. Around 70% of epilepsy cases are thought to be genetic in origin. The genetic and phenotypic heterogeneity seen in epilepsy can make it difficult to diagnose and to treat. By studying rare large families with monogenic forms of epilepsy we have been able to carry out linkage analysis to identify a chromosomal region harbouring the causal variant in a family. By exome sequencing affected individuals and extracting the unique variants from within the linkage interval we have been able to chase down the putative causal variant. The identification of independent mutations in the same gene in patients who are phenotypically similar validates our identification of the casual gene. Using this approach we have recently identified two new genes in autosomal dominant forms of focal epilepsy, a severe form of nocturnal frontal love epilepsy (NFLE) with intellectual disability and psychiatric features and a syndrome known as FFEVF. The findings will lead to improved molecular diagnostic tests in epilepsy and reveal new pathways involved in the pathogenesis of epilepsy.

1705 – 1720

GENETIC VARIATION NEAR GRB2 AND KCNB2 IDENTIFIED BY A GENOME-WIDE ASSOCIATION STUDY ARE REPRODUCIBLY ASSOCIATED WITH DIABETIC RETINOPATHY

Contributed Paper – Dr Kathryn Burdon
FLINDERS UNIVERSITY, AUSTRALIA

BIOGRAPHY

Dr Kathryn Burdon is a molecular genetiscist based in the Department of Ophthalmology at Flinders University in Adelaide. Her work is focused on the identification and characterisation of genes for blinding diseases including both Mendelian disorders and complex diseases. She uses a variety of techniques to map disease genes, including both family studies genome-wide association studies in large case-control cohorts. Her talk today will describe the results from an NHMRC funded genome-wide association scan for diabetic retinopathy.

ABSTRACT

Purpose: Diabetic Retinopathy (DR) is a potentially blinding diabetes complication with evidence for genetic determinants. We conducted a well powered genome-wide association study (GWAS) to identify susceptibility loci for DR, focusing on sight-threatening disease.

Method: 2120 Caucasians with type 2 diabetes (T2D) were recruited from multiple hospitals and assessed for DR. A GWAS using the OmniExpress array (Illumina) was conducted in the first 1090 samples collected. Following data cleaning, association analysis was performed for any DR, sight-threatening DR (proliferative, severe non-proliferative or significant macular edema), plus each DR subtype alone, compared to individuals without DR. Top ranking SNPs were genotyped in the remainder of the cohort. Two replicated SNPs were typed in further replication cohorts from Aravind Eye Hospital, India and an Australian type 1 diabetes (T1D) cohort.

Results: SNP rs783992 near KCNB2 (voltage gated potassium channel, Kv2.2), was associated with any DR (p=2.01x10-8) and proliferative DR (p=7.5x10-8) and also showed replication in the second half of the initial cohort (p=0.01), but not in the Indian or T1D cohorts. SNP rs9896052 near GRB2 (Growth factor receptor-bound protein 2, an adaptor protein involved in signal transduction/cell communication) was associated with sight-threatening DR (p=4.71x10-8) and macular edema (p=1.0x10-6) and showed replication in the Caucasian (p=0.035), the Indian (p=0.005) and the T1D (p=0.022) cohorts.

Conclusion: SNPs rs783992 and rs9896052 are reproducibly associated with DR in Caucasians with T2D and rs9896052 is also more broadly associated in other populations. This is the first GWAS for DR to provide genome-wide significant association with replication.

1720 – 1735

LIONS AND TIGERS AND SPERM TAILS? OH MY! MAKING SENSE OF VARIANT DIVERSITY FROM WHOLE-GENOME SEQUENCING

Contributed Paper – Dr Jac Charlesworth

MENZIES RESEARCH INSTITUTE,
UNIVERSITY OF TASMANIA, AUSTRALIA

BIOGRAPHY

Dr Charlesworth is a statistical geneticist with experience in complex disease genetics and the analysis of a wide range of genomic data and complex phenotypes. She received her PhD from the University of Tasmania in 2006. She then completed a four year postdoctoral fellowship in statistical and computational genetics at Texas Biomedical Research Institute USA. She is now a Research Fellow at the Menzies Research Institute Tasmania. Dr Charlesworth's analytical expertise includes the analysis of linkage, association, expression and next-generation sequence data; in particular data derived from extended pedigrees with complex phenotypes.

ABSTRACT

Complex disease genetics has recently made a rapid transition from the paucity of functionally relevant signals seen in the early GWAS to the overwhelming number of apparently functional coding variants detected by whole-genome sequencing. We have 50x whole-genome sequencing and detailed neuroimaging data available for 365 individuals from extended pedigrees, as part of the US based Genetics of Brain Structure study, a spin-off from the long running San Antonio Family Heart Study. We are utilising this resource to investigate the underlying genetic architecture of highly heritable and disease relevant neuroimaging traits such as deep white matter lesions. To deal with the sheer volume of genetic variation we defined a set of 'highly deleterious non-synonymous' variants that had PolyPhen2 scores >0.8 and were present in at least 5 individuals in the pedigrees. A surprising number of variants (28,990) passed these thresholds. We then used standard variance components based measured genotype association testing in SOLAR to determine the influence of these variants on our traits of interest. The results are surprising and can be just as difficult to interpret. For example we identified an uncommon splice variant (16 copies) in *SPEF2,* encoding sperm flagellar 2, that was significantly associated with a decrease in white matter lesion volume. This gene was a previous GWAS hit that we overlooked based on the gene name, but now leaves us asking questions like "could a sperm tail gene be relevant in brain lesion pathology?"

DAY 3 TUESDAY 15 OCTOBER

SESSION 5
RNA/NONCODING

CHAIR: DR RUBY LIN AND DR MARK COWLEY

0830 – 0910

THE SECRET LIVES OF STEM CELLS

Plenary Speaker – Prof Jeanne F. Loring
THE SCRIPPS RESEARCH INSTITUTE, USA

Sponsored by

STEMCELLS AUSTRALIA

BIOGRAPHY

Jeanne Loring, Ph.D. is a Professor and the founding Director of the Center for Regenerative Medicine at The Scripps Research Institute in La Jolla. Her research team primarily focuses on large-scale analysis of the genomics and epigenomics of human pluripotent stem cells (hPSCs) and their derivatives. Their translational projects include development of stem cell applications for multiple sclerosis, Alzheimer disease, Parkinson's disease, and Fragile X and Rett syndromes, and they are creating collections of iPSC lines for disease modeling and ethnicity associated drug toxicity. Dr. Loring is also involved in the societal issues associated with stem cell research, including the ethics of stem cell generation and clinical use, the legal implications of stem cell patents, and public education about the dangers of unregulated stem cell treatments ("stem cell tourism"). Dr. Loring serves on the ethics boards of the Bill and Melinda Gates Foundation and Merck KGaA.

ABSTRACT

Human pluripotent stem cells are remarkable for their abilities to self-renew indefinitely and differentiate into virtually every cell type. In the last 15 years since human embryonic stem cell (hESC) lines were first reported, there has been tremendous progress toward development of hESCs and induced pluripotent stem cells (iPSCs) for cell replacement therapy and for modeling human disease. With the explosion of excitement and the rush to the clinic, careful examination of the cells has been overlooked; there is a disconnect between the hopes and fears for therapy and the realities of the cells themselves. What are these cells really doing in their dishes as they proliferate and differentiate? Our work shows that in the dark, in their incubators, hESCs and hiPSCs are undergoing dramatic changes. The culture dish is a hotbed of natural selection, with cells with duplications and deletions and epigenetic changes gaining an advantage, and gradually, without us noticing it, taking over the dish. We have documented these changes in hundreds of hPSC lines; some changes make sense and others are quite surprising. We are using this information to understand what stem cells are doing under cover of darkness, and what it means for the future of clinical applications.

0910 – 0935

POST-TRANSCRIPTIONAL REGULATION OF MITOCHONDRIAL GENE EXPRESSION

Keynote Speaker – Prof Aleksandra Filipovska
WESTERN AUSTRALIAN INSTITUTE FOR MEDICAL RESEARCH, AUSTRALIA

BIOGRAPHY

Aleksandra Filipovska received her PhD in 2002 from the University of Otago, New Zealand. From 2003-2005 she was a NZ Foundation for Research, Science and Technology Fellow at the MRC Mitochondrial Biology Unit in Cambridge, the United Kingdom. In 2006 she relocated to Australia as a NHMRC Howard Florey Fellow and established her research group at the Centre for Medical Research at the University of Western Australia in Perth. She is currently an Australian Research Council Future Fellow at the University of Western Australia and a group leader at the Western Australian Institute for Medical Research. Her research efforts to date have been recognised with the award of several prizes such as the Australian Academy of Sciences Ruth Stephens Gani Medal, ANZSCDB Young Investigator award and the WA Tall Poppy Award.

Aleksandra has published in internationally journals such as Cell, Nature Chemical Biology, Journal of Cell Biology and Journal of Biological Chemistry. Her research has been featured on the cover of Cell, Metallomics and Nature Chemical Biology, and it has been highlighted in Nature Biotechnology, Nature Chemical Biology and Cell Cycle.

Her research focuses on the study of mitochondrial gene expression and function in health and disease, and the development of methods for studying and developing therapeutics for inherited mitochondrial diseases.

ABSTRACT

Human mitochondria contain a small and compact genome that is transcribed as long polycistronic transcripts that encompass each strand of the genome, which are processed into mature mRNAs, tRNAs and rRNAs within the mitochondrial matrix. Recently we provided the first comprehensive map of the human mitochondrial transcriptome by near-exhaustive deep sequencing of long and small RNA fractions from purified mitochondria. We have identified previously undescribed transcripts, including small RNA and long non-coding RNAs encoded by the mitochondrial genome. Furthermore despite their common polycistronic origin, we observed wide variation between individual tRNAs, mRNAs, and rRNA amounts, indicating the importance of RNA-binding proteins in the regulation of mitochondrial gene expression. We have investigated the roles of the mammalian pentatricopeptide repeat (PPR) proteins and found that these RNA-binding proteins are all localized to mitochondria where they regulate mitochondrial gene expression. Mammalian PPR proteins have diverse roles in RNA metabolism and translation that are important for mitochondrial function and cell health. To investigate the importance of RNA-binding proteins in mitochondria globally we have established new methods for massively parallel sequencing and analyses of RNase-accessible regions of human mitochondrial RNA. We have identified specific regions within mitochondrial transcripts that are bound by RNA-binding proteins. These mitochondrial protein footprints indicate that RNA-binding proteins as well as small RNAs play a significant role in the regulation of mitochondrial gene expression.

0935 – 1000

FUNCTIONAL ROLE OF RNA-PROTEIN INTERACTIONS IN CELLS

Keynote Speaker – Dr Minna-Liisa Änkö
WALTER AND ELIZA HALL INSTITUTE OF MEDICAL RESEARCH, AUSTRALIA

BIOGRAPHY

Minni (Minna-Liisa) Änkö is a senior research fellow at the Walter and Eliza Hall Institute of Medical Research (WEHI) in Melbourne. She obtained her PhD in biology from the Åbo Akademi University in Finland on a scholarship from the Finnish Graduate School of Neuroscience. From 2006-2011 Minni was a post-doctoral fellow at the Max Planck Institute of Cell Biology and Genetics, in Dresden, Germany. She received the Sigrid Juselius fellowship for her postdoctoral work in the laboratory of Dr Karla Neugebauer. In 2011 Minni received the Jane and Aatos Erkko fellowship and was a visiting scientist at the Australian Regenerative Medicine Institute, Monash University, Melbourne before moving to her current position at WEHI. Minni's work has been published in world leading journals including Molecular Cell, Nature Structural & Molecular Biology and Genome Biology. She has contributed invited articles to Molecular Cell, Trends in Biochemical Sciences and Handbook of RNA Biochemistry. She is an active member of the RNA Society, Australian Society of Biochemistry and Molecular Biology, Victorian RNA Network and Australasian Society for Stem Cell Research. Minni's current work focuses in understanding the functions of SR protein splicing factors during early development, in stem cells and cancer.

ABSTRACT

Pre-mRNA splicing is a key process involving almost all mammalian genes. The spliceosome is the catalytic molecular machine but additional splicing factors are required for the correct recognition of the exon-intron boundaries. Furthermore, alternative splicing increases the mRNA and protein repertoire and provides a source for qualitative differences between cells because the vast majority of alternatively spliced transcripts are expressed only in specific cells, or at specific times during development. SR proteins comprise a family of essential splicing factors which play numerous roles in gene expression. How SR proteins coordinate gene expression programs in cells is poorly understood and comprehensive knowledge of their RNA targets has been lacking. We have used various genome-

wide approaches to identify SR protein binding sites and regulated genes. Our RNA immunoprecipitation experiments with SRSF3 and SRSF4 followed by microarray analysis revealed that these SR proteins are components of mRNPs containing distinct, functionally-related groups of transcripts. To identify exact RNA binding sites of SRSF3 and SRSF4 in cells, we used UV cross-linking and immunoprecipitation (CLIP) coupled with RNA sequencing. We determined distinct in vivo consensus binding motifs and detected interactions with intronless and intron-containing mRNAs, and non-coding RNAs. Functional studies of candidate targets revealed unexpected regulatory roles of SR proteins. Our studies have shown that SR proteins contribute to the control of gene expression

programs in cells by acting as positive and negative regulators of splicing but also by executing functions beyond splicing. The identification of SR protein gene targets allows us now to establish connections between SR proteins and the biological processes they regulate. We are currently investigating SR protein functions in vivo using mouse models, especially focusing on their roles in the regulation of pluripotency, differentiation and early development.

1000 – 1015

ASSESSING THE TARGET SPECIFICITY OF TRIPLEX FORMING OLIGONUCLEOTIDES USING DNA MICROARRAYS

Contributed Paper – Dr Denis Bauer
CSIRO, AUSTRALIA

BIOGRAPHY

Dr. Bauer is interested in high performance compute systems for integrating large volumes of data to inform strategic interventions in human health, livestock breeding and biosecurity.

She has a BSc(Hons) and PhD (2010) in Bioinformatics (Germany, UQ) and Post-Docs in machine learning and genetics (IMB, QBI). She published in high impact factor journals (Nature Genetics and Genome Research) and her 15 peer-reviewed publications (7 first, 1 last author) have been cited more than 200 times (h-index 5). She was invited speaker at RIKENs FANTOM meeting (2006) and Bio-IT World Asia conference (2013) and attracted more than A$120,000 in funding.

ABSTRACT

Nucleic acid triple helix-based genomic addressing systems have the potential to offer high genome-wide binding accuracy, while keeping the involved molecular machinery to a minimum, making triplexes an attractive alternative to zinc finger nucleases or CRISPRs in genomic engineering applications.

To quantify the DNA binding profile of short DNA or RNA Triplex-Forming Oligonucleotides (TFO) under physiological conditions, we developed a 2-stage high-throughput assay utilizing a custom DNA microarray design with all possible double stranded DNA targets for a specific TFO. We first convert the single stranded microarray probes to double stranded TFO-target sites by means of primer extension (reproducibility cc=0.95). We then incubate the TFOs and measure their binding potential to the dsDNA targets by means of fluorescence.

We observe that binding signatures for triplex-helices are different from those of RNA-DNA hybrids or G-quadruplexes (two orders of magnitude in $\log_2$ signal intensity). While the binding signature of the DNA-TFO is consistent with the expected antiparallel triplex, the observed pattern for the RNA-TFO suggests a different binding mechanism, possibly involving strand-invasion. From the DNA-TFO we find that target length (8 bp *versus* 20bp; *p*-value 8e−16) and nucleotide composition (pyrimidine interruptions) towards the center influence binding, hence offering a way to fine tune affinity and specificity of the TFO.

Our method of systematically quantifying binding preferences and assessing ligand-design features for reducing off-target risk provides a first step toward an accurate addressing system for genomic engineering applications.

1015 – 1030

LOC134466: A NOVEL TRANSCRIPT REGULATING CELL POLARITY THAT IS FREQUENTLY HYPERMETHYLATED IN TYPE II EPITHELIAL OVARIAN CANCER

Contributed Paper – Dr Brian Gloss
GARVAN INSTITUTE, AUSTRALIA

BIOGRAPHY

Brian Gloss is an early postdoctoral researcher currently working for associate professor Marcel Dinger at the new center for clinical genomics in the garvan

institute. His interest is on noncoding RNA and how it is regulated and how it regulates cellular processes, particularly in cancer.

ABSTRACT

Rationale: We have previously identified the novel locus *LOC134466* (a lincRNA, also known as *ZNF300P1*) as frequently silenced in ovarian cancer as a result of DNA methylation.

Objective: We sought to investigate the role of the epigenetic repression og *LOC134466* expression in cancer.

Methods and Results: In this study we demonstrate that *LOC134466* methylation is also a feature of breast and colorectal cancers. Furthermore, increased methylation is observed in breast cancer metastases versus primary tumors. Further evaluation implies that *LOC134466* is specifically targeted for epigenetic deregulation by DNA methylation in ovarian cancer rather than being silenced as part of a broader epigenetic defect. Transcriptional analysis of siRNA-mediated knockdown of *LOC134466* in human ovarian surface epithelial cells has revealed alterations in key cell cycle and cell motility networks.

Using siRNA to examine the effect of *LOC134466* suppression, we demonstrate a decrease in cell proliferation and colony formation with loss of *LOC134466*, indicating that the lincRNA does not act as a typical tumor suppressor. However, knockdown of the *LOC134466* transcript results in less persistent migration in wound-healing assays, as a result of a loss of cellular polarity. Using an ex-vivo peritoneal adhesion assay, we also identify a potential role for *LOC134466* in metastasis of ovarian cancer cells by enhancing adhesion to the peritoneal membranes.

Conclusions: These findings suggest that *LOC134466* likely serves as a tumor suppressor, aberrantly silenced in a high proportion of epithelial ovarian cancers and potentially crucial to the spread of cancer cells from their primary site.

SESSION 6
PROTEOMICS

CHAIR: PROF KIRILL ALEXANDROV

1100 – 1120

IN VITRO RECONSTITUTION AND ANALYSIS PROTEIN INTERACTION NETWORKS

Keynote Speaker – Prof Kirill Alexandrov
INSTITUTE FOR MOLECULAR BIOSCIENCE,
THE UNIVERSITY OF QUEENSLAND, AUSTRALIA

BIOGRAPHY

Professor Kirill Alexandrov obtained his Masters degree in Invertebrate Zoology at the Leningrad State University, Russia in 1989 and completed his Ph.D in Cell Biology at EMBL Heidelberg, Germany in 1995. He went on to postgraduate work at the Department of Physical Biochemistry at the Max-Planck Institute in Dortmund, Germany, and remained with the Institute for 12 years, becoming a group leader in 1999. He co-founded the German biotechnology company JenaBioscience in 1998. He joined the Institute for Molecular Bioscience and the Australian Institute for Bioengineering and Biotechnology of the University of Queensland, Australia in 2008 as an ARC Future Fellow.

His research interests are in the development of platform technologies for production, evolution and analysis of proteins, structure and function of macromolecular protein complexes, design of protein biosensors and gene expression in Leishmania protozoa.

ABSTRACT

The ability to replicate the building blocks of life such as DNA and proteins represents the core technologies of Life sciences. The exponential increase in the number of sequenced genomes has focused attention on how best to produce, study and modify the encoded gene products. While the structural and functional information is encoded in a single protein molecule, accessing this information is technically and economically challenging. This constitutes a critical technological bottleneck that determines the pace of progress in many areas of biology. The problems become particularly aggravating when analysis of complex protein machines is attempted. To address these issues we developed a pipeline for rapid production of recombinant proteins their interaction analysis in high throughput format. To

achieve that we created a suite of in vitro expression vectors amenable to Gateway cloning. We used these vectors express genes from human ORFeome libraries in recently developed Leishmania tarentolae cell-free system. We then combined single molecule fluorescence spectroscopy and Amplified Luminescent Proximity Homogeneous Assay Screen to analyse the interactions of subunits of several in vitro reconstituted multisubunit eukaryotic protein complexes such as Cavin, HOPS and Mediator that were previously refractory to reconstitution and analysis. Reconstitution of the complexes and analysis of subunit interactions could be carried out within hours. The approach allows rapid extraction of information on oligomerisation state, interaction affinities and the stabilities of individual subunits and the entire complexes. The implications of the developed technology for protein research will be discussed.

1120 – 1140

PROTEOMICS IN NON-MODEL ORGANISMS: FROM PLANTS TO MAMMALS AND BACK AGAIN

Keynote Speaker – Dr Michelle Colgrave
CSIRO ANIMAL, FOOD AND HEALTH SCIENCES, AUSTRALIA

BIOGRAPHY

Dr Michelle Colgrave is a Senior Research Scientist (Proteomics) with CSIRO Animal, Food and Health Sciences, based at the Queensland Bioscience Precinct in Brisbane, Australia.

Working within CSIRO's Food Futures Flagship, she leads a number of proteomics research projects, complementing genomic efforts to benefit Australia's livestock and plant industries and improve human and animal health. Dr Michelle Colgrave is using mass spectrometry techniques to help identify and quantify key proteins that will benefit Australia's livestock and plant industries and improve human health.

ABSTRACT

A great deal of biological research has focused on model organisms, for example mouse, rat, fruit fly, nematodes, yeast and arabidopsis. As such much of the genomic research has been aimed at such organisms and is somewhat lacking in the field of agricultural sciences. However, numerous non-model plants are essential as food, feed or energy resources and likewise animals such as cattle and sheep have received limited attention. The power of proteomics, i.e. the high-throughput identification of protein products, can be lost in non-model species due to the lack of genomic information available or due to the sequence divergence to a related model organism. Nevertheless, proteomics has great potential in the study of non-model species. This presentation will focus on proteomic applications in the study of non-model plants and animals addressing the issues commonly encountered when initiating the proteome analysis of a non-model organism. I will discuss aspects such as sample preparation, data analysis and interpretation, protein identification via MS and automated database searching. I will specifically focus on our progress examining barley (gluten protein identification), abalone (evaluation of reproductive performance) and applications of proteomics to the dairy and beef industries.

1140 – 1200

GETTING GENOMICS AND PROTEOMICS DATA TO WORK TOGETHER – WHAT CAN THEY TELL US ABOUT EACH OTHER?

Keynote Speaker – Dr Jason Wong
LOWY CANCER RESEARCH CENTRE, UNIVERSITY OF NEW SOUTH WALES, AUSTRALIA

BIOGRAPHY

Jason Wong heads the Bioinformatics and Integrative Genomics group at the Lowy Cancer Research Centre and is Lecturer within the Prince of Wales Clinical School at the University of New South Wales. He received his B.Sc (Hons I), specialising in Bioinformatics from the University of Sydney in 2002. As an Oxford-Australia scholar, he went on to complete a D.Phil in Bioanalytical Chemistry at the University of Oxford, UK in 2007. This was followed by an Irish Government post-doctoral fellowship at the Conway Institute of Biomolecular & Biomedical Research, University College Dublin, specialising in Chemical Proteomics. In 2008 he returned to Sydney to take up his current position. Since 2008, he has been awarded a UNSW Vice-Chancellor's Post-doctoral Fellowship and is currently a Cancer Institute NSW Early Career Development Fellow. His current research is focused on the application of bioinformatics and proteomics to study gene regulation and function in cancer.

ABSTRACT

Genomics and proteomics are closely related fields of research. An understanding of one is generally required for the other, yet in many ways, the methods used to study the two cannot be more different. With the emergence of massive parallel sequencing vast quantities of genomics and transcriptomics data are being generated. At the same time, improvements in mass spectrometry technologies are enabling proteins to be identified with greater specificity and sensitivity. This now provides new opportunities to investigate ways to integrate genomics and proteomics data and understand how the two can complement each other to advance biological knowledge. Using HeLa cells as a model system, we have comprehensively examined the gene models derived from genomics and transcriptomics data and integrated these with proteomics and phosphoproteomics datasets. Reanalysis of proteomics data using HeLa specific gene models enable significant increases in the number of peptides/proteins to be identified, providing new insights into both the genome and proteome of HeLa cells. Technical challenges and methods required for integrating genomics and proteomics data will also be discussed. In summary, given that massive parallel sequencing data are now available for many popular cell lines in public data repositories, our study provides further support for the need and benefit of an integrative data analysis for both genome and proteome analysis.

1200 – 1215

USING MULTI-LEVEL OMICS DATA TO INFER CAUSAL RELATIONSHIPS BETWEEN CORRELATED TRANSCRIPTS AND METABOLITES

Contributed Paper – Anita Goldinger
THE UNIVERSITY OF QUEENSLAND, AUSTRALIA

BIOGRAPHY

Anita Goldinger is a PhD student working in the Visscher lab in Brisbane. Her research is focused on studying the genetic architecture driving gene expression and endophenotypes in healthy human populations.

ABSTRACT

One of the key objectives of omics research is the elucidation of complex interactions between genetic variants, gene expression, proteins and metabolites that drive phenotypic traits. The majority of genetic studies to date have focused on single transcripts and metabolites. Whilst these studies have been successful in identifying numerous regulatory variants they do not investigate the pervasive co-regulation found between transcripts and metabolic endophenotypes. Here we present the results of a study investigating nine highly-correlated modules composed of 90 probes within the Brisbane Systems Genetic Study (BSGS). BSGS is comprised of 862 individuals from 314 families, and contains multi-level omics data including high-density genotype data, whole blood transcriptomics, mass-spectrometry metabolite data and a wide-range of phenotypic traits. Using principal components to decompose these nine modules into common axes variation, we observed a considerable genetic component h^2 >0.5 driving a number of these axes with an average h^2 of 0.27 across all axes. These axes demonstrated association to several metabolic traits and higher-order phenotypes. Using eSNPs driving gene expression of probes within these modules we were able to identify SNPs driving correlated transcripts and associated metabolites. These joint associations provide novel biological interpretations of the underlying genetic architecture driving these transcripts and gives strong evidence of causal relationships. The dissection of this genomic architecture driving these multi – level omics data has important applications in functional annotations for genomic loci uncovered in GWAS studies and also in exploring the basis of disease susceptibility between individuals.

OPTIONAL LUNCHTIME PRESENTATION

1245 – 1315

SEQCAP EPI: ULTRA-HIGH COMPLEXITY PROBE POOLS FOR TARGETED BISULFITE SEQUENCING

Dr Jeff Jeddeloh
ROCHE APPLIED SCIENCE, USA

Sponsored by Roche

BIOGRAPHY

Jeff Jeddeloh is currently Director of Technology Innovation at Roche Applied Science, based in Madison, WI. In this role, Jeff identifies and secures innovative, disruptive technologies for the future life science portfolio of Roche. Prior to joining Roche, Jeff was the Science and Technology Director at Orien Genomics where he managed platform development and implementation for bio marker discovery and the epigenetics molecular diagnostics business. Jeff has a PhD in Molecular Genetics from Washington University in St. Louis.

ABSTRACT

DNA methylation has been shown to have a role in a host of biological processes, including silencing of transposable elements, stem cell differentiation, embryonic development, genomic imprinting, and inflammation, as well as many diseases, including cancer, cardiovascular disease, and neurologic diseases. Epigenetic modifications can also affect drug efficacy by modulating the expression of genes involved in the metabolism and distribution of drugs, as well as the expression of drug targets, contributing to variability in drug responses among individuals. There are currently a number of tools to study DNA methylation status, either at a single locus level, using methods like methylation-specific PCR or MALDI-TOF-MS, or at a broader, genome-wide level, using DNA microarrays, reduced representation bisulfite sequencing (RRBS), or even whole genome shotgun bisulfite sequencing. The latter method is preferred by many researchers, as it provides DNA methylation status at base pair resolution and allows for the assessment of percent methylation at each position in the genome. However, it is expensive, in terms of money and analysis, to generate such data for the entire genome, when generally only a subset of the genome is of interest to most researchers. We describe a system for the targeted enrichment of bisulfite treated DNA, allowing researchers to focus on a subset of the genome for high resolution cytosine methylation analysis. Regions ranging in size from 10 kb to 75 Mbp may be targeted, and multiple samples may be multiplexed and sequenced together to provide an inexpensive method of generating methylation data for a large number of samples in a high throughput fashion. Innovations in probe design, selection and manufacture, as well as optimization techniques to improve capture uniformity differentiate our system are highlighted.

SESSION 7
POPULATION VARIATION

CHAIR: DR CARSTEN KULHEIM

1330 – 1350

ANALYSING GENOMIC NETWORKS IN CANCER

Keynote Speaker – Assoc Prof Cristin Print
NEW ZEALAND BIOINFORMATICS INSTITUTE,
UNIVERSITY OF AUCKLAND, NEW ZEALAND

BIOGRAPHY

Cris graduated in Medicine from Auckland Medical School in 1989. He worked as a house surgeon and undertook asthma research in Dunedin before completing an immunology PhD at the University of Auckland. He researched leukocyte and germ cell apoptosis at the Walter and Elisa Hall Institute in Melbourne for four years with Suzanne Cory, before moving to Cambridge University, UK, where he was a fellow of St Edmund's College and researched endothelial cell and reproductive genomics. Since 2005 he has worked in Auckland University using bioinformatics to better understand cancer pathology, and is joint director of the university's Bioinformatics Institute.

ABSTRACT

The availability of cancer genomic data is rapidly expanding. Researchers can now easily access linked clinical, pathological, DNA and RNA data on thousands of individual patients with any one tumour type. Simultaneously, a growing number of targeted therapies are in clinical trial, with some of these already accompanied by genomic diagnostic assays for patient stratification. This talk will discuss our research group's use of bioinformatics to identify linked clinical and genomic tumour biomarkers and therapeutic stratification tools for breast cancer, colon cancer and melanoma.

1350 – 1410

MAKING SENSE OF NEXT GENERATION SEQUENCE DATA WITH GENERATIVE MODELS

Keynote Speaker – Dr Lachlan Coin

INSTITUTE FOR MOLECULAR BIOSCIENCE,
THE UNIVERSITY OF QUEENSLAND, AUSTRALIA

BIOGRAPHY

Lachlan Coin completed a Bachelor of Science degree at the Australian National University, and was awarded the University Medal in Mathematics, in 1999. He worked as a strategy consultant in 2000/2001 then returned to science to do a PhD in Bioinformatics at Cambridge University, based at the Sanger Institute. During his PhD Lachlan investigated patterns of protein domain evolution. After his PhD Lachlan worked on developing phylogenetic methods for the Treefam database at the Sanger Institute, and in 2005 moved to Imperial college to work as a research fellow in the Department of Epidemiology and Public Health. At Imperial college Lachlan worked on methodology for identifying genetic variation associated with disease phenotypes, focussing on structural variation. This work lead to development of new algorithms for identifying structural variation from high-throughput genotyping and sequencing data, as well as identifcation of structural variation associated with Obesity and lipid metabolism. Lachlan has recently begun to establish a new Population Genomics group within the division of genomics and bioinformatics at the Institute of Molecular Biosciences.

ABSTRACT

Sequencing of the first reference genomes for many species, including humans, was completed at great expense using automated Sanger sequencing to provide high quality long read DNA sequence. Sanger sequencing has now largely been superseded by next generation sequencing, which refers to a set of technologies and sample preparation techniques which have in common their ability to produce short read sequence data in a very high throughput manner at relatively low cost. This technology is ideally suited to re-sequence 100s of individuals from multiple populations in each species in order to provide a complete catalogue of genomic variation in that species, such as has being undertaken by the 1000 genomes sequencing consortium, and to search for trait associated variation not captured by genotyping arrays. However, next generation sequence data is noisy, with the rate of base-calling error exceeding the rate at which novel mutations occur. We have developed a suite of tools which model next generation sequence data at the population level for inference of multiple features, including: detection and genotyping of indels; detection and genotyping of copy number variation; association of copy number variation with phenotype from exome sequence data; identification of highly differentiated SNP and indel sites between populations; Y and mitochondrial haplotyping. I will present data which show that our methods have substantially greater sensitivity than other methods, at the same false discovery rate. Moreover our methods have been designed to work at arbitrary levels of ploidy, making them ideally suited to applications in crop science.

1410 – 1430

SINGLE CELL GENE EXPRESSION ANALYSIS TO REFINE CELL POPULATIONS IN MAMMARY EPITHELIUM HIERARCHY

Keynote Speaker – Dr Bhupinder Pal
THE WALTER AND ELIZA HALL INSTITUTE
OF MEDICAL RESEARCH, AUSTRALIA

Sponsored by

millennium science Fluidigm

BIOGRAPHY

Bhupinder Pal received his PhD degree from the University of Melbourne in 2008. He is currently a NHMRC Postdoctoral research fellow working in the Breast Cancer Research Laboratory at the Walter and Eliza Hall Institute of Medical Research, Melbourne. His research interests include molecular characterisation of mammary epithelial cells in mouse and human. Currently, he is studying mammary epithelium at single cell level to understand the role of individual epithelial cell types in the normal development of mammary gland.

ABSTRACT

Mouse and human mammary epithelium consists of functionally homologous cell types: adult mammary stem cell, luminal progenitor and mature luminal cells. These cell types are marked by unique gene expression signatures during different mammary developmental stages. Microarray gene expression profiling carried out on the epithelial populations prospectively isolated by FACS has led to the identification of many important molecular regulators that dictate self-renewal, lineage commitment and differentiation programs. However, accumulating evidence suggests that those isolated populations are most likely heterogeneous and the gene expression information gathered from large number of cells can mask a minor, but functionally important cell compartment within each population.

In this study, we used microfluidic based single-cell gene expression techniques on freshly isolated single epithelial cells. Our results have unmasked unique gene clusters within each test population, potentially marking novel cellular subsets. Expression patterns of proliferation, differentiation and hormone receptor genes were sufficient to identify subtle genetic differences between subsets and gain insight into pathways important for controlling their functions. Despite their heterogeneous nature, mouse and human epithelial populations show conserved gene expression at the single cell level.

1430 – 1445

IDENTIFICATION OF SINGLE NUCLEOTIDE POLYMORPHISMS (SNPS) INVOLVED IN THE DETERMINATION OF HUMAN PHYSICAL APPEARANCE

Contributed Paper – Mark Barash
BOND UNIVERSITY, AUSTRALIA

BIOGRAPHY

Before starting his PhD studies at Bond University, Mark enjoyed a successful career as a Forensic DNA Officer in the Israeli Police Force. He has been involved in solving many serious crime cases by the means of forensic DNA analysis methods. He has also published numerous articles in the forensic literature and participated in many international conferences. Mark moved to Australia to undertake a groundbreaking forensic research project, which aims to establish the genetic factors behind human facial appearance. In other words to understand why we look as we are? The results of this project will not only make a significant impact in forensic science, but also in other scientific disciplines, such as anthropology and medical diagnostics of craniofacial disorders.

ABSTRACT

Forensic Molecular Photofitting is a new area of forensic DNA profiling that seeks to obtain visible information, such as skin, eye and hair pigmentation and more recently, facial morphology, about the depositor of a DNA sample.

This study aimed to identify a set of single nucleotide polymorphisms (SNPs) involved in normal human craniofacial variation and develop an assay that can predict the phenotype of the person who is the source of the DNA. To achieve this, more than 1200 SNPs in 173 candidate genes, potentially involved in normal embryonic craniofacial development and various malformations were targeted for genotyping. In addition, approximately 700 markers previously shown to be associated with pigmentation traits such as eye, skin and hair colour, along with ancestry lineage markers and identity informative SNPs, STRs and INDELs were chosen for genotyping.

More than 500 DNA samples along with pigmentation phenotype, ancestry information and 3-Dimentional (3D) facial images have been collected. The 3D images were analysed for more than 90 linear and angular craniofacial measurements, including various craniofacial indexes.

The DNA samples were sequenced using a Next Generation Sequencing platform at a set of candidate SNPs and were evaluated for statistically significant associations with pigmentation, ancestry and anthropometric craniofacial measurements.

The results of this study will be presented.

1445 – 1500

DETECTION AND REPLICATION OF EPISTASIS INFLUENCING TRANSCRIPTION IN HUMANS

Contributed Paper – Dr Joseph Powell
THE UNIVERSITY OF QUEENSLAND, AUSTRALIA

BIOGRAPHY

Joseph moved to Brisbane early in 2010 after a stint in Edinburgh where he did his postgraduate degrees. He has been fortunate to work on a range of projects involving methods, theory and application around the nexus of quantitative, statistical and population genetics. This has provided a good foundation for his more recent research investigating the genetic architecture regulating gene expression and its role within a systems genetics framework. Most of his current work involves the investigation of genetic (co) variance oftranscriptlevels both within and between different tissues.

ABSTRACT

Epistasis is the phenomenon whereby one polymorphism's effect on a trait depends on other polymorphisms present in the genome. The extent to which epistasis influences complex traits and contributes to their variation is a fundamental question in evolution and human genetics. Though epistasis has been demonstrated in artificial gene manipulation studies in model organisms, and examples have been reported in other species, few convincing examples with independent replication exist for epistasis amongst natural polymorphisms in human traits. Its absence from empirical findings may simply be due to its low incidence in the genetic control of complex traits, but an alternative view is that it has previously been too technically challenging to detect due to statistical power and computational issues. Here we show that, using advanced computation techniques and a gene expression study design, many instances of epistasis are found between common single nucleotide polymorphisms (SNPs). In a cohort of 846 individuals with data on 7339 gene expression levels in peripheral blood, we found 501 significant pairwise epistatic interactions between common SNPs acting on the expression levels of 238 genes ($p < 2.91 \times 10^{-16}$). Replication of these signals in two independent data sets showed both concordance of direction of epistatic effects ($p = 5.56 \times 10^{-31}$) and enrichment of interaction *p*-values, with 30 being significant at a conservative threshold of $p < 0.05 / 501$. There was evidence of functional enrichment for the interacting SNPs, for instance 44 of the genetic interactions are located within 5Mb of regions of known physical chromosome interactions ($p = 1.8 \times 10^{-10}$). Epistatic networks of three SNPs or more influence the expression levels of 129 genes, whereby one *cis*-acting SNP is modulated by several *trans*-acting SNPs. For example MBNL1 is influenced by an additive effect at rs13069559 which itself is masked by *trans*-SNPs on 14 different chromosomes, with nearly identical genotype-phenotype (GP) maps for each *cis-trans* interaction. This study presents the first evidence for multiple instances of natural genetic polymorphisms interacting to influence human traits.

SESSION 8
COMPUTATIONAL BIOLOGY

CHAIR: DR DENIS BAUER AND DR JAC CHARLESWORTH

1530 – 1610

BRIDGING THE GAP: ENABLING TOP RESEARCH IN TRANSLATIONAL RESEARCH

Plenary Speaker – Prof Knut Reinert
FREIE UNIVERSITÄT BERLIN, GERMANY

BIOGRAPHY

I am currently Professor for Algorithmic Bioinformatics at the Freie Universität Berlin in the department for Computer Science and Mathematics. In particular, I am interested in developing mathematical models for analyzing large genomic sequences, especially in the context of next generation sequencing (NGS), and data derived from mass spectrometry experiments, for example for detecting differential expression of proteins between normal and diseased samples.

ABSTRACT

In this talk I will convey to you my view about the necessary steps for enabling efficient research in biomedical research in the times where biotechnology can give us comprehensive views of certain data.

I will start by arguing that the NGS technologies developed in the recent years changed the research landscape to a degree similar to the beginning of the millennium when the human genome was initially sequenced.

As a consequence, the research tools of many biomedical researcher have or will change in the sense that they will conduct large scale, complex computations. Hence, as a community, we have to turn our focus to how we develop such tools.

Thinking about this becomes essential since in the near future clinical decisions concerning the treatment of individuals (personalised medicine) will be based on such computations. I will talk about the past and future role of software libraries for enabling translational research and exemplify some points with the SeqAn C++ library developed in my lab.

1610 – 1635

RNA-SEQ: FROM READS TO GENES TO PATHWAYS

Keynote Speaker – Prof Gordon Smyth
WALTER AND ELIZA HALL INSTITUTE OF MEDICAL RESEARCH, AUSTRALIA

BIOGRAPHY

Professor Smyth is a statistical bioinformatician working on gene expression problems. His group develops novel statistical methodology, algorithms and software for the analysis of microarray, RNA-seq and ChIP-seq data. His group has developed some well known software packages including limma, edgeR, goseq and subread.

ABSTRACT

This talk will discuss our latest pipelines for RNA-seq read mapping and differential expression analysis. Until recently, most RNA-seq pipelines have made it difficult to access the rich range of tools previously developed for microarrays. This talk will discuss the voom pipeline, which gives access to some powerful and statistically rigorous tools for pathway and expression signature analysis.

1635 – 1700

DATA INTEGRATION OF HIGHLY DIMENSIONAL BIOLOGICAL DATA SETS WITH MULTIVARIATE ANALYSIS

Keynote Speaker – Dr Kim-Anh Lê Cao
QUEENSLAND FACILITY FOR ADVANCED BIOINFORMATICS, THE UNIVERSITY OF QUEENSLAND, AUSTRALIA

BIOGRAPHY

Dr Kim-Anh Lê Cao was awarded her Ph.D in 2008 in Université de Toulouse, France. She was awarded the "Marie-Jeanne Laurent-Duhamel" prize 2009 of the Société Française de Statistique (French Statistical Society) for her Ph.D thesis.

Since the beginning of her Ph.D she has initiated a wide range of valuable collaborative and research opportunities in both statistics and molecular biology. Her research interests are multidisciplinary as they focus on mathematical statistics characterization of molecular biological systems, and she is interested

in developing sound statistical frameworks applied to addressing new biological questions arising from these frontier molecular technologies.

Since 4 years she has been employed as a research only academic in the Queensland Facility for Advanced Bioinformatics (QFAB), in the Institute for Molecular Bioscience (University of Queensland).

ABSTRACT

Recent advances in high throughput 'omics' technologies enable quantitative measurements of expression or abundance of biological molecules of a whole biological system. The transcriptome, proteome and metabolome are dynamic entities, with the presence, abundance and function of each transcript, protein and metabolite being critically dependent on its temporal and spatial location.

Whilst single omics analyses are commonly performed to detect between-groups difference from either static or dynamic experiments, the integration or combination of multi-layer information is required to fully unravel the complexities of a biological system. Data integration relies on the currently accepted biological assumption that each functional level is related to each other. Therefore, considering all the biological entities (transcripts, proteins, metabolites) as part of a whole biological system is crucial to unravel the complexity of living organisms.

With many contributors and collaborators, we are currently establishing a global analytical framework to extract relevant information from high throughput `omics' platforms such as genomics, proteomics, metabolomics and other types of biological data. Specifically, the statistical methodologies that we developed and implemented in our R package mixOmics focus on the so-called multivariate projection-based approaches, which can handle such large data sets, deal with multi collinearity and missing values. These methodologies enable dimension reduction by projecting these large data sets into a smaller subspace, to capture the largest sources of variation in the biological studies.

In this presentation, I will illustrate how various techniques enable exploration, data integration and visualisation of the data, with a particular focus on time course experiments analysis.

1700 – 1715

BLUE: FAST, ACCURATE ERROR CORRECTION USING K-MER CONSENSUS AND CONTEXT

Contributed Paper – Paul Greenfield
CSIRO COMPUTATIONAL INFORMATICS, AUSTRALIA

BIOGRAPHY

Paul Greenfield is a bioinformatician (coming from Computer Science) working for CSIRO Computational Informatics. Paul works mostly on metagenomic sequence data analysis, in both environmental and human gut studies. Much of this work makes use of k-mer tools, including Blue, and large bacterial k-mer databases.

ABSTRACT

Sequence data is now cheap to obtain, and becoming cheaper all the time. The only problem is that this data contains errors, some random, some more systematic. Commonly used tools, such as aligners and assemblers, have been written to cope with these errors in various ways, such as looking for consensus or doing error-tolerant string matching. Another way of dealing with sequencing errors is to run the reads through an error-corrector early in the processing pipeline. There have been a number of published error-correction algorithms but none of these have been yet widely adopted by the bioinformatics community. This presentation looks at Blue – a k-mer correction tool based on consensus and correctness – and comparing its effectiveness and performance to a number of the other available algorithms.

Our test results show that Blue was the only tool that was accurate, effective and fast enough to be a practical component in a processing pipeline. Using Blue resulted in much improved assemblies (longer contigs containing fewer errors), and many more reads aligning with no mismatches (typically over 99% in our tests). The other algorithms tested varied considerably in their effectiveness (from 'not very' to 'useful'), and were also often an order of magnitude slower and more memory-demanding. Blue can also cross-correct 454/Ion Torrent data with a consensus derived from Illumina reads, and this has capability has been used within CSIRO to produce better quality assemblies for both microbial and insect genomes.

1715 – 1730

BUILDING TISSUE-SPECIFIC PROMOTER-ENHANCER REGULATORY MAPS

Contributed Paper – Timmothy O'Connor
INSTITUTE FOR MOLECULAR BIOSCIENCES, THE UNIVERSITY OF QUEENSLAND, AUSTRALIA

BIOGRAPHY

Tim O'Connor is a first-year PhD student in the IMB at the University of Queensland. He has previously studied bioinformatics at Washington State University, Temple University, and Princeton University. Prior to becoming a student at the UQ he worked as a software developer at Microsoft for five years integrating technologies from Microsoft Research into consumer products. His current research is focussed on mechanisms of gene regulations, particularly the identification of long-distance regulatory interactions in the genome.

ABSTRACT

Transcription factor ChIP-seq data present a wealth of binding events but a dearth of potential regulatory targets of many binding events due to the large genomic distances between them and potential gene targets. Knowledge of tissue-specific distal regulatory regions, or enhancers, and the set of gene promoter targets these regions regulate, would provide putative gene targets for transcription factors (TFs) that bind in these enhancer regions. A method in the ENCODE project proposes a promoter-enhancer regulatory mapping mechanism using high correlation between histone modifications at putative enhancers and transcription at distal promoters (up to 1Mbp away) across a set of tissues to identify regulatory relationships. We improve this mapping methodology to generate tissue-specific promoter-enhancer maps. We rigorously demonstrate that our correlation-based maps identify novel, distal TF regulatory relationships. We do this by showing that models that predict gene expression from TF binding data (ChIP-seq peaks) are more accurate when based on our promoter-enhancer maps than when based on the commonly used nearest-neighbor heuristic. Our promoter-enhancer models are also far more accurate than promoter-only models of gene expression, demonstrating that our maps capture important regulatory relationships.

DAY 4 WEDNESDAY 16 OCTOBER

SESSION 9
APPLIED GENOMICS

CHAIR: DR BROOKE GARDINER AND LIAM WILLIAMS

0910 – 0950

HIGH RESOLUTION TRANSCRIPTOME ANALYSIS – ONE CELL AT A TIME

Plenary Speaker – Dr Jian-Bing Fan

SENIOR DIRECTOR OF SCIENTIFIC RESEARCH, ILLUMINA

Sponsored by

illumina®

BIOGRAPHY

Jian-Bing Fan, Senior Director of Scientific Research, Illumina, Inc., has over 20 years of experience in human genome research and microarray/next-gen sequencing technology development. He played important roles in developing the BeadArray technology and a series of genomics products at Illumina for high throughput genotyping, gene expression and methylation analysis. Currently, Dr. Fan is leading the Illumina Dx research efforts to develop highly sensitive and quantitative assays for use with medical specimens such as archived tissue samples, blood samples, circulating tumor cells and cell-free DNA. He has applied all these technologies in a broad spectrum of clinical applications such as non-invasive prenatal testing (NIPT), preimplantation genetic diagnosis and screening (PGD/PGS), and cancer profiling and early detection. Prior to joining Illumina, Dr. Fan was the Manager of Polymorphism Research at Affymetrix. While at Affymetrix, he spearheaded the development of high-density oligonucleotide array technology and its application in large-scale SNP identification and genotyping. Before that, Dr. Fan was a senior research fellow at UCSF/Stanford Human Genome Center. He was a key scientist involved in all aspects of the Center's scientific operation, such as large-scale physical mapping and sequencing of the human genome. Dr. Fan received his PhD in Genetics from Columbia University. He has authored more than 100 original articles and book chapters, published one book, been granted 18 US/world patents, and served on the editorial board of the international journal Genome Research from 2004 to 2007 and is currently an associate editor of Genomics, and Editor-in-Chief, Genomics Data.

ABSTRACT

Gene function and regulation inside mammalian cells occurs spatially and temporally within the context of local microenvironment. Each individual cell is at a particular expression stage of gene activities which defines specific cellular functions/phenotypes such as cell growth, proliferation, and interactions with other cells. A comprehensive molecular characterization of individual cells will help uncover the structure and dynamics of the cell lineage tree within a tissue/organ, in health and in disease, thus leading to a leapfrog advance in biology and medicine.

This talk will focus on some of the recent development of single cell transcriptome methodologies and their applications in cancer and stem cell research. The criteria for effective single-cell transcriptome analysis are (1) to be able to measure gene expression reliably and (2) to be able to profile a large number of individual cells cost-effectively. This talk will also discuss efforts toward the development of novel *in-situ* sequencing platforms that could carry out targeted expression analysis of 100s to 1000s of genes in millions of individual cells simultaneously, in either the tissue at a spatial resolution of single cell or a heterogeneous cell population in tissue culture.

0950 – 1015

CLINICAL IMPLEMENTATION OF NEXT-GENERATION SEQUENCING FOR DIAGNOSTICS

Keynote Speaker – Dr Karin Kassahn

GENETICS AND MOLECULAR PATHOLOGY,
SA PATHOLOGY, AUSTRALIA

BIOGRAPHY

Karin Kassahn is a medical scientist with 10 years research experience in biological and biomedical sciences. She obtained her PhD in marine genomics from James Cook University and then pursued post-

doctoral positions at the Institute for Molecular Bioscience, UQ, in comparative genomics, bioinformatics and medical genomics. Since 2009 she has been an active member of the International Cancer Genome Consortium where her research has focused on somatic mutation detection from next-generation sequencing. In 2013 she took up the position of Head of Technology Advancement, SA Pathology to implement NGS in clinical diagnostics and patient management. She is an active member of the ReFuGe Consortium to sequence coral communities for better reef management.

ABSTRACT

Advances in next-generation sequencing are enabling the use of genomics in the clinical management of patients. As the clinical applications of next-generation sequencing expand, clinical validation studies are becoming increasingly important. The challenges and successes of applying these novel technologies in a clinical setting are enormous. Over the past 6 months we have been working towards the implementation of targeted gene panels to stratify cancer patients for targeted therapies and the use of comprehensive gene panels and exome sequencing for the diagnosis of complex, inherited disorders. We are evaluating sample quality from various clinical lung biopsy methods and evaluate their suitability for molecular tumor profiling. We share some of the successes of translating these technologies to diagnostics and outline the remaining challenges. These include the establishment of standardised bioinformatics pipelines for variant calling and annotation, the clinical interpretation and reporting of genomic data, data management within established health IT systems, the ethical implications of offering more comprehensive gene testing than has ever been possible before, and the dynamic relationship between research and diagnostics. We discuss some of the approaches and solutions we have considered. The increasing demand for genomics in health management is challenging current funding models for genetic testing in Australia, but the potential benefits for patients are enormous and are driving implementation of these technologies across the country. In the years to come, genetic diagnosis enabled by NGS will likely become increasingly important in the clinic.

1015 – 1030

DETECTION OF MARKERS THAT DISTINGUISH HUMAN FETAL AND MATERNAL DERIVED MESENCHYMAL STEM/STROMAL CELLS (MSCS)

Contributed Paper – Dr Celena Heazlewood
AUSTRALIAN INSTITUTE FOR BIOENGINEERING AND NANOTECHNOLOGY, AUSTRALIA

BIOGRAPHY

Celena was awarded her PhD in June 2013. She has investigated the placenta as a source for mesenchymal stromal cells (MSCs) for cellular therapy. She found two distinct populations of fetal and maternal MSCs and identified genetic profiles that may be used to distinguish the two populations. Celena is currently a postdoctoral research fellow at the Australian Institute for Bioengineering and Nanotechnology where she is continuing her work on MSCs.

ABSTRACT

Mesenchymal stem/stromal cells are commonly isolated from post-partum placenta, which is composed of both fetal and maternal components. It is thought that fetal-derived cells offer superior properties for clinical application, but isolation of homogeneous fetal MSC populations is difficult without molecular markers that can distinguish them from maternal cells. We isolated fetal (amniotic membrane) and maternal (decidua) components and characterised them extensively using a range of phenotypic properties including genotype, cell surface expression of traditional MSC markers; mesodermal differentiation capacity; and their ability to suppress T cell alloproliferation. Concurrently, transcriptomic profiles were generated, with the aim to identify markers of each MSC population. We applied a modified differential expression analysis using the 'attract' pathway analysis tool to generate a list of coexpressed genes that provided the best discrimination between fetal and maternal cells. Comparison of the top ranked candidates to public MSC datasets using the YuGene cross-platform comparison tool in **www.stemformatics.org** provided an indication of the reproducibility of these candidates. The markers have been validated in freshly isolated placental cultures using SmartFlare RNA technology, and antibodies, where appropriate. We have successfully established markers that can distinguish fetal and maternal MSC populations within the term placenta, which will be valuable for storing both the mother's and child's MSC for potential stem cell therapeutic applications.

SESSION 10
EPIGENETICS

CHAIR: DR STACEY EDWARDS AND DR LUTZ KRAUSE

1100 – 1140

ENHANCER RNA BIOGENESIS AND AN ATLAS OF TRANSCRIBED ENHANCERS ACROSS THE HUMAN BODY

Plenary Speaker – Assoc Prof Albin Sandelin
THE BIOINFORMATICS CENTRE, DENMARK

BIOGRAPHY

Albin Sandelin took his PhD at Karolinska Institute, and continued as a postdoc at RIKEN Yokohama Institute, working on the FANTOM and ENCODE projects. In 2006, he was recruited as a group leader at Copenhagen University. He is now running a joint computational-experimental group focusing on promoter and enhancer discovery and characterization in fundamental biology and disease.

ABSTRACT

In higher organisms, cellular development and diversity is highly controlled by enhancers, which regulate the correct temporal and cell type-specific activation of gene expression. Despite their obvious importance for development and disease, the exact locations and target genes as well as mechanisms of enhancers are still poorly defined. Thus, there is an urgent need not only to identify enhancer locations, but their specific usage in the wealth of cells within the human body, their impact in regulation in healthy and diseased individuals, how enhancer can be linked to target genes and what sets enhancers apart from TSSs in terms of biological signals. Although several recent studies have substantially increased our understanding of the complex architectures of enhancers in mammalian genomes, many of these studies have been limited to a few cell lines.

Interestingly, enhancers may bind RNAPII that, in turn, may produce capped enhancer-associated RNAs, reflecting the activity of an enhancer in a given cell. In this talk, I will present new data on the biogenesis and degradation of these RNAs, and then show how we can use the he FANTOM5 panel of tissue and primary cell samples covering the vast majority of human tissues and cell types to define an atlas of active, in-vivo transcribed enhancers. Uniquely, this makes it possible to identify truly cell/tissue-specific active enhancers, and to compare regulatory programs between different cells and tissues at unprecedented depth. Comparison of enhancer usage over multiple tissues and cell types revealed that most enhancers are not exclusively used in one cell type or tissue. In those cases, however, where we find cell type-specific enhancers, these are enriched in motif signatures associated with known key regulators of the relevant cell type. Clustering samples by enhancer usage groups related tissues together while forming distinct clusters of fetal samples, which may reflect fetal-specific regulatory programs.

Our atlas of active enhancers across the human body may serve as a resource of unprecedented scale for future studies on enhancer regulation and the biogenesis of enhancer-associated RNAs.

1140 – 1200

THE ROLE OF MYST LYSINE ACETYLTRANSFERASES IN TRANSCRIPTIONAL REGULATION AND CHROMATIN CONFORMATION

Keynote Speaker – Dr Anne Voss
WALTER AND ELIZA HALL INSTITUTE OF MEDICAL RESEARCH, AUSTRALIA

BIOGRAPHY

Anne K. Voss completed her doctoral degree in Hannover, Germany. After post-doctoral positions at Cornell University, Ithaca, NY, USA and at the Max-Planck-Institute of Biophysical Chemistry in Göttingen, Germany, Anne established her laboratory at the Walter and Eliza Hall Institute of Medical Research in Melbourne in 2000. Since 2012 Anne is the Head of the Development and Cancer Division at the Walter and Eliza Hall Institute. Anne investigates the genetic regulation embryonic development and adult stem cells with particular emphasis on transcriptional regulation through chromatin modifications in health and disease.

ABSTRACT

Chromatin modifications are known to correlate with the transcriptional state of gene loci. However, functional data in multi-cellular organisms are scarce due to our inability to change individual chromatin modifications. Acetylation of histone lysine residues is associated with the transcriptionally active state of gene loci and is catalysed by lysine acetyltransferases.

Previously termed histone acetyltransferases, these enzymes exhibit acetyltransferase activity towards a broad spectrum of lysine residues on histones and non-histone proteins in cell-free biochemical assays. However, we found that within the cellular context, the MYST histone acetyltransferases show surprising substrate specificity for just one histone lysine residue. This has enabled us to examine the effects of histone 4 lysine 16 (H4K16), H3K9 and H3K14 acetylation individually. We observed that genome-wide changes in acetylation of H4K16 and H3K14, regulated by MYST 1 (MOF) and MYST2 (HBO1), respectively, has profound effects on gene transcription and chromatin distribution in the nucleus. In contrast MYST 3 (MOZ) regulates H3K9 acetylation at specific gene loci and so affects gene expression only at these gene loci. Overall, a model of the specific function of individual histone acetyltransferases emerges that re-defines the role of these proteins in chromatin organisation and regulation of gene transcription.

1200 – 1220

MOTIF-BASED ANALYSIS OF CHIP-SEQ DATA

Keynote Speaker – Assoc Prof Tim Bailey
INSTITUTE FOR MOLECULAR BIOSCIENCE,
THE UNIVERSITY OF QUEENSLAND, AUSTRALIA

BIOGRAPHY

Dr Bailey is an Associate Professor and a Principal Research Fellow at the Institute for Molecular Bioscience at The University of Queensland. His research is in bioinformatics and computational biology with an emphasis on the development of methods for understanding transcriptional regulation. Dr Bailey is best known for the MEME Suite of motif-based sequence analysis tools for analyzing motifs in DNA, RNA and protein sequences. The MEME Suite is now used by about 18,000 individual users per month, and was cited over 1200 times in the last 18 months. For example, in 2012 the MEME Suite was involved in discovering a group of cyclic peptides in plants, profiling protein motifs involved in lipid-protein interactions, characterizing a new superfamily of fungal effector proteins, exploring DNA recognition by homeodomain proteins, discovering motifs involved in origins of replication in yeast, and discovering a purine motif involved in chromatin silencing at the nuclear lamina. Dr Bailey received his undergraduate degree in Mathematical Sciences from Stanford University in 1977, and his PhD in Computer Science from the University of California, San Diego in 1995.

ABSTRACT

ChIP-seq experiments are the current method of choice for surveying the targets of DNA-binding transcription factors (TFs). Motif-based sequence analysis of ChIP-seq data can provide extremely valuable insights into the biological mechanisms underlying transcriptional regulation. For example, it can determine the DNA-binding affinity (motif) of the factor, distinguish regions bound directly or indirectly by the factor, and suggest the identities of TFs that regulate cooperatively. Motif analysis can also be employed in differential mode to identify explanatory differences in the motif content of genomic regions bound by the factor under different cellular contexts. I will describe several types of motif-based analysis and give concrete examples of biological insights they can yield.

1220 – 1235

EPIGENETIC REGULATION OF RIBOSOMAL GENE TRANSCRIPTION DURING MALIGNANT TRANSFORMATION

Contributed Paper – Dr Jeannine Diesch
PETER MACCALLUM CANCER CENTRE, AUSTRALIA

BIOGRAPHY

Jeannine Diesch completed her undergraduate studies in 2008 at the University of Ulm in Germany with Honours (BSc(Hons)) in Molecular Biology and Endocrinology. She conducted her PhD studies (PhD in Biochemistry and Molecular Biology) in the laboratory of Dr Amardeep Dhillon at the Peter MacCallum Cancer Centre in Melbourne investigating the role of AP-1 transcription factors in cancer invasion and metastasis. Since the beginning of this year she is a postdoctoral fellow in the Oncogenic Signalling and Growth Control Program headed by Assoc Prof. Ross Hannan and Prof. Grant McArthur and investigates the regulation of ribosome biogenesis in malignant transformation.

ABSTRACT

The ribosomal genes (rDNA) are transcribed by RNA polymerase I (Pol I) in the nucleoli from 200 rDNA repeats per haploid mammalian genome of which only a fraction are active at any given time. rDNA transcription accounts for approximately 80% of the total cellular transcription, making it rate-limiting for ribosome synthesis and cellular growth. Ribosome biosynthesis is often elevated in human cancers and plays an important role in cancer progression (Bywater et al, Cancer Cell 2012). However, its regulation during malignancy is not well understood.

We are utilising the c-Myc-driven Em-*Myc* transgenic mouse model of B-cell lymphoma to investigate the regulation of rDNA transcription during cancer progression. Compared to wild type B-cells, malignant Em-*Myc* B-cells have hyperactivated Pol I and increased rDNA transcription. This increase in transcription is associated with a dramatic increase in the number of active rDNA repeats, indicating changes in chromatin structure during transformation.

In order to assess epigenetic changes and enrichment of rDNA-binding factors at the rDNA loci we are utilising chromatin-immunoprecipitation (ChIP) followed by rDNA hybridisation capture using wild type, pre-malignant and malignant Em-*Myc* B-cells.

As changes in the three-dimensional organisation of the nucleus and nucleoli have been shown to affect gene expression, we are employing chromosome conformation capture sequencing (4C-seq) to identify long distance rDNA interactions occurring within all cell stages, followed by functional testing of those loci found to have altered interactions and therefore expression.

These studies will continue to develop our understanding of rDNA transcription and its contribution to malignant transformation.

SESSION 11
RNA/TRANSCRIPTOMICS

CHAIR: ASSOC PROF MARCEL DINGER
AND DR MARK WALTHAM

1330 – 1410

POWER AND LIMITATIONS OF RNA-SEQ: A PROGRESS REPORT FROM THE SEQC PROJECT

Plenary Speaker – Dr Leming Shi
SCHOOL OF PHARMACY, FUDAN UNIVERSITY, CHINA

BIOGRAPHY

Dr. Leming Shi's research focuses on pharmacogenomics, bioinformatics, and cheminformatics aiming to realize personalized medicine. As a Principal Investigator at the US Food and Drug Administration (FDA), Dr. Shi conceived and has been leading the community-wide MicroArray and Sequencing Quality Control (MAQC/SEQC) project (www.fda.gov/MicroArrayQC/, www.nature.com/nbt/focus/maqc/ and www.nature.com/focus/maqc2/). Dr. Shi was a co-founder of Chipscreen Biosciences Ltd. in Shenzhen, China where he helped develop a chemogenomics-based drug discovery platform leading to several novel small-molecule drug candidates with promising efficacy and safety profiles in anticancer and antidiabetic clinical trials. Dr. Shi is a co-inventor on nine issued patents about novel therapeutic molecules and has published over 160 peer-reviewed papers (eight of them appeared in Nature Biotechnology). Dr. Shi received his Ph.D. in computational chemistry from the Chinese Academy of Sciences in Beijing.

ABSTRACT

In the US FDA-led SEQC (i.e., MAQC-III) project, different RNA-Seq platforms were tested across more than ten sites using well-established reference RNA samples with built-in truths in order to assess the discovery and expression-profiling performances of platforms and analysis pipelines. The results demonstrate that novel exon-exon junctions can still be discovered beyond existing comprehensive annotations and sequencing depth. With extensive investigations encompassing diverse performance metrics (accuracy, precision, reproducibility, titration consistency, mixing ratio recovery, and mutual information) and comparisons to qPCR and microarray platforms, we found high levels of inter-site and cross-

platform concordance for differentially expressed genes. However, performance is clearly platform and pipeline dependent, and transcript-level profiling shows larger variation. Together with applications of RNA-Seq to several preclinical and clinical problems, the entire SEQC data sets comprise over 100 billion reads (10 Tb) and provide a unique resource for testing future developments of RNA-Seq in clinical and regulatory settings.

1410 – 1430

DECODING MIRNA REGULATED GENETIC CIRCUITS

Keynote Speaker – Dr Nicole Cloonan
QUEENSLAND INSTITUTE OF MEDICAL RESEARCH, AUSTRALIA

BIOGRAPHY

Nicole Cloonan is an ARC Future Fellow who has recently established the "Genomic Biology Laboratory" at the QIMR Berghofer Medical Research Institute. Her work is multi-disciplinary in nature, involving bioinformatics, biochemistry, cell biology, and molecular biology – all of which she uses to understand the complexity, function, and systems biology of RNA. Her previous work has included the development of protocols and analysis techniques for RNAseq, sequencing the genomes of pancreatic cancer patients and their tumours, but most recently has been heavily involved in the miRNA field.

ABSTRACT

Maintenance of stem cell pluripotency requires a balance between cellular plasticity and cellular integrity. MicroRNAs have long been useful as markers of cellular states (for example, they can classify tumour tissue of origin or molecular subtypes of cancer), however it is less clear whether these molecules are causal to or correlated with these stable cellular states. "Project Grandiose" has collected matched proteomic and transcriptomic data sets from cells transitioning between stable states during iPS reprogramming, and this resource offers a unique opportunity to both ascertain the general role of miRNAs in cell state transition, and to understand some of the large genomic circuitry regulated by miRNAs.

1430 – 1450

GENOMIC DARK MATTER: THE COMPLEXITY OF LONG NON-CODING RNAS FROM MECHANISM TO THERAPEUTIC

Keynote Speaker – Prof Kevin Morris
SCHOOL OF BIOTECHNOLOGY AND BIOMEDICAL SCIENCES, UNIVERSITY OF NEW SOUTH WALES, AUSTRALIA

BIOGRAPHY

Kevin Morris received his BS from Humboldt State University and PhD from the University of California Davis. During his post-doctoral training at the University of California San Diego he determined that non-coding RNAs were capable of epigenetically modulating gene transcription in human cells. The Morris lab has since determined the mechanistic underpinnings of non-coding RNA mediated regulation of gene transcription in human cells, with evidence suggesting a role for longer forms of non-coding RNAs function as endogenous effectors involved in epigenetically remodeling target loci in human cells. The lab is currently interested in utilizing the recently described endogenous non-coding RNA pathway in human cells to epigenetically modulate gene expression in those genes involved in HIV, cystic fibrosis, and cancer.

ABSTRACT

Observations over the past eight years have determined that in human cells non-coding RNAs (ncRNAs) transcriptionally modulate gene expression and epigenetic states at particular loci. One mechanism involved in antisense ncRNA directed transcriptional gene silencing (TGS) appears to involve DNA methyltransferase 3a (DNMT3a) and Ehancer of Zeste 2 (EZH2). Some of the ncRNAs found to operate in this pathway have been shown to exhibit bimodal functionalities, whereby they can function in both a nuclear and cytoplasmic context via higher ordered complex RNA:RNA interactions. We present evidence here suggesting that several disease relevant genes involved in HIV, cystic fibrosis, and cancer are also under this form of RNA directed epigenetic regulation, with each exhibiting distinctly unique and inherent complexities. Collectively, the data presented here offer subtle insights into the fabric of ncRNA based regulation in human cells while simultaneously suggesting that a mechanism of action can be taken advantage of to either transcriptionally silence or activate a gene.

1450 – 1505

INTRON RETENTION COUPLED WITH NONSENSE MEDIATED DECAY: A PREVIOUSLY OVERLOOKED MECHANISM OF GENE EXPRESSION CONTROL IN GRANULOPOIESIS

Contributed Paper – Dr Justin Wong
CENTENARY INSTITUTE, AUSTRALIA

BIOGRAPHY

Dr. Justin Wong is a Senior Research Fellow of the Sydney Medical School, University of Sydney. He is also a Research Officer at the Gene and Stem Cell Therapy Program, Centenary Institute headed by Professor John Rasko. Dr. Wong have published numerous high profile papers in the field of epigenetics including articles in the New England Journal of Medicine, Cancer Cell and Nature Structural & Molecular Biology. He recently published a first author paper in Cell on the role of intron retention and granulocyte maturation.

ABSTRACT

Alternative splicing (AS) is an important mechanism that promotes protein diversity. Of all modes of AS, intron retention (IR) is the least understood. IR has been widely regarded as a failure in the splicing machinery to excise intronic sequences from pre-messenger RNAs, but its role in normal physiology is unclear. Using mRNA-seq, we measured differential intron IR in FACS purified cells at three progressive stages of mouse granulopoiesis; $CD34^{+}Kit^{+}Gr\text{-}1^{low}$ promyelocytes, $CD34^{-}Kit^{-}Gr\text{-}1^{mid}$ myelocytes and $CD34^{-}Kit^{-}Gr\text{-}1^{high}$ granulocytes. We found that IR affects 86 genes, including those specific to granulocytes (*Lyz2* and *MMP8*) and nuclear architecture (*Lmnb1* and *Lbr*). IR was associated with the downregulation of protein measured by mass spectrometry (P=0.0015, binomial test). There was a significant overlap of IR between human and mouse (P=2.85E-22, hypergeometric test), showing that IR is conserved. Inhibition of nonsense-mediated dacay (NMD) in granulocytes resulted in marked accumulation of 39/86 intron retaining mRNAs (P<0.05, RUV with Holm-Bonferroni correction), indicating that IR triggers NMD to downregulate mRNA and protein expression. Knockdown of the core NMD factor Upf1 also resulted in the accumulation of IR in 6/6 genes examined. Importantly, analysis of nascent RNA transcripts demonstrated that IR-mediated NMD occurred independently of transcriptional regulation in mRNA expression control. Enforced re-expression of *Lmnb1*, which displayed the highest levels of IR, decreased granulocyte cell count, increased nuclear volume by 30% and altered nuclear morphology. We conclude that IR coupled with NMD provides yet another mechanism of gene expression control during normal granulopoiesis[1].

SESSION 12
PLATFORM GENOMICS

CHAIR: DR NICOLE CLOONAN AND
ASSOC PROF CHRISTINE WELLS

1530 – 1610

BIOINFORMATICS SERVICES IN AUSTRALIA – A COLLABORATION WITH THE EUROPEAN BIOINFORMATICS INSTITUTE

Keynote Speaker – Graham Cameron
BIOINFORMATICS RESOURCE AUSTRALIA – EMBL, AUSTRALIA

BIOGRAPHY

Graham Cameron joined EMBL in 1982 working on the world's first public DNA database. In 1986 he took over the leadership of that project and developed the concept for the European Bioinformatics Institute (EBI) and oversaw its launch. Until April 2012 he was responsible for the EBI's databases and services. In October 2012 he joined the University of Queensland to direct the Bioinformatics Resource Australia–EMBL (BRAEMBL).

ABSTRACT

Bioinformatics is crucial to all life science research. The European Bioinformatics Institute (EBI) is one of a few major centres in the world that provide data and services for bioinformatics and, with Australia's membership of EMBL, a natural collaborator of Australia. In 2010 a project was launched to mirror EBI services from the University of Queensland (UQ). The goal was to improve Australian bioinformatics by removing barriers of geographical remoteness.

We have revisited the Mirror's mission in light of experience and with input from a survey of Australian bioinformatics needs, and are creating the

Bioinformatics Resource Australia – EMBL (BRAEMBL) with a mission to:

- enable optimal exploitation of the tools and data of bioinformatics by Australian scientists
- contribute to the global biomolecular information infrastructure in a way which showcases Australian science.
- engage in Australia-wide training in support of these goals

Key findings of the survey and the rationale for the BRAEMBL project will be presented.

BRAEMBL will work with the EBI to create a part of the EBI in Australia and to ensure that Australian

scientists have access to the data and methods of bioinformatics and the necessary IT resources, though integrated high-quality services to rival those available anywhere in the world. This will draw on the support of Australian partners including BioPlatforms Australia (BPA) and the existing eResearch infrastructure. It will work with UQ's Research Computing Centre to be early adopters of modern IT methodologies, in particular cloud computing.

The evolving plan for the BRAEMBL and its contribution to Australian bioinformatics will be presented.

1610 – 1630

HIGH THROUGHPUT FUNCTIONAL GENOMICS APPROACHES TO IDENTIFYING NOVEL DRIVERS OF MELANOMA

Keynote Speaker – Assoc Prof Brian Gabrielli
DIAMANTINA INSTITUTE, THE UNIVERSITY OF QUEENSLAND, AUSTRALIA

BIOGRAPHY

Associate Professor Gabrielli heads the Cell Cycle Group and is the Head of Cancer Program at The University of Queensland Diamantina Institute He is also the director of the ARVEC functional genomics facility developed in collaboration with Prof Tom Gonda. The facility has established high throughput genome wide over and under expression to identify genes that functionally contribute to a wide range of biological processes. Dr Gabrielli's main research interest is targeting defective cell cycle controls in cancer.

ABSTRACT

SNP genotyping, genomic sequencing and transciptomic analysis of cancers provides an unbiased approach to identifying drivers of transformation and metastasis. However, these commonly rely on statistical methods to identify recurrent changes in the genome that are significantly associated tumorigenesis and progression. This approach has been successful in identifying common and highly penetrant changes, but fails where multiple different individual changes on pathways producing a similar outcome are the targets. Melanoma with its high mutation rate is an example where most of the recurrent changes have been identified but these account for a relatively small proportion of melanomas. We are using two different high throughput over expression gain of function screens to identify dominant genes and mutations that are early drivers of transformation. These offer the opportunity to identify target pathway that are commonly defective in melanomas and may be targets for future therapies.

Poster Abstracts

CANCER GENOMICS

1 CLONAL PROLIFERATION OF HEPATOCYTES AND THE DEVELOPMENT OF HEPATOCELLULAR CARCINOMA IN CHRONIC LIVER DISEASE
Thomas Tu, Magdalena Budzinska, Robert Cheng, Fabio Luciani, Nicholas Shackel

2 IDENTIFICATION OF DIFFERENTIAL EXPRESSION OF VARIANT PROSTATE SPECIFIC TRANSCRIPTS AT THE SLC45A3 PROSTATE CANCER LOCUS USING RNASEQ
John Lai, Jiyuan An, Raja Vasireddy, Melanie Lehman, Srilakshmi Srinivasan, Leire Moya, Jyotsna Batra, Judith Clements, Colleen Nelson

3 NEXT-GENERATION SEQUENCING AS A TOOL FOR RAPID ASSESSMENT OF DEFECTIVE SIGNALLING PATHWAYS IN CANCER
Aedan Roberts, Linda Prasad, Tasfia Chowdhury, Daniel Catchpoole

4 ASSOCIATION OF MIRSNPS WITH PROSTATE CANCER RISK: RESULTS FROM LARGE COLLABORATIVE ONCOLOGIC GENETIC STUDIES (COGS)
Jyotsna Batra, Shane Stegeman, Amanda Spurdle, Judith Clements, The PRACTICAL Consortium

5 GENOMIC EVOLUTION OF HEPATOCYTES TOWARDS HEPATOCELLULAR CARCINOMA
Magdalena Budzinska, Thomas Tu, Robert Cheng, Fabio Luciani, Nicholas Shackel

6 PROFILING THE EVOLUTION OF MALIGNANT MELANOMA
Vincent Corbin, Clare Fedele, Samantha Boyle, Robert Fuller, Mark Shackleton, Anthony Papenfuss

7 EPIGENOMIC AND GENOMIC ANALYSIS OF FAMILIAL PROSTATE CANCER
Emma Cazaly, Adele Holloway, Ms Joanne Dickinson, Jac Charlesworth

8 COMPREHENSIVE MOLECULAR CHARACTERIZATION OF PHEOCHROMOCYTOMA
Aidan Flynn, Diana Benn, Roderick Clifton-Bligh, Bruce Robinson, Annette Hogg, Dr Alison Trainer, Dr Paul James, Dr Richard Tothill, Dr Rodney Hicks

9 C-MYC AND HER2 COOPERATE TO DRIVE A STEM-LIKE PHENOTYPE WITH POOR PROGNOSIS IN BREAST CANCER
Radhika Nair, Daniel Roden, Wee Teo, Andrea McFarland, Simon Junankar, Jason Herschkowitz, Charles Perou, Warren Kaplan, Sandra O'Toole, Alexander Swarbrick

10 A NEW METHOD FOR DETECTING DIFFERENTIAL VARIABILITY SHOWS CPG ISLANDS ARE HIGHLY VARIABLY METHYLATED IN CANCER
Belinda Phipson, Alicia Oshlack

11 REPURPOSING THERAPIES FOR TREATMENT OF HER2-AMPLIFIED PANCREATIC CANCER
Mark Cowley, Angela Chou, Nicola Waddell, David Chang, Amber Johns, Katia Nones, Ann-Marie Patch, Sean Grimmond, Andrew Biankin

ECOGENOMICS

12 GENOME AND DRUG SUSCEPTIBILITY ANALYSIS OF MYCOBACTERIUM BOLLETII M24 ISOLATED FROM BRONCHOALVEOLAR LAVAGE (BAL) OF A MALAYSIAN PATIENT

Mee Lian Leong, Siew Woh Choo, Yun Fong Ngeow

13 SPUTUM GENE EXPRESSION OF SIX MARKERS IDENTIFIES ASTHMA INFLAMMATORY PHENOTYPE AND CORTICOSTEROID RESPONSE

Katie Baines, Jodie Simpson, Lisa Wood, Rodney Scott, Naomi Fibbens, Heather Powell, Douglas Cowan, Robin Taylor, Jan Cowan, Peter Gibson

14 BEETLE (GUT) JUICE - EXPLORING A UNIQUE SOURCE FOR NOVEL ENZYME ACTIVITIES

Aurelie Laugraud, Paul Maclean

CLINICAL GENOMICS

15 EXPANDING THE PHENOTYPE OF ARID1B-MEDIATED INTELLECTUAL DISABILITY AND IDENTIFICATION OF ABNORMAL CELL CYCLE FUNCTION AS A CONTRIBUTING PATHOGENIC MECHANISM

Joe C H Sim, Susan M. White, Elizabeth Fitzpatrick, Gabrielle R Wilson, Hayley S Mountford1,2, Greta Gillies, Kate Pope, Han Brunner, Pernille Torring, Richard J Leventer, Martin B Delatycki, David J Amor, Paul J Lockhart

16 GENOME-WIDE DNA METHYLATION PATTERNS IN PANCREATIC DUCTAL ADENOCARCINOMA

Katia Nones, Nic Waddel, Sarah Song, Ann-Marie Patch, Karin Kassahn, Amber Johns, Jianmin Wu, John Pearson, Andrew Biankin, Sean Grimmond

17 USE OF ION TORRENT PGM SEQUENCE ANALYSIS TO DETERMINE GENOMIC BREAKPOINTS IN THE LOW-DENSITY LIPOPROTEIN RECEPTOR GENE

Fathimath Faiz, Richard Allcock, Amanda Hooper, Frank van Bockxmeer

18 INVESTIGATING FRAGMENTATION SIGNATURES OF CELL-FREE DNA IN MATERNAL PLASMA USING NEXT-GENERATION SEQUENCING

Dineika Chandrananda, Natalie Thorne, Devika Ganesamoorthy, Damien Bruno, Yuval Benjamini, Terence Speed, Howard Slater, Melanie Bahlo

19 RNA-SEQ ANALYSIS OF THE DORSOLATERAL PREFRONTAL CORTEX IN SCHIZOPHRENIA

Susan Corley, Dr Fillman, Cyndi Shannon Weickert, Marc Wilkins

20 CHARACTERISATION OF URINE CELL-FREE DNA DEEP SEQUENCING DATA FROM KIDNEY TRANSPLANT PATIENTS REVEALS A NOVEL SIGNATURE FOR CELL-MEDIATED REJECTION

Natalie Thorne, Damien Bruno, Devika Ganesamoorthy, Dineika Chandrananda, David Power, Francesco Ierino, Melanie Bahlo, Howard Slater

21 MICROARRAY ANALYSIS OF MIGRAINE PATIENTS ANDHEALTY CONTROLS FROM NORFOLK ISLAND POPULATION

Astrid Rodriguez, Miles Benton, Rodney A Lea, Lyn Griffiths, Bridget Maher

22 IDENTIFICATION OF NOVEL THERAPEUTICS FOR COMPLEX DISEASES FROM GENOME-WIDE ASSOCIATION DATA

Mani Grover, K.A. Mohanasundaram, Sara Ballouz, Tamsyn Crowley, Craig Sherman, Merrridee Wouters

23 GENETIC ANALYSIS OF AN EPILEPTIC SYNDROME REVEALS A MAJOR CAUSE OF A COMMON FORM OF EPILEPSY
Michael Ricos, Yeh Sze Ong, Bree Hodgson, Sarah Heron, Xeniua Ionia, Boukje de Vries, Ingrid Scheffer, Samuel Berkovic, Massimo Pandolfo, Leanne Dibbens

24 STRATIFICATION OF INHERITED DISEASE: OVERLAPPING GENOTYPES BETWEEN KALLMANN SYNDROME (KS) AND HYPOPITUITARISM-ASSOCIATED DISORDERS HIGHLIGHT NEED FOR GENOME SEQUENCING OF KNOWN GENETIC DISORDERS.
Mark McCabe, Carles Gaston-Massuet, Youli Hu, Louise Gregory, Juan-Pedro Martinez-Barbera, Pierre-Marc Bouloux, Mehul Dattani

25 RIBOSOMAL PROTEIN 6S MRNA IS A BIOMARKER UPREGULATED IN MULTIPLE SCLEROSIS, DOWNREGULATED BY INTERFERON TREATMENT, AND AFFECTED BY SEASON
Grant Parnell, Prudence Gatt, Joseph Powell, Peter Visscher, Grant Montgomery, Mark Slee, Malgorzata Krupa, Jeannette Lechner-Scott, Simon Broadley, David Booth

RNA/NONCODING

26 ASSESSMENT OF CARDIAC MICRORNA HIGH THROUGHPUT SEQUENCING DATA SETS GENERATED FROM RNA OF VARYING QUALITY
David Humphreys, Kavitha Muthiah, Liza Thomas, Peter Macdonald, Christopher Hayward

27 LONG NONCODING RNAS IN FORMATION OF MEMORY, THROUGH LONG-TERM POTENTIATION
Jesper Maag, Debabrata Panja, Marcel Dinger, Clive Bramham, Karin Wibrand

28 THE BIOINFORMATICS BEHIND NOVEL TARGET DISCOVERY USING HIGH THROUGHPUT FUNCTIONAL GENOMICS
Cathryn M Gould, Kaylene J Simpson

29 TO BE, OR NOT TO BE: MIRNA PROFILE OF A NOVEL STABLE STATE OF INDUCED PLURIPOTENT STEM CELLS
Hardip Patel, Jennifer Clancy, Nicole Cloonan, Andrew Corso, Mira Puri, Pete Tonge, Andras Nagy, Thomas Preiss

30 A COMPARISON OF PLASMA MIRNA LEVELS IN INSULIN SENSITIVE, INSULIN RESISTANT AND TYPE 2 DIABETIC OBESE INDIVIDUALS VERSUS LEAN CONTROLS: DIFFERENTIAL MIRNA LEVELS BETWEEN GROUPS
Donia Macartney-Coxson, Angela Jones, Miles Benton, Richard Stubbs

PROTEOMICS

31 INTEGRATING FUNCTIONAL SOMATIC MUTATION BURDEN WITH GENE EXPRESSION
Shila Ghazanfar

32 SYSTEMS ANALYSIS OF INTER-TISSUE SIGNALING GENE REGULATION IN ORGANOGENESIS
Joshua Ho, Daniel O'Connell, Peter Park, Richard Maas

33 MULTIPLEXED ANALYSIS OF IN VITRO RECONSTITUTED HUMAN PROTEIN COMPLEXES
Kirill Alexandrov, Yann Gambin, Emma Sierecki, Loes Stevers, Nichole Giles, Mark Polinkovsky,, Mehdi Moustaqil, Sergey Mureev, Wayne Johnston

POPULATION VARIATION

34 PEDIGREE STRATIFIED GENOME-WIDE MINING FOR PAIRWISE SNP INTERACTIONS IN TYPE II DIABETES
Timothy Hancock, Justin Bedo, Gerard Wong, Melanie A. Carless Joanne E. Curran, John Blangero, Peter Meikle, Adam Kowalczyk

35 STABILITY OF BIVARIATE GWAS BIOMARKER DETECTION
Justin Bedo, David Rawlinson, Cheng Soon Ong, Ben Goudey, Qiao Wang, Adam Kowalczyk

36 AN EXPLORATION OF MUTATION DETECTION USING HALOPLEX TARGETED SEQUENCING
Simon Sadedin, Alicia Oshlack

37 A BIOINFORMATIC PIPELINE FOR SNP CALLING IN NEXT GENERATION GENOTYPING-BY-SEQUENCING
David Kainer, Amanda Padovan, Justin Borevitz, William Foley, Carsten Kulheim

38 GENE EXPRESSION VARIABILITY AS A UNIFYING ELEMENT OF THE PLURIPOTENCY NETWORK
Elizabeth Mason, Jessica Mar, Dr Andrew Laslett, Martin Pera, John Quackenbush, Associate Professor Ernst Wolvetang, Christine Wells

39 DEFINING A HOMOGENEOUS PROGENITOR POPULATION
Jessica Schwaber, Lars Nielsen

COMPUTATIONAL BIOLOGY

40 COMBINING TRANSCRIPTOME ASSEMBLIES FROM MULTIPLE DE NOVO ASSEMBLERS TO GENERATE FULL LENGTH RNA SILENCING GENE TRANSCRIPTS IN NICOTIANA BENTHAMIANA
Kenlee Nakasugi, Peter Waterhouse

41 WHAT TO DO WHEN IT'S ALL RELATIVE: HOW TO ANALYSE RELATIVE ABUNDANCES WITHOUT BEING MISLED
David Lovell, Vera Pawlowsky-Glahn, Juan Jos Egozcue

42 EXPLORING DIFFERENCES IN CHROMATIN ORGANIZATION WITH HI-C DATA
Aaron Lun, Gordon Smyth

43 IMPROVED GENOTYPE CALLING FOR RARE VARIANTS
Ruijie Liu, Zhiyin Dai, Rafael A. Irizarry and Matthew E. Ritchie

44 NEXT-GENERATION READ ALIGNMENT
Wei Shi

45 A SYSTEMATIC EVALUATION OF OBSERVED ALLELIC SPECIFIC EXPRESSION ACROSS TWO HUMAN TISSUES
David Wood, Katia Nones, Anita Steptoe, Angelika Christ, Ivon Harliwong, Felicity Newell, Tim Bruxner, David Miller, Nicole Cloonan, Sean Grimmond

46 FINDING MSC GENE EXPRESSION SIGNATURE REGARDLESS OF THE TISSUE OF ORIGIN
Florian Rohart, Celena Heazlewood, Kim-Anh Le Cao, Christine Wells

47	A WORKFLOW TO INCREASE VERIFICATION RATE OF CHROMOSOMAL STRUCTURAL REARRANGEMENTS USING HIGH THROUGHPUT NEXT GENERATION SEQUENCING **Quek K**, Nones K, Patch A-M, Cloonan N, Newell F, Fink L, Christ A, Miller D, Bruxner T, APGI, Waddell N, Biankin AV, Pearson J and Grimmond S
48	MILKING THE PIGEON **Tamsyn Crowley**, Meagan Gillespie, Volker Haring, Paul Monaghan, John Donald, Rob Moore, Kevin Nicholas
49	PHOSPHORYLATION SITE PREDICTION **Ralph Patrick**, Kim-Anh Le Cao, Bostjan Kobe, Mikael Boden
APPLIED GEONOMICS	
50	STEMFORMATICS: STEM CELL GENE EXPRESSION **Rowland Mosbergen**, Othmar Korn, Jarny Choi, Nick Seidenman, Nick Matigian, Christine Wells
51	POST-LIGHT SEQUENCING: PROGRESS AND PROSPECTS **Lauren Bragg**, Glenn Stone, Margaret Butler, Philip Hugenholtz, Gene Tyson
52	A VIRTUAL APPLIANCE APPROACH TO DELIVERING BIOINFORMATICS PIPELINES: AN EXAMPLE WITH TREVA A TARGETED RESEQUENCING ANALYSIS PIPELINE DEVELOPED AT PETER MAC **Jason Li**, Maria Doyle, Isaam Saeed, Richard Tothill
53	NORMALISATION AND DIFFERENTIAL ENRICHMENT ANALYSIS OF CHIP-SEQ DATA **Christoffer Flensburg**, Alicia Oshlack
54	THE VALUE OF CONTROLS IN PEAK CALLING FROM CHIP-SEQ EXPERIMENTS **Fabian Buske**, Phillippa Taberlay, Susan Clark
55	THE TRANSCRIPT AND THE NOISE: COULD PERSONALIZED TRANSCRIPTOMICS BEAT THE DRUG CHEATS? **Sarah Hausner**, Aaron Chuah, Christopher Gore, Hardip Patel, Simon Easteal
EPIGENETICS	
56	DE NOVO IDENTIFICATION OF DIFFERENTIALLY METHYLATED REGIONS IN THE HUMAN GENOME **Tim Peters**, Mike Buckley, Susan Clark, Peter Molloy
57	GLOBAL DNA METHYLATION PROFILING REVEALS SIGNIFICANT DIFFERENTIAL METHYLATION OF SUBCUTANEOUS AND OMENTAL ADIPOSE TISSUE BEFORE AND AFTER WEIGHT LOSS **Donia Macartney-Coxson**, Miles Benton, Alice Johnstone, David Eccles, Brennan Harmon, Mark Hayes, Rod Lea, Lyn Griffiths, Eric Hoffman, Richard Stubbs
58	CORRECTING FOR UNWANTED VARIATION IN 450K METHYLATION ARRAY STUDIES **Jovana Maksimovic**, Johann Gagnon-Bartsch, Terry Speed, Alicia Oshlack

RNA/TRANSCRIPTOMICS

59	ORCHESTRATED INTRON RETENTION REGULATES NORMAL GRANULOCYTE DIFFERENTIATION **William Ritchie**, Justin Wong, John Rasko, Jeff Holst
60	DIFFERENTIAL EXPRESSION ANALYSIS OF MULTIFACTOR RNA-SEQ EXPERIMENTS **Yunshun Chen**, Gordon Smyth
61	GENE-LEVEL DIFFERENTIAL EXPRESSION ANALYSIS FOR DE NOVO ASSEMBLED TRANSCRIPTOMES **Nadia Davidson**, Alicia Oshlack
62	THE EMERGING PARADIGM OF REGULATED INTRON RETENTION IN MRNA: MAPPING THE PROGRAM OF RNA-BINDING PROTEINS WITH RNA-SEQ **Dan Andrews**, Vicky Cho, Arleen Sanny, Andy Tan, Chris Goodnow
63	COMPARISON OF IN VITRO AND IN VIVO GENE EXPRESSION IN BOVINE OVARIAN FOLLICLES BY MICROARRAY ANALYSES **Nick Hatzirodos**, Katja Hummitzsch, Helen Irving-Rodgers, Claire Glister, Philip Knight, Ray Rodgers

Don't forget to vote for the AGRF Student Poster – People's Choice Award!

Voting forms available from the registration desk or Tweet your poster choice by using #AMATA

Voting closes at the conclusion of afternoon refreshments on Tuesday 15 October 2013.

CANCER GENOMICS

POSTER 1

CLONAL PROLIFERATION OF HEPATOCYTES AND THE DEVELOPMENT OF HEPATOCELLULAR CARCINOMA IN CHRONIC LIVER DISEASE

Dr Thomas Tu, Ms Magdalena Budzinska, Dr Robert Cheng, Dr Fabio Luciani, Dr Nicholas Shackel

ABSTRACT

While many risk factors for hepatocellular carcinoma (HCC) are known, e.g. chronic hepatitis virus infection, and alcoholic injury, the early stages of HCC development are not clear. Further, heterogeneity in HCC genotypes and phenotypes make it difficult to develop specific cancer biomarkers using altered expression of particular genes. In evolutionary theories of cancer, clonal proliferation is a necessary step in carcinogenesis. Supporting this model, progressive clonal proliferation of hepatocytes, even prior to cellular dysplasia, has been observed leading up to HCC in various aetiologies. We aim to use altered clonal proliferation as a more sensitive marker of HCC risk. We hypothesise that the distribution of single nucleotide polymorphisms (SNP) in mitochondrial DNA (mtDNA) can be used as a measure of altered clonal proliferation and thus act as a biomarker. A computer simulation of a liver undergoing random cell turnover showed that the frequency distribution of mtDNA SNPs was altered when cells with a survival advantage were introduced, simulating a subclone of cells with a driver mutation. Significantly greater numbers and more abundant SNPs were observed in simulations with an introduced survival advantage. The majority (>99%) of the SNPs were at a frequency of <1/10000. Therefore, we are in the process of ultra-deep (>100000x) sequencing of hepatocyte mtDNA from patients with liver inflammation, liver cirrhosis, and HCC as well as normal donor controls. With these studies we intend to show that the diversity, not only the geno/phenotype, of cell subpopulations should be considered in the progressive alterations leading up to cancer.

POSTER 2

IDENTIFICATION OF DIFFERENTIAL EXPRESSION OF VARIANT PROSTATE SPECIFIC TRANSCRIPTS AT THE SLC45A3 PROSTATE CANCER LOCUS USING RNASEQ

John Lai, Jiyuan An, Raja Vasireddy, Melanie Lehman, Srilakshmi Srinivasan, Leire Moya, Jyotsna Batra, Judith Clements, Colleen Nelson

ABSTRACT

Increased expression of the SLC45A3-ELK4 fusion correlates with prostate cancer progression. Here, we further characterise the SLC45A3 and ELK4 loci in prostate cancer cells. We identified two variant prostate specific SLC45A3 transcripts (PST-1 and PST-2) from the Illumina Body Map 2.0 project. Androgens promote prostate cancer and PST-1 and PST-2 was found to be androgen regulated in this study, but importantly, PST-1 and PST-2 are more highly expressed in cells which have lower expression of SLC45A3-ELK4, and vice-versa. Given the association of higher SLC45A3-ELK4 expression in metastatic prostate cancers, we propose that the SLC45A3 PST-1 and PST-2 transcripts might be protective of prostate cancer progression, and that this expression is locally co-ordinated through shifts in alternative transcription. RNAseq analysis of SLC45A3 and ELK4 variants (including SLC45A3-ELK4) from prostate cancer cells treated with (anti)-androgens indicate that variant expression differs in response to these agents. Due to the lack of correlation of chromatin features with gene expression profiles in this study, we suggest that a post-transcriptional mechanism is more likely to mediate the opposite expression of SLC45A3-ELK4 with the SLC45A3 PSTs and in mediating the differences in variant expression in response to (anti)-androgen treatment. Given the higher expression of SLC45A3-ELK4 in metastatic prostate cancer and the opposite expression of PSTs with SLC45A3-ELK4, and the different sensitivities to (anti)-androgens of SLC45A3 and ELK4 variants (including SLC45A3-ELK4), we propose that a deregulation of the expression equilibrium in prostate cancer cells at this loci is an important event in the progression of this disease.

POSTER 3

NEXT-GENERATION SEQUENCING AS A TOOL FOR RAPID ASSESSMENT OF DEFECTIVE SIGNALLING PATHWAYS IN CANCER

Aedan Roberts, Linda Prasad, Tasfia Chowdhury, Daniel Catchpoole

ABSTRACT

Cancers are increasingly defined on the basis of defects in particular signalling pathways, which can inform diagnosis and prognosis, and determine likely response to different therapies.

Massively parallel sequencing allows the comprehensive sequencing of a small number of genes with sufficient redundancy to detect even low-level somatic mutations: it is possible to sequence the coding regions of an entire signalling pathway in one clinical specimen with greater than 1000x coverage in one sequencing run. In a clinical setting, this translates to the ability to determine whether a patient's tumour harbours variations in a particular pathway within hours of obtaining a sample, allowing diagnostic or treatment decisions to be made in a single day.

Prognosis in neuroblastoma is highly variable, with a subset of tumours spontaneously differentiating to a benign form. Several lines of evidence implicate interleukin-7 (IL7), secreted from surrounding stroma and signalling through a JAK-STAT pathway, as the factor that induces differentiation in these tumours.

Using the Ion Torrent Personal Genome Machine and the Ion 314 chip, coding regions of the 10 key genes involved in IL7 signalling will be sequenced in 13 neuroblastoma cell lines. This can be accomplished in a single run with greater than 100x coverage, and is expected to provide further evidence of the role of IL7 signalling in neuroblastoma. Our experimental design will also act as a proof-of-concept for the application of massively parallel sequencing to rapidly provide actionable information on defective signalling pathways with the potential to inform decisions on patient treatment.

POSTER 4

ASSOCIATION OF MIRSNPS WITH PROSTATE CANCER RISK: RESULTS FROM LARGE COLLABORATIVE ONCOLOGIC GENETIC STUDIES (COGS)

Jyotsna Batra, Shane Stegeman, Amanda Spurdle, Judith Clements, The PRACTICAL Consortium

ABSTRACT

Shane Stegeman1, A. Spurdle2, J. Clements1, Practical Consortium, Jyotsna Batra1

1Australian Prostate Cancer Research Centre-Queensland And Institute Of Health And Biomedical Innovation, Qut, Brisbane, Queensland, Australia – 4102

2Molecular Cancer Epidemiology Laboratory, Queensland Institute Of Medical Research, Brisbane, Queensland, Australia – 4006

miRNA are small non-coding RNA regulating transcriptional and post-transcriptional gene expression by binding mostly to the 3′ untranslated region of the target genes. Genetic variants in the 3′ UTRs of the target genes (miRSNPs) can affect the base-pairing between miRNAs and mRNAs, and hence disrupt or create novel target sites. To investigate the role of miRSNPs in prostate cancer, we evaluated association of 2,169 miRSNPs within the cancer related genes in a large cohort of 25,074 prostate cancer cases and 24,272 controls.

A total of 242 miRSNPs were found to be associated with the risk of prostate cancer ($p<0.05$). However, 23 SNPs within 16 genes remain significant after correction for multiple testing. The most significant association was observed for a miRSNP rs1058205 present within the KLK3 3′UTR with an OR= 0.86(0.83-0.9), p=1.2x10-14. Reporter vector assays with over-expression of miR-3162-5p resulted in a 20% decrease in luciferase signal (p=0.04) for the rs1058205 SNP T allele vector construct whilst no significant change was observed for the C allele suggesting that miR-3162-5p has specific affinity for the T allele. Further validating these results, we observed a 36% reduction in cellular KLK3 protein expression following miR-3162-5p over-expression in the LNCaP cell line (TT rs1058205 SNP).

Findings from this large study provide evidence that variants in miRNA-related genes may be associated with prostate cancer risk and suggest for the first time that a prostate cancer associated SNP rs1058205 in the KLK3 3′ UTR affects miR-3162-5p targeting in prostate cancer.

POSTER 5

GENOMIC EVOLUTION OF HEPATOCYTES TOWARDS HEPATOCELLULAR CARCINOMA

Magdalena Budzinska, Thomas Tu, Robert Cheng, Fabio Luciani, Nicholas Shackel

ABSTRACT

The molecular pathogenesis of hepatocellular carcinoma (HCC) is poorly understood. There are a number of genetic abnormalities that are typically associated with HCC development. However, no single genomic rearrangement or mutation in a single gene or pathway is consistently abnormal in all cases of HCC. Moreover, driver mutations are only a minority of the total number of genetic changes that occur in HCC.

Cancer genome accumulates a large number of genetic alterations, such as genomic rearrangements, deletions, insertions or translocations. The vast majority of those events are assumed to be phenotypically neutral and are called passenger mutations. They are nonrecurrent and dispersed across a cancer genome. Passenger mutations are believed to have no role in cancer development, however many of them fall within protein-coding genes or other functional elements, thus can have potentially deleterious effects on cancer cells. Although protein-coding passengers may individually exert small effect on a disease development, their collective accumulation may be significant for the cancer progression.

We developed a method to investigate the possible role of passenger mutations in HCC progression. We used publically available HCC Next Generation Sequencing (NGS) datasets from the International Cancer Genome Consortium (ICGC). Our preliminary results show significant differences in the number of SNPs that are predicted to cause an alteration in phenotype between tumor and non-tumor samples, but not total SNP number. We are in the process of sequencing patient cohorts with liver inflammation, liver cirrhosis, HCC and normal donor as a control to investigate cancer progression in different stages of liver disease.

POSTER 6

PROFILING THE EVOLUTION OF MALIGNANT MELANOMA

Vincent Corbin, Clare Fedele, Samantha Boyle, Robert Fuller, Mark Shackleton, Anthony Papenfuss

ABSTRACT

Malignant melanoma is an aggressive, notoriously treatment resistant and fatal disease. The high genomic instability of melanoma is associated with tumour evolution, and the eventual loss of pigmentation in some cases. When de-pigmented, melanoma has been observed to be more aggressive and associated with a poorer prognosis. It is therefore highly important to understand melanoma evolution via characterisation of the genomic differences between various forms.

We examined the evolution of human melanoma in a mouse xenograft model, which is believed to be a highly accurate model of tumour evolution in the patient. Xenografted melanoma cells were continuously propagated through multiple generations of mice, and surprisingly, some melanoma cell populations evolved into a more aggressive non-pigmented form.

Genomic analysis was performed on different generations of melanoma, including pigmented and non-pigmented forms arising from the same tumour cell ancestry. Chromosomal SNP microarrays were used to explore variations in copy number profile between generations. We then developed a novel statistical method to quantify genomic differences between tumours without relying on allelic copy number estimation, as copy number assessment can be imprecise in highly unstable genomes. Our method utilizes statistical analysis across the genome to simultaneously compare both log R ratio and B allele frequency. In addition, we sequenced extracted RNA from a collection of pigmented and non-pigmented tumours, detecting differentially expressed genes to explain the aggressivity associated with tumour depigmentation. Differentially expressed genes were then analysed to identify the gene pathways involved in tumour evolution in the xenograft model.

POSTER 7

EPIGENOMIC AND GENOMIC ANALYSIS OF FAMILIAL PROSTATE CANCER

Emma Cazaly, Adele Holloway, Ms Joanne Dickinson, Jac Charlesworth

ABSTRACT

More thaABn 20,000 men are diagnosed annually with prostate cancer in Australia, yet prognosis and treatment lack sensitivity and specificity, creating challenges involving early detection and determining likelihood of metastasis. A family history of prostate cancer has consistently been identified as a major risk factor and genome wide association studies have identified more than thirty prostate cancer susceptibility variants. However these variants account for less than a third of familial risk of the disease. Epigenetic factors may prove key to understanding this 'missing heredity link' as it is becoming more evident that high-risk variants may occur in the regulatory regions of genes.

DNA methylation, the most commonly studied epigenetic mark, is frequently disrupted in disease, particularly cancer, where global demethylation occurs in conjunction with promoter specific hypermethylation. While many gene promoters have been found to be hypermethylated in prostate cancer and associated with transcriptional repression, the mechanism initiating this aberrant methylation is yet to be fully understood. By examining the methylation pattern of specific gene regions using Illumina's novel Infinium HumanMethylation450 BeadChip platform we aim to identify additional high-risk variants with potential uses as biomarkers and treatment targets. Around 50 blood DNA samples from affected men from the Tasmanian Familial Prostate Cancer Resource together with unaffected relatives across three generations have been interrogated on the platform. Analysis of these methylation patterns will not only allow the investigation of disease associated aberrant marks, particularly in relation to SNPs, but also the exploration of transgenerational epigenetic inheritance.

POSTER 8

COMPREHENSIVE MOLECULAR CHARACTERISATION OF PHEOCHROMOCYTOMA

Aidan Flynn, Diana Benn, Roderick Clifton-Bligh, Bruce Robinson, Annette Hogg, Dr Alison Trainer, Dr Paul James, Dr Richard Tothill, Dr Rodney Hicks

ABSTRACT

Pheochromocytoma (PC) are neuroendocrine tumours that arise from adrenal medullary and extra-adrenal chromaffin cells. These tumours are relatively rare with an estimated incidence of around 1 per 100,000 people (Beard, Sheps et al. 1983) and can be classified as either functional or non-functional depending on their ability to secrete catacholamines. There are currently 13 genes linked to an increased risk of developing pheochromocytoma alone or as part of a multi-site cancer syndrome. Mutations in these genes account for more than 90% of ostensibly heritable PC but only around 15% of sporadic PC, as such, the genetic events contributing to sporadic PC remain largely unknown. We aim to combine RNA-Seq, exome sequencing, and copy number variation (CNV) analysis on a cohort of 50 PC tumours to better understand the underlying genetic changes driving tumorigenesis. As a starting point to this work we have used gene-expression profiling to molecularly subtype PC. Previously, unsupervised clustering of microarray data has revealed two major groups; a pseudo-hypoxia group which consists of tumours with a mutation in VHL or any of the SDHx genes, and a Receptor Tyrosine Kinase (RTK) signaling group harboring mutations in RET, NF1, or TMEM127 (Dahia, Ross et al. 2005). We aim to further refine this paradigm into additional sub-groups and, through development of a cross platform classifier, leverage existing microarray data to classify our RNA-Seq data. This classification will be used to guide the selection of candidate genes found through exome sequencing and/or found to be in a region of chromosome instability by CNV analysis.

POSTER 9

C-MYC AND HER2 COOPERATE TO DRIVE A STEM-LIKE PHENOTYPE WITH POOR PROGNOSIS IN BREAST CANCER

Radhika Nair, Daniel Roden, Wee Teo, Andrea McFarland, Simon Junankar, Jason Herschkowitz, Charles Perou, Warren Kaplan, Sandra O'Toole, Alexander Swarbrick

ABSTRACT

The HER2 (ERBB2) and MYC genes are commonly amplified in breast cancer, yet little is known about their molecular and clinical interaction. Using a novel chimeric mammary transgenic approach and in vitro models, we demonstrate markedly increased self-renewal and tumour propagating capability of cells transformed with Her2 and c-Myc. Co-expression of both oncoproteins in cultured cells led to activation of a c-Myc transcriptional signature and acquisition of a self-renewing phenotype independent of an EMT programme or regulation of conventional cancer stem cell markers. Instead, Her2 and c-Myc cooperated to induce expression of lipoprotein lipase, which was required for proliferation and self renewal in vitro. HER2 and MYC were frequently co-amplified in breast cancer, associated with aggressive clinical behaviour and poor outcome. Lastly, we show that in HER2+ breast cancer patients receiving adjuvant chemotherapy (but not targeted anti-HER2 therapy), MYC amplification is associated with a poor outcome. These findings demonstrate the importance of molecular and cellular context in oncogenic transformation and acquisition of a malignant stem-like phenotype and have diagnostic and therapeutic consequences for the clinical management of HER2+ breast cancer.

POSTER 10

A NEW METHOD FOR DETECTING DIFFERENTIAL VARIABILITY SHOWS CPG ISLANDS ARE HIGHLY VARIABLY METHYLATED IN CANCER

Belinda Phipson, Alicia Oshlack

ABSTRACT

Methylation of DNA is known to be dramatically altered in cancers. New technologies such as the Illumina 450k methylation arrays are a cost-effective way to measure methylation at over 480,000 CpG sites across the genome. Hence they have been used to profile thousands of samples of both normal and tumour tissues. The primary focus of analysis methods to date has concentrated on detecting sites of differential methylation, i.e. a difference in the mean level of methylation between two or more groups. However, there have been a few reports in the literature claiming increased variability of methylation in cancer samples compared with normal tissue. The statistical methods used previously (F test and Bartlett test) are not well characterised for this data and are known to be highly sensitive to non-normality and outliers. We have therefore developed a new method to detect differential variability between sample groups. Our method is based on Levene's z test and uses the limma framework, giving it the flexibility to handle any experimental design. We show our test is robust as well as being computationally fast and outperforms previously used methods. We have applied our method to various large cancer/normal methylation datasets from The Cancer Genome Atlas (TCGA) and also find increased variability in tumour samples. Interestingly, a large proportion of the top differentially variable CpG sites are found in CpG islands and are not associated with CpG sites that are differentially methylated. This phenomenon is seen across different tumour types we have analysed from TCGA.

POSTER 11

REPURPOSING THERAPIES FOR TREATMENT OF HER2-AMPLIFIED PANCREATIC CANCER

Mark Cowley, Angela Chou, Nicola Waddell, David Chang, Amber Johns, Katia Nones, Ann-Marie Patch, Sean Grimmond, Andrew Biankin

ABSTRACT

Pancreatic ductal adenocarcinoma (PDAC) is one of the most lethal, and molecularly diverse malignancies, with few therapeutic options. We have used next generation sequencing to define the molecular features of PDAC. Repurposing of therapeutics that target specific molecular mechanisms in different disease types offers considerable potential for rapid improvements in outcome. Herceptin®, a monoclonal antibody targeting HER2 has been successfully used to treat HER2-amplified breast cancer, and has recently been repurposed for gastric cancer; the incidence of HER2-amplified PDAC is poorly understood.

Through a combination of whole exome, whole genome, copy number and microarray analysis, we identified 1/99 PDAC tumours with a ~1MB high-level amplification, affecting HER2, resulting in high mRNA and protein expression. From an archival cohort of 369 PDAC tissues, and using IHC and SISH, we identified 10 patients (2.1%) with high-level HER2-amplification, and defined standardized guidelines for identifying HER2-amplified PDAC.

Strikingly, the distribution of metastatic sites in HER2 amplified PDAC showed strong preference to brain and lung, over the normal pattern of liver (P<5E-7; relative to non-HER2-amplified PDAC), similar to that observed in HER2-amplified breast cancer.

The substantial molecular heterogeneity of PDAC implies that with an incidence of 2.1%, HER2-amplification potentially represents a valuable therapeutic target. Since collecting these data, we observed a significant response to Herceptin therapy in a patient with HER2-amplified PDAC. We have established the IMPACT clinical trial to assess the efficacy of three genotype-guided therapies in PDAC, including Herceptin.

ECOGENOMICS

POSTER 12

GENOME AND DRUG SUSCEPTIBILITY ANALYSIS OF MYCOBACTERIUM BOLLETII M24 ISOLATED FROM BRONCHOALVEOLAR LAVAGE (BAL) OF A MALAYSIAN PATIENT

Mee Lian Leong, Siew Woh Choo, Yun Fong Ngeow

ABSTRACT

Mycobacterium abscessus is an opportunistic human pathogen has caused wide spectrum human diseases include serious pulmonary diseases and healthcare associated infections predominantly in immune deficient patient. M. abscessus has recently been dissociated into 3 subspecies: M. abscessus sensu stricto, M. bolletii and M. massiliense. In general, M. bolletii is the most drug resistant compared to the other subspecies.

Here we described M. bolletii M24 which isolated from bronchoaveolar lavage (BAL) of a Malaysian patient. The M24 strain is been sequenced using both Illumina and PacBio platform, annotated and analysed using different bioinformatic approaches. The draft genome of M24 was assembled into 67 contigs with Illumina reads. Scaffolds of 20 were resulted from hybrid scaffolding of Illumina reads and error corrected PacBio data. In the assembled genome which has 5,488,620 bp with 63.9% GC content, the total predicted coding sequences and RNAs was 5,560 and 89 respectively. Interesting findings include a few incomplete prophages and a 37-kb intact prophage, suggesting that phages have played an important role in the diversification of M24. We have discovered a high number of tRNAs(86), of which 35 organised in a 9,000bp island and verified by PacBio sequencing data.

In drug susceptibility pattern analysis, an intact erm(41) gene revealed in M24 confers inducible clarithromycin resistance and existence of bla gene determined the imipenem resistance. These have been verified by minimum inhibitory concentration (MIC) test. This study elucidates the genetic basis of M. abscessus and will form the benchmark for future functional works on this pathogen.

POSTER 13

SPUTUM GENE EXPRESSION OF SIX MARKERS IDENTIFIES ASTHMA INFLAMMATORY PHENOTYPE AND CORTICOSTEROID RESPONSE

Katie Baines, Jodie Simpson, Lisa Wood, Rodney Scott, Naomi Fibbens, Heather Powell, Douglas Cowan, Robin Taylor, Jan Cowan, Peter Gibson

ABSTRACT

Rationale: Airway inflammation is associated with asthma exacerbation risk, treatment response, and disease mechanisms.

Objective: This study aimed to identify and validate a sputum gene expression signature that discriminates asthma inflammatory phenotypes.

Methods: Gene expression profiles were generated from 47 asthmatics using Illumina HumanRef8 BeadChips, and analysed for differential expression. Potential gene biomarkers were further confirmed by qPCR: in the discovery study, a clinical validation study (n=59 asthmatics and 20 healthy controls), and an inhaled corticosteroid (ICS) responsiveness study (n=71). Marker expression was analysed using multiple logistic regression and area under the curve (AUC) values.

Results: There were 277 genes altered between asthma inflammatory phenotypes. Expression results were confirmed by qPCR, and 23 out of 35 genes showed highly significant differential expression in both the discovery and validation populations. A signature of six genes was derived, including CLC, CPA3, DNASE1L3, IL1B, ALPL, and CXCR2, that could significantly ($p<0.0001$) discriminate eosinophilic asthma from each other phenotype including non-eosinophilic asthma (AUC=89.6%), paucigranulocytic asthma (AUC=92.6%), neutrophilic asthma (AUC=91.4%) and healthy controls (AUC=97.6%). This signature could also significantly ($p<0.0001$) distinguish neutrophilic asthma from non-neutrophilic asthma (AUC=84.5%), paucigranulocytic asthma (AUC=85.7%) and healthy controls (AUC=90.8). The six gene signature predicted ICS response (>12% change in FEV1, AUC=91.5%), and inflammatory phenotype independently.

Conclusions: A sputum gene expression signature of six biomarkers significantly discriminates inflammatory phenotypes of asthma, and predicts treatment response. This signature has the potential to become a useful diagnostic tool to assist in the clinical diagnosis and management of asthma.

POSTER 14

BEETLE (GUT) JUICE - EXPLORING A UNIQUE SOURCE FOR NOVEL ENZYME ACTIVITIES

Aurelie Laugraud, Paul Maclean

ABSTRACT

The geographic separation of New Zealand (NZ) from other major land masses for >80 million years has resulted in the evolution of unique flora and fauna. Among these are scarab beetles, whose larvae consume plant biomass including roots, wood, and humus. Like ruminant herbivores, NZ scarab beetles rely on a community of cellulolytic endosymbiotic microorganisms in their gut to convert recalcitrant plant matter into useful end products, such as energy. We hypothesized that scarab insect guts function as miniature bioreactors and are, thus, a useful model to obtain fundamental insights into the digestion of complex plant material. To begin testing this idea we employed next generation sequencing (NGS) to look for candidate genes involved in a range of enzymatic activities from the hindgut of NZ scarab larvae. Here we wish to discuss bioinformatics methods to analyse the NGS dataset. More than 100 Gbp of RNA-Seq (Illumina HiSeq, paired-end, 100bp, 400bp insert) were generated. We want to mine this data for cellulolytic genes and species/organisms present within scarab hindgut tissue. The absence of a reference genome for scarab makes this a challenging meta-transcriptomics study. We tested some of the commonly used tools available and discuss to what extend they are relevant for our dataset. A strategy combining the best parts of publicly available tools with our own adjustments is presented.

POSTER 15

EXPANDING THE PHENOTYPE OF ARID1B-MEDIATED INTELLECTUAL DISABILITY AND IDENTIFICATION OF ABNORMAL CELL CYCLE FUNCTION AS A CONTRIBUTING PATHOGENIC MECHANISM

Joe C H Sim, Susan M. White, Elizabeth Fitzpatrick, Gabrielle R Wilson, Hayley S Mountford1,2, Greta Gillies, Kate Pope, Han Brunner, Pernille Torring, Richard J Leventer, Martin B Delatycki, David J Amor, Paul J Lockhart.

ABSTRACT

ARID1B is a conserved DNA-binding component of the SWItch/sucrose nonfermentable (SWI/SNF)-like chromatin remodelling complex. The protein appears to be important for normal brain development and function as mutations in ARID1B were recently shown to cause Coffin-Siris syndrome (1, 2). ARID1B mutations have also been identified in patients with nonsyndromic intellectual disability and when intellectual disability co-presents with autism and agenesis of the corpus callosum (3, 4).

We identified a patient with moderate intellectual disability, absent speech and dysmorphic features with fetal finger and toe pads and plantar lipomas. Karyotype analysis showed an apparently balanced de novo 4;6 reciprocal translocation. Microarray analysis demonstrated a heterozygous deletion of 1.2Mb on 6q25.3, involving five genes including ARID1B. We subsequently ascertained four other patients with a strikingly similar phenotype of dysmorphism, plantar fat pads and intellectual disability and identified heterozygous de novo ARID1B frameshift mutations in all cases, resulting in haploinsufficiency of ARID1B.

SWI/SNF complexes are involved in regulation of gene expression and ARID1B may play a role in priming cells for cell cycle entry by activating the transcription of cell cycle genes. FACS analysis of primary fibroblast cells derived from affected patients demonstrated delayed cell cycle re-entry after serum starvation. Similarly, biochemical analysis suggested temporal and quantitative alterations in key regulators of cell cycle transition. In conclusion, we have expanded the spectrum of features associated with haploinsufficiency for ARID1B and provided the first functional data that alterations in cell cycle are an important contributor to the pathogenic mechanism underlying SWI/SNF complex disorders.

1. Santen et al, 2012, Nat Genet, 44(4):379-80
2. Tsurusaki et al, 2012, Nat Genet, 44(4):376-8
3. Hoyer et al, 2012, Am J Hum Genet, 90, 565–572
4. Halgren et al, 2012, Clin Genet, 82(3):248-55

POSTER 16

GENOME-WIDE DNA METHYLATION PATTERNS IN PANCREATIC DUCTAL ADENOCARCINOMA

Katia Nones, Nic Waddel, Sarah Song, Ann-Marie Patch, Karin Kassahn, Amber Johns, Jianmin Wu, John Pearson, Andrew Biankin, Sean Grimmond

ABSTRACT

Pancreatic cancer is the 4th leading cause of cancer death in western societies with a 5-year survival rate of less than 5%, which did not change in the last 45 years. As a consequence there is an urgent need to develop novel therapeutic and early detection strategies. DNA methylation has been increasingly implicated in cancer. Here we examined the DNA methylation profile of 156 pancreatic ductal adenocarcinomas (PDAC) and 23 non-malignant pancreas, using high-density arrays. More than 90,000 CpG sites were significantly differentially methylated (DM) in PDAC relative to non-malignant pancreas, with pronounced alterations in a sub-set of 13,517 CpG sites. This sub-set of DM CpG sites segregated PDAC from non-malignant pancreas, regardless of tumour cellularity. A total of 3981 genes were aberrantly methylated and approximately 36% of these genes showed significant correlation between methylation and mRNA expression levels. Pathway analysis revealed an enrichment of aberrant methylation in genes involved in key molecular mechanisms important to PDAC (TGF-β, WNT, Integrin signaling, Cell adhesion, Stellate cell activation and Axon Guidance). Bisulfite amplicon deep sequencing and qRT-PCR analyses of Axon Guidance genes SLIT2, SLIT3, ROBO1, ROBO3, SRGAP1, and MET suggested epigenetic suppression of SLIT-ROBO signaling and up-regulation of MET expression. Hypo-methylation of MET and ITGA2 correlated with high gene expression, which correlated with poor survival of PDAC patients. These data suggest that aberrant methylation plays an important role in pancreatic carcinogenesis affecting known core signaling pathways further adding to the disease complexity with important implications to ongoing research and future therapeutic development.

POSTER 17

USE OF ION TORRENT PGM SEQUENCE ANALYSIS TO DETERMINE GENOMIC BREAKPOINTS IN THE LOW-DENSITY LIPOPROTEIN RECEPTOR GENE

Fathimath Faiz, Richard Allcock, Amanda Hooper, Frank van Bockxmeer

ABSTRACT

Background: The LDLR has an unusually high content of repetitive regions with Alu repeats accounting for 85% of total intronic sequence. The presence of such repetitive regions throughout the human genome has been linked to enhance likelihood of recombination events, with non-allelic homologous recombination (NAHR) occurring more frequently than non-homologous end joining (NHEJ) recombination. Besides single nucleotide variants (SNVs) and small indels, approximately 10–15% of mutations causing familial hypercholesterolemia (FH) are reported to involve much larger structural variants, usually detected by Multiplex Ligation-dependent Probe Amplification (MLPA). The breakpoints involved are invariable unknown and it has been presumed that the common MLPA variants are identical.

Aim: Identify genomic breakpoints for 12 large deletions in LDLR previously identified by (MLPA) in unrelated patients with FH.

Methods: The MLPA variants were analysed by sequencing on an Ion Torrent PGM sequencer followed by Sanger sequencing.

Results: Approximate (within 1 Kb) genomic breakpoints for the MLPA variants were established from the PGM BAM files using subsequent Sanger sequencing to establish exact breakpoints. In the 12 samples, eleven different rearrangements/mutational events were found, with eight out of eleven occurring in *Alus*. Two of the three samples with the relatively common exons 2-6del variant had identical breakpoints. Furthermore, two exons 11-12del variants also had unique breakpoints, indicating separate ancestral origin or mutational events that were previously considered to be identical by descent from a common ancestor.

Conclusions: This study showed that Ion Torrent PGM sequencing is an effective method to detect large deletions and can be utilized to determine genomic breakpoints for disease causing variants in population studies.

POSTER 18

INVESTIGATING FRAGMENTATION SIGNATURES OF CELL-FREE DNA IN MATERNAL PLASMA USING NEXT-GENERATION SEQUENCING

Dineika Chandrananda, Natalie Thorne, Devika Ganesamoorthy, Damien Bruno, Yuval Benjamini, Terence Speed, Howard Slater, Melanie Bahlo

ABSTRACT

Fragmented cell-free DNA that circulates in plasma is a versatile biomarker that has been used to non-invasively detect fetal aneuploidy in pregnant women, monitor the health of transplanted organs and scan for copy number changes in tumour DNA. In all these situations, Next-Generation Sequencing (NGS) has proved to be a powerful tool for studying circulating DNA on a genome-wide level.

Since most NGS studies have focused on the diagnostic applications of cell-free DNA, only a few studies have investigated its biological characteristics. These studies have reported on the origin, fragment sizes and quantity, but the effects of the natural fragmentation of plasma DNA have not been investigated in depth.

We demonstrate the existence of significant fragmentation bias in cell-free DNA. Nucleotide frequencies show a position specific pattern in the region spanning 8-9 positions on either side of the DNA cleavage site. This fragmentation signature is a true biological effect and is highly replicable across independent samples. Using deep-sequencing data, we also compare this effect between the maternal and fetal DNA fragments.

Our investigations also show a specific motif that is over-represented in plasma DNA and that this effect is confounded with the GC bias that occurs in NGS data. We apply a new bias correction that accounts for both GC & fragmentation based on Benjamini and Speed (2012), and investigate its effect on the sensitivity of fetal aneuploidy detection using maternal plasma.

POSTER 19

RNA-SEQ ANALYSIS OF THE DORSOLATERAL PREFRONTAL CORTEX IN SCHIZOPHRENIA

Susan Corley, Dr Fillman, Cyndi Shannon Weickert, Marc Wilkins

ABSTRACT

Schizophrenia is a complex disorder. As such, identification of specific genes involved in the disease process is challenging. Massively-parallel sequencing of mRNA (RNA-Seq) allows genes of interest to be identified in the transcriptome and can provide exquisite insights into the proteins likely to be active within a tissue. We have analysed RNA-Seq data from dorsolateral prefrontal cortex (DLPC) obtained post-mortem from 19 schizophrenia patients and 19 controls. The reads were previously produced on the SOLiD platform and mapped with XMATE [1]. We performed alternative mapping using TopHat. Differential expression analysis was performed using DESeq and edgeR. Our results indicate potential dysregulation of nuclear receptors and differential expression of the genes activated by these transcription factors. Our results show that RNA-Seq is a powerful tool for identifying aberrant pathways or networks in complex disorders such as schizophrenia.

1. Fillman, S.G., et al., Increased inflammatory markers identified in the dorsolateral prefrontal cortex of individuals with schizophrenia. Mol Psychiatry, 2012.

POSTER 20

CHARACTERISATION OF URINE CELL-FREE DNA DEEP SEQUENCING DATA FROM KIDNEY TRANSPLANT PATIENTS REVEALS A NOVEL SIGNATURE FOR CELL-MEDIATED REJECTION

Natalie Thorne, Damien Bruno, Devika Ganesamoorthy, Dineika Chandrananda, David Power, Francesco Ierino, Melanie Bahlo, Howard Slater

ABSTRACT

Development of a sensitive, non-invasive test that can be used on a daily basis for early detection of rejection has been identified as the highest priority research need in the transplantation field. Currently, kidney transplant patients undergo 3-6 monthly biopsies under general anaesthetic in order to detect signs of allograft rejection.

A common form of kidney transplant rejection, cell-mediated rejection (CMR), involves damage to the tubules of the kidney. We hypothesised that due to this damage, donor cell-free DNA would be released directly from renal tubule cells into the urine. In the absence of rejection, cell-free DNA in the urine would instead originate predominantly by filtration of circulating plasma through the renal glomeruli. We predicted this trans-renal cell-free DNA would be fragmented differently compared to CMR-released cell-free DNA in the urine.

We analysed urine cell-free DNA fragments in kidney transplant patients using whole genome Illumina HiSeq paired-end sequencing and found a distinct fragment length signature for CMR rejection. We explored sequence motifs at the fragmentation sites and other potential patterns in the coverage of fragments across the genome. With these bioinformatic analyses we hope to prompt a better understanding of the natural history of cell-free DNA. We will present our results from these investigations and outline future experiments to confirm our findings and assess inter and intra patient heterogeneity. We will discuss the challenges involved in translating this approach into a simple, inexpensive test that can be used in the clinical setting.

POSTER 21

MICROARRAY ANALYSIS OF MIGRAINE PATIENTS AND HEALTY CONTROLS FROM NORFOLK ISLAND POPULATION

Astrid Rodriguez, Miles Benton, Rodney A Lea, Lyn Griffiths, Bridget Maher

ABSTRACT

Migraine has been defined as a common disabling primary headache disorder by the International Headache Society. So far, a number of genes have been identified to have a major effect on the disease and it has been recognized that its wide phenotypic variation is caused by an important number of genes interacting with each other in a complex network of reaction. In an attempt to discover new genes involved in the pathophysiology of migraine we compared the gene expression patterns of individuals suffering migraine (79) with healthy controls (254) from Norfolk Island population, considered to be a genetic isolate. The analysis was performed using an Illumina

HumanHT-12 v4 array. The statistical significance of differential expressed genes between cases and controls was determined by a standard t-test and a logistic regression analysis (R) . Approximately 3.5% (825) of transcripts representing 684 unique genes, showed differential expression patterns in patients with migraine ($p<0.05$). Transcripts like NUTD3, CREBBP, CACNA1D, NR1D2, PTK2B and ZDHHC1 are over-expressed in migraine patients. Pathway analysis was undertaken using DAVID Bioinformatics Resources and GSEA and Huntington´s disease, MAPK signalling pathway, Chemokine signalling pathway and Calcium signaling pathway were found to be significantly enriched by our gene list. A GWAS analysis showed that SNPs in genes CACNA1D (rs898423) and PTK2B (rs9773817) are significantly associated with migraine in the same population. Finally, data from an eqtl-analysis in the same population showed how the expression of HOXB2 could be affecting the expression of CREBBP, a gene detected to be increased in patients.

POSTER 22

IDENTIFICATION OF NOVEL THERAPEUTICS FOR COMPLEX DISEASES FROM GENOME-WIDE ASSOCIATION DATA

Mani Grover, K.A. Mohanasundaram, Sara Ballouz, Tamsyn Crowley, Craig Sherman, Merrridee Wouters

ABSTRACT

Candidate gene prediction systems identify genes likely to be of functional relevance to a phenotype from associated genetic loci. Gentrepid, a human candidate gene discovery platform, utilizes two algorithms- Common Module Profiling and Common Pathway Scanning - to prioritize candidate genes for human inherited disorders. Recently, several protocols were developed to apply Gentrepid to the analysis of data from Genome Wide Association Studies (GWAS) using the Wellcome Trust Case Control Consortium (WTCCC) data set on seven complex diseases as an example (Ballouz et al, 2011).We are integrating drug databases now to enable researchers to immediately associate potential therapeutics with candidate genes. In this work presented here, we associated drugs with seven WTCCC phenotypes. For instance, Gentrepid predicted Peroxisome proliferator activated receptor delta (PPARD) as a candidate gene for Type II diabetes. Using the reference drug databases, we identified a dozen drugs that target PPARD. Drug Bank (Wishart et al, 2006) suggested 10 drugs used to treat lipid and glucose metabolic diseases, the Therapeutic Target Database (TTD) (Chen et al, 2002) indicated two drugs currently used to treat obesity and hyperlipidemia, and Pharm-GKB database (Hernandez et al, 2008) suggested two drugs used to treat prostatic neoplasms. For Carbohydrate (chondroitin 6) sulfotranferase 3 (CHST3), another Gentrepid candidate gene for Type II diabetes, Pharm-GKB suggested the same two drugs to treat prostatic neoplasms as identified for the PPARD gene. Thus, these drugs can be immediately utilized in further laboratory studies and in phase III clinical trials.

POSTER 23

GENETIC ANALYSIS OF AN EPILEPTIC SYNDROME REVEALS A MAJOR CAUSE OF A COMMON FORM OF EPILEPSY

Michael Ricos, Yeh Sze Ong, Bree Hodgson, Sarah Heron, Xeniua Ionia, Boukje de Vries, Ingrid Scheffer, Samuel Berkovic, Massimo Pandolfo, Leanne Dibbens

ABSTRACT

The majority of epilepsies are focal in origin, with seizures arising in one part of the brain. The aetiology of most focal epilepsy cases has remained unknown and are not generally thought to be genetic. We investigated the monogenic disorder 'Familial Focal Epilepsy with Variable Foci' (FFEVF) in which family members have seizures originating from different cortical regions. FFEVF shows autosomal dominant inheritance in large pedigrees. We set out to identify the gene responsible for the disorder by linkage analysis followed by exome sequencing two large families, which mapped to Chromosome 22q12. We detected DEPDC5 nonsense mutations and subsequently identified mutations in 7/8 published families with FFEVF.

Mutation analyses of DEPDC5 in focal epilepsies from smaller families, too small for a conventional clinical diagnosis of FFEVF, were performed using high-resolution melt curve analysis & verified by Sanger sequencing. Mutations in DEPDC5 account for approximately 12% of cases of familial focal epilepsy, becoming the most common 'known cause' of familial focal epilepsy. Detection of de-novo DEPDC5 mutations indicate that the gene contributes to sporadic cases of focal epilepsy, as well familial cases.

Shared homology with G-protein signaling molecules and recent findings suggest a role for DEPDC5 in mTOR signaling. Identification of DEPDC5 advances our understanding of the pathogenesis of epilepsy by implicating another new gene pathway contributing broadly to inherited focal epilepsies. Detection of DEPDC5 mutations will aid in patient diagnosis and treatment.

POSTER 24

STRATIFICATION OF INHERITED DISEASE: OVERLAPPING GENOTYPES BETWEEN KALLMANN SYNDROME (KS) AND HYPOPITUITARISM-ASSOCIATED DISORDERS HIGHLIGHT NEED FOR GENOME SEQUENCING OF KNOWN GENETIC DISORDERS

Mark McCabe, Carles Gaston-Massuet, Youli Hu, Louise Gregory, Juan-Pedro Martinez-Barbera, Pierre-Marc Bouloux, Mehul Dattani

ABSTRACT

Background and aims: PROKR2 and X-linked KAL1 are implicated in KS; combination of hypogonadotropic hypogonadism and anosmia. We hypothesised that overlapping phenotypes with congenital hypopituitarism and craniofacial/midline disorders including septo-optic dysplasia (SOD) suggest an overlap in genotypes, and thus aimed to screen a cohort of 422 patients with SOD for mutations in PROKR2 and KAL1.

Methods: Mutations in coding regions were identified by direct sequencing. Functional assays were conducted on novel mutations in vitro and the role of Prokr2 in the developing murine hypothalamo-pituitary axis was assessed by in situ hybridisation using various hormone markers in a Prokr2-/- model.

Results: Three functionally deleterious PROKR2 mutations were found in nine patients with SOD with pituitary abnormalities. Surprisingly, the mother of an affected heterozygous child was homozygous, yet unaffected. Her phenotype was supported by the largely normal Prokr2-/- mice. Additionally, three female SOD patients presented two mutations in KAL1; their mothers the unaffected carriers. Both mutations resulted in functional impairment in vitro with cellular secretion of one variant restricted.

Conclusions: PROKR2 is not causative of SOD and its role in KS needs further study given the normal fertility of the homozygous mother, who had three other children, unassisted. KAL1 is localised to the pseudo-autosomal region of the X-chromosome; potentiating a female phenotype. However, the lack of maternal phenotype is unexplained. The ambiguity in genotypes and their low incidence in our cohort (~97% wt), highlights the need for a whole genome/exome sequencing approach to identify causative/contributing genes to these conditions.

POSTER 25

RIBOSOMAL PROTEIN 6S MRNA IS A BIOMARKER UPREGULATED IN MULTIPLE SCLEROSIS, DOWNREGULATED BY INTERFERON TREATMENT, AND AFFECTED BY SEASON

Grant Parnell, Prudence Gatt, Joseph Powell, Peter Visscher, Grant Montgomery, Mark Slee, Malgorzata Krupa, Jeannette Lechner-Scott, Simon Broadley, David Booth

ABSTRACT

Background: Multiple Sclerosis (MS) is an immune-mediated disease of the central nervous system that responds to therapies targeting circulating immune cells.

Objective: To test if the T cell activation gene expression pattern (TCAGE) we had previously described from whole blood was replicated in an independent cohort.

Methods: We used RNA-seq to interrogate the whole blood transcriptomes of 72 individuals (40 healthy control, 32 untreated MS). A cohort of 862 control individuals from the Brisbane Systems Genetics Study (BSGS) was used to assess heritability and seasonal expression. The effect of interferon beta (IFNB) therapy on expression was evaluated.

Results: The MS/TCAGE association was replicated and rationalised to a single marker, ribosomal protein S6 (RPS6). Expression of RPS6 was higher in MS than controls ($p<0.0004$), and lower in winter than summer ($p<4.6E-06$). The seasonal pattern correlated with monthly UV light index ($R=0.82$, $p<0.002$), and was also identified in the BSGS cohort ($p<0.0016$). Variation in expression of RPS6 was not strongly heritable. RPS6 expression was reduced by IFNB therapy.

Conclusions: These data support investigation of RPS6 as a potential therapeutic target and candidate biomarker for measuring clinical response to IFNB and other MS therapies, and of MS disease heterogeneity.

RNA/NONCODING

POSTER 26

ASSESSMENT OF CARDIAC MICRORNA HIGH THROUGHPUT SEQUENCING DATA SETS GENERATED FROM RNA OF VARYING QUALITY

David Humphreys, Kavitha Muthiah, Liza Thomas, Peter Macdonald, Christopher Hayward

ABSTRACT

Profiling microRNAs (miRNA) by high throughput sequencing (miRNA-Seq) provides quantification of expression level and reveals the complexity of processing variants. We recently published a software tool called miRspring which completely integrates miRNA-Seq data and sophisticated analysis tools into a HTML research document, and is functional without internet connectivity. Our initial use of the miRspring software was to establish quality parameters from 73 publically availably human miRNA-Seq data sets. In this study we have used the miRspring document to analyse miRNA-Seq data generated from cardiac biopsies taken from human patients at the time of LVAD (left ventricle assist device) implant and subsequent explant. The quality of RNA extraction from biopsies varied considerably with some preparations being significantly degraded (i.e. RIN scores < 6). We find that the established quality parameters identified from the publically available data sets were maintained in majority of data sets generated from degraded RNA. In particular there was no significant decrease in miRNA length and no appreciable increase in isomiRs or non-canonical processing across the data sets. This indicates that the degraded RNA preparations had intact miRNAs and are suitable for further analysis. From this analysis we find specific cardiac miRNAs to be upregulated at the time of explant indicating that transcriptional remodeling was occuring in hearts with LVAD support. We conclude that low quality RNA material can still be used for miRNA profiling if the appropriate quality control measures are implemented, and the miRspring software is proving to be a valuable tool for this analysis.

POSTER 27

LONG NONCODING RNAS IN FORMATION OF MEMORY, THROUGH LONG-TERM POTENTIATION

Jesper Maag, Debabrata Panja, Marcel Dinger, Clive Bramham, Karin Wibrand

ABSTRACT

Long-term potentiation (LTP) is a process of strengthening of signals across the synaptic cleft, and the process is believed to be involved in the formation of long-term memory. The protein-coding mechanisms of this phenomenon have been widely studied, however, the involvement of non-coding elements in the formation of increased synaptic connectivity has hitherto not been reported. Long noncoding RNAs have been shown to exert a wide selection of biological functions in the cell, and they are expressed in a cell specific manner.

To explore the role of ncRNA in LTP, we stimulated the dentate gyrus region in the right brain hippocampus in four anesthetised male Sprague-Dawley rats with a high frequency stimulation protocol widely used to induce LTP. RNA was extracted two hours after stimulation from the dentate gyrus, and total RNA sequencing was conducted on a HiSeq 2000. The left, unstimulated, dentate gyrus was used as transcriptional control.

In accordance with previous studies, various protein-coding genes known to be involved in neurogenesis and LTP were shown to be differentially expressed between the HFS stimulated side and control. Novel noncoding transcripts expressed in cis to known neuroplasticity genes exhibited elevated expression in stimulated samples. Moreover, miRNA not previously affiliated with neurogenesis was found to have an upregulated expression pattern after stimulation. Functional studies will be pursued to characterise the molecular mechanisms of these noncoding transcripts.

These data suggests novel evidence that noncoding RNAs function in memory formation and synaptic plasticity.

POSTER 28

THE BIOINFORMATICS BEHIND NOVEL TARGET DISCOVERY USING HIGH THROUGHPUT FUNCTIONAL GENOMICS

Cathryn M Gould, Kaylene J Simpson

ABSTRACT

RNA interference (RNAi) is an endogenous eukaryotic mechanism of gene regulation which inhibits gene expression. Harnessing this mechanism synthetically, RNAi can be used for large-scale genome-wide screens that systematically knockdown each gene in the cell.

Such high throughput screens require automated liquid handling robotics, plate readers, high content microscopy and bioinformatics. The technology platforms available in the VCFG and accessible to researchers Australia-wide are: siRNA, miRNA, shRNA and, in 2014, long non coding RNA (lncRNA).

Large-scale screens can address many biological questions including understanding the molecular basis of tumourigenesis, identifying therapeutic targets, and basic biology for discovery of genes that regulate cellular processes. Since RNAi screens are run over many weeks, we have developed a bioinformatics pipeline to ensure basic QC measures of robustness and performance are maintained from week to week to permit legitimate biological conclusions and identification of hit candidates at the end of the screen. Integration of transcriptome profiling data and computational approaches to address potential off-target effects are also included. Following the primary screen, users perform a secondary screen and subsequent downstream analysis of the validated high confidence hits can begin in earnest.

The downstream analysis embraces a more systems biology approach to integrate the high confidence hits with other functional datasets, including model organism resources, interaction databases, disease repositories, as well as data-driven predictive tools in the case of poorly characterized candidates. Another burgeoning source of potential functional association info from screens is the quantitative multivariate cellular analysis from high dimensional microscopy image-derived datasets.

POSTER 29

TO BE, OR NOT TO BE: MIRNA PROFILE OF A NOVEL STABLE STATE OF INDUCED PLURIPOTENT STEM CELLS

Hardip Patel, Jennifer Clancy, Nicole Cloonan, Andrew Corso, Mira Puri, Pete Tonge, Andras Nagy, Thomas Preiss

ABSTRACT

Pluripotent embryonic stem cells (ESCs), derived from the inner cell mass of a mouse blastocyst, possess the ability to differentiate to all three germ layers, contributing to embryonic development and the generation of adult animals. Somatic cells can be induced to acquire a pluripotent stem cell (iPSC) fate through the over-expression of key transcription factors. Like ESCs, iPSCs can fulfill the strictest of all developmental assays: the production of completely iPSC-derived embryos and mice. An international consortium has gathered global proteomics, long and short RNAseq and epigenetics data to describe key changes during the reprogramming of mouse fibroblasts to iPS cell state. The consortium has also discovered a novel stable state of iPSCs expressing high levels of key transcription factors (Oct4, Sox2, Klf4, and c-Myc) compared to the ESCs. We analysed miRNA profiles to identify key molecular signatures differentiating canonical iPSC state and a novel iPSC state. We discovered that miRNA expression profiles in both states are very different. For example, the novel state displays low expression of many core miRNA-mediators of pluripotency, including the miR-302 cluster. Nevertheless highly expressed miRNAs in this novel state possess the same target spectrum to promote reprogramming pathways (e.g. cell cycle). Additionally, we also show that the novel state is associated with increased expression of a small RNA species from the mitochondria. I will also describe the rapid changes in miRNA populations during the first four days after the induction of expression of four transcription factors (Oct4, Sox2, Klf4, and c-Myc).

POSTER 30

A COMPARISON OF PLASMA MIRNA LEVELS IN INSULIN SENSITIVE, INSULIN RESISTANT AND TYPE 2 DIABETIC OBESE INDIVIDUALS VERSUS LEAN CONTROLS: DIFFERENTIAL MIRNA LEVELS BETWEEN GROUPS

Donia Macartney-Coxson, Angela Jones, Miles Benton, Richard Stubbs

ABSTRACT

Type 2 diabetes is a consequence of insulin resistance and strongly associated with obesity. However, not all obese individuals develop type 2 diabetes and some obese individuals are metabolically healthy. MiRNAs are small, non-coding RNA molecules which inhibit translation and/or direct mRNA degradation; they are reported to be involved in the pathogenesis of complex diseases including obesity and type 2 diabetes. MiRNAs are increasingly recognized as potential disease biomarkers, and their transportation in plasma and delivery to recipient cells where they can modulate target mRNAs has been demonstrated.

Therefore, we carried out a pilot study in 45 obese individuals who were insulin sensitive, insulin resistant or had type 2 diabetes (n=15 each group) and compared miRNA levels with those of 12 lean controls. Using an Exiqon QPCR array panel we assayed 168 miRNAs. Two miRNAs (mir103a, mir-425) were identified as normalization controls as they showed minimum variability among all samples irrespective of group. Using normalized (deltaCt) miRNA levels we compared the median level of each miRNA for each clinical group with that of the controls using the Kolmogorov-Smirnov test. Most miRNAs were more abundant in obese individuals than controls. The levels of 85/168 miRNAs were significantly different for at least one patient group in comparison to controls after multiple testing adjustment (adjusted $p<0.05$). These included both novel observations and concordance with the literature. This study provides a strong basis for future work aimed at investigating differential miRNA levels in larger independent cohorts and understanding the role miRNA play in disease.

The project was supported by ESR Pioneer Funding. Miles Benton is supported by a Corbett Postgraduate Research grant.

PROTEOMICS

POSTER 31

INTEGRATING FUNCTIONAL SOMATIC MUTATION BURDEN WITH GENE EXPRESSION

Shila Ghazanfar

ABSTRACT

Tumour progression is believed to be characterised by higher rates of somatic mutation compared to surrounding normal tissue (Gerlinger et al, New England Journal of Medicine, 2012). The challenge is differentiating between mutations that are responsible for the tumour progression, amongst a background of other mutations, both somatic and germline, and genetic differences that are not of interest in regards to the disease.

Currently, somatic mutations are typically detected through whole exome sequencing. It has been recently demonstrated that these somatic mutations, and in particular the functional mutation burden of the corresponding genes is well represented in melanoma (Hodis et al, *Cell*, 2012). However for these genes to have a phenotypical impact, they require transcription to RNA. We integrate information on functional mutation burden of such genes with expression levels measured by RNA-Seq, thereby integrating two data sources to predict patient survival outcome. We compare existing methods for detecting somatic mutations (Cibulkis et al, *Nat Biotech*, 2013; Evans et al, *NanoBioscience*, 2013) and determining functional mutation burden (Lawrence et al, *Nature*, 2013; unpublished **http://www.broadinstitute.org/cancer/cga/invex/**), as well as methods for integrating these with gene expression data.

POSTER 32

SYSTEMS ANALYSIS OF INTER-TISSUE SIGNALING GENE REGULATION IN ORGANOGENESIS

Joshua Ho

ABSTRACT

Development of many organs depends on the sequential and reciprocal exchange of various signaling molecules between juxtaposed epithelial and mesenchymal tissue, but the detailed mechanism controlling these epithelial-mesenchymal (E-M) interactions remains unknown. We used the developing mouse molar tooth as a model to decipher the gene regulatory network (GRN) that underlies the complex E-M signaling dynamics during organogenesis. Through the extensive profiling of over 100 microdissected embryonic mouse dental E-M tissues, mutant tissues and signaling molecule treated tissues, our analysis reveals two surprising new insights: (1) Despite the reciprocal exchange of signaling molecule expression, the overall temporal genome-wide expression change in E-M tissues is highly concordant, and (2) among key signaling pathways, the Wnt and Bmp pathways are the primary driver of odontogenesis. We developed a statistical approach to integrate our expression datasets with over 1,000 pieces of perturbation evidence from the literature to generate an inter-tissue GRN for early odontogenesis. Within this GRN, we identified a novel feedback circuit that connects the Wnt and Bmp pathways across the E-M tissue compartments through the action of the Wnt and Bmp4 ligands. Moreover, our inter-tissue Wnt/Bmp circuit was validated with two sets of in vivo mouse genetic crosses designed to "short-circuit" or to "break" the feedback circuit. Computer simulation demonstrates that the circuit structure alone can explain the observed signaling molecule expression dynamics in wild-type and mutant mice. This work illustrates how complex signaling dynamics like the E-M interactions in organogenesis represent an intrinsic property of the underlying GRN structure.

POSTER 33

MULTIPLEXED ANALYSIS OF IN VITRO RECONSTITUTED HUMAN PROTEIN COMPLEXES

Kirill Alexandrov, Yann Gambin, Emma Sierecki, Loes Stevers, Nichole Giles, Mark Polinkovsky, Mehdi Moustaqil, Sergey Mureev, Wayne Johnston

ABSTRACT

The ability to replicate the building blocks of life such as DNA and proteins represents the core technologies of Life sciences. The exponential increase in the number of sequenced genomes has focused attention on how best to produce, study and modify the encoded gene products. While the structural and functional information is encoded in a single protein molecule, accessing this information is technically and economically challenging. This constitutes a critical technological bottleneck that determines the pace of progress in many areas of biology. The problems become particularly aggravating when analysis of complex protein machines is attempted. We developed an integrated approach for rapid and flexible production, derivatization and analysis of proteins and protein complexes staring from human ORFeome libraries. By integrating the cell-free protein production, multiplexed handling and single molecule fluorescence spectroscopy we were able to in vitro reconstitute several multisubunit eukaryotic protein complexes such as Cavin, HOPS,BBSome and Mediator that were previously refractory to reconstitution and analysis. Reconstitution of the complexes and analysis of subunit interactions is carried out within hours using a combination of spectroscopic and bead- based interaction assays. The approach allows rapid extraction of information on oligomerisation state, interaction affinities and the stabilities of individual subunits and the entire complexes. The implications of the developed technology for protein research will be discussed.

POPULATION VARIATION

POSTER 34

PEDIGREE STRATIFIED GENOME-WIDE MINING FOR PAIRWISE SNP INTERACTIONS IN TYPE II DIABETES

Timothy Hancock, Justin Bedo, Gerard Wong, Melanie A. Carless Joanne E. Curran, John Blangero, Peter Meikle, Adam Kowalczyk

ABSTRACT

Type II Diabetes (T2D) is a complex metabolic disease which is thought to be triggered by lifestyle and environmental factors and can be characterized by numerous defective physiological processes such as insulin resistance and loss of pancreatic beta cell function. The predominant lifestyle component of T2D means that to identify loci of genomic predisposition involves exhaustive analysis of the interactions between an individual's genetic information in the context of confounding lifestyle and environmental factors. Evidence for the interaction between genomic predisposition, lifestyle and environment is clear from previous studies which found the risk of T2D to increase in siblings by 3.4-3.5 (OR: odds ratio) if one parent is affected and by 6.1 (OR) if both parents are affected. In stark contrast, the results of several GWA studies into T2D have yielded low power results with individual SNPs providing minimal predictive value. In this research, we present an application of our Genome Wide Interaction Search (GWIS) platform for fast genome-wide identification of the most significant pairwise SNP interactions in the context of the family structure within the San Antonio Family Heart Study. To factor in family structure within the GWIS platform we perform a stratification analysis by considering all relatives within each individual's descent vectors over all pedigrees. Our results show that through stratification, we gain power in detecting many more loci significantly associated with T2D. Furthermore, by considering SNP interactions we identity individual pairs that perform as well as current multiple single SNP models.

POSTER 35

STABILITY OF BIVARIATE GWAS BIOMARKER DETECTION

Justin Bedo, David Rawlinson, Cheng Soon Ong, Ben Goudey, Qiao Wang, Adam Kowalczyk

ABSTRACT

Given the difficulty and effort required to confirm candidate biomarkers detected in genome-wide association studies (GWAS), there is no practical way to definitively filter false positives, however we can measure the replicability of candidate biomarkers. Recent advances in algorithmics and statistics now enable repeated exhaustive search for bivariate features within a practical amount of time and computational resources. We used cross-validation to evaluate the stability of three feature selection methods. We performed 10 trials of 2-fold cross-validation of exhaustive bivariate analysis on seven Wellcome–Trust Case–Control Consortium GWAS datasets, comparing the traditional ^{2} statistic, the high-performance GBOOST method and the recently proposed GSS statistic. To compare the incomplete lists of ranks, we propose an extension to Spearman's correlation. The extension allows us to consider a natural threshold for feature selection where the correlation is zero.

We found that stability between ranked lists was superior for GSS in the majority of diseases, compared to ^{2} and GBOOST. A more thorough analysis of the correlation between SNP-frequency and univariate ^{2} score demonstrated that the ^{2} test is highly confounded by main effects: SNPs with high univariate significance replicably dominate the ranked results. We show that removal of the univariately significant SNPs improves ^{2} replicability but risks filtering pairs involving SNPs with univariate effects.

Finally, we investigate the relationship between stability, FWER, and FDR. We show that FDR and FWER can exhibit wildly varying numbers of significant pairs. ZIC does not have this undesirable behaviour.

POSTER 36

AN EXPLORATION OF MUTATION DETECTION USING HALOPLEX TARGETED SEQUENCING

Simon Sadedin, Alicia Oshlack

ABSTRACT

Targeted High Throughput Sequencing (HTS) is an important technology for enabling cost effective resequencing in clinical and research contexts. Recently, Agilent has introduced the HaloPlex technology as a cost effective and labour saving alternative to other targeted capture techniques. However in order to achieve these benefits, HaloPlex uses enzymatic shearing to fragment input DNA, which results in a starkly different profile of read coverage. In this work we explore the differences between HaloPlex enrichment and other more conventional targeted sequencing methods. We investigate coverage variability, rates of off target capture, and variant calling performance. Our results show that HaloPlex data benefits from special handling and we provide recommended step for analysis. We further demonstrate that Haloplex has some unique advantages that may allow increased performance in applications such as mutation detection if exploited correctly.

POSTER 37

A BIOINFORMATIC PIPELINE FOR SNP CALLING IN NEXT GENERATION GENOTYPING-BY-SEQUENCING

David Kainer, Amanda Padovan, Justin Borevitz, William Foley, Carsten Kulheim

ABSTRACT

The ever-increasing throughput and downward trending cost-per-base of DNA sequencing technologies has unlocked enormous opportunity for population genetics studies. At the same time, however, it has created difficulties in handling the larger scales of sequence and variant data, as well as challenges due to sequencing error and how to deal with the presence of rare variants within populations. This study investigates new bioinformatics techniques to confidently call SNPs, account for sequencing error and rare variants, filter data and produce inputs for downstream population genetic analyses such as diversity, structure and genotype-phenotype association.

This study is currently in its early stages, though techniques are being developed for an active population genetic study of Yellowbox Eucalypts (Eucalyptus melliodora), an iconic species that is highly valued for timber and honey production. Its fragmented habitat is critically endangered due to land clearing for farming. The goal of the Yellowbox study is to identify and type thousands of SNP markers to determine rare individual and population variation, outcrossing rate, and population structure across the geographic range. Two hundred individuals from five populations have been sampled and their genomes have been sequenced after targeted complexity reduction through restriction digestion. A new bioinformatics pipeline from raw reads to allele calling has been developed. These methods should help produce greater understanding of the population structure and improve conservation efforts of E. melliodora while also improving the efficiency and reliability of future population genetics studies.

POSTER 38

GENE EXPRESSION VARIABILITY AS A UNIFYING ELEMENT OF THE PLURIPOTENCY NETWORK

Elizabeth Mason, Jessica Mar, Dr Andrew Laslett, Martin Pera, John Quackenbush, Associate Professor Ernst Wolvetang, Christine Wells

ABSTRACT

The term "pluripotent stem cell" encompasses a range of cellular phenotypes, with the dual ability to self renew and generate cells of the three embryonic germ layers. Phenotypic heterogeneity is a hallmark of pluripotent stem cell populations, and is a function of the oscillatory behaviour of Oct4 and Nanog, the master-regulators of the pluripotency gene regulatory network. Here we present evidence that variability in the expression of such genes provides a new metric with which to characterize human pluripotent stem cell populations. We conclude that even in homogenous populations, the underlying genetic network displays high and low-variance elements. We show that those genes with the least expression variability are also highly connected, demonstrating a high level of regulatory constraint of the core network elements. Oct4 and Nanog exhibit different patterns of expression

variability, indicating changes in the stabilisation of expression of these key factors in different stem cell phenotypes. Expression variability provides insight into the heterogeneity of a pluripotent stem cell population, as well as insight into the regulation of key network elements within a well-defined pluripotency network.

POSTER 39

DEFINING A HOMOGENEOUS PROGENITOR POPULATION

Jessica Schwaber, Lars Nielsen

ABSTRACT

Hematopoiesis is typically described as a deterministic hierarchy of transient cell populations defined by cell surface markers. Lineage commitment, however, is speculated to be stochastic. Commitment of the common myeloid progenitor (CMP) to downstream linages is governed by lineage-affiliated transcription factors, GATA_1/PU.1. We have formulated a stochastic model based on GATA_1/PU.1 signaling and confirmed that the model is capable of driving fate decision. We are exploring an expanded model that incorporates additional factors linked to external signaling events with the potential to modulate the stochastic switch and guide differentiation rather than merely permissively rescue already committed cells.

We are currently developing ways of parameterizing and validating the model using single cell assays. Modeling with a complex initial population requires the study of a very large number of individual cells, something difficult with current single cell transcriptomics techniques. Therefore, we need synchronicity in our initial population to monitor differentiation.

The CMP population spans a quiescent to a lineage primed progenitor state. A key challenge is to identify factors that define a narrower CMP phenotype extractable from the broader population. Time series preculture experiments were conducted for CD34+ UCB isolated cells to increase the cycling cell population. Preliminary data for single cell model validation was collected from these cultures by isolating cells sorted off CMP defining cell surface markers and Ki67 transcript levels. Single cell CFU assays were used to confirm growth of multi-lineage colonies from the defined/isolated population and qRT-PCR were run verify for a priori determined CMP defining transcript targets.

COMPUTATIONAL BIOLOGY

POSTER 40

COMBINING TRANSCRIPTOME ASSEMBLIES FROM MULTIPLE DE NOVO ASSEMBLERS TO GENERATE FULL LENGTH RNA SILENCING GENE TRANSCRIPTS IN NICOTIANA BENTHAMIANA

Kenlee Nakasugi, Peter Waterhouse

ABSTRACT

In an effort to produce an assembly that contained (but not limited to) full length RNA silencing gene transcripts to facilitate more informative first pass searches, and to increase the chances of finding paralogous transcripts while limiting redundancy, we have combined the sequences from multiple assemblies generated by four popular de novo transcriptome assemblers: Trans-Abyss, Trinity, Soap-denovo-trans and Oases. The subject organism is Nicotiana benthamiana, an allopolyploid plant.

Two methods were implemented to reduce the redundancy of combined assemblies - a clustering based approach (TGI clustering tools), and one that selects a 'best set' of mRNA sequences rather than producing longest possible transcripts (EvidentialGene pipeline). Metrics used to assess the quality of assemblies include the average length of the 1000 longest proteins, average bit-scores from blast comparisons against reference databases, and feature response curves.

By combining the output of different assemblers by varying k-mer sizes and input read counts, we were able to detect all 35 query RNA silencing gene transcripts as full length from simple first pass blast searches. Only 24 RNA silencing transcripts could previously be detected as complete using one assembler. While the TGI clustering tool could produce longer transcripts, the average bit-scores of blast searches and feature response curves show that the Evidential Gene pipeline produced higher quality assemblies.

By using a combined assemblies approach as recommended by the EvidentialGene pipeline, one can recover more completely assembled transcripts while limiting redundancy and maximising the quality of the assembly.

POSTER 41

WHAT TO DO WHEN IT'S ALL RELATIVE: HOW TO ANALYSE RELATIVE ABUNDANCES WITHOUT BEING MISLED

David Lovell, Vera Pawlowsky-Glahn, Juan Jos Egozcue

ABSTRACT

Most measurement processes in molecular bioscience yield information about relative abundances only.

This can be a result of sample preparation steps where nucleic acids are brought to a specified concentration before measurement; or by the measurement step itself, as in the case of next-generation sequencing.

Awareness is gradually growing that these data need special treatment and interpretation but it is not yet widely appreciated just how misleading traditional statistical approaches (such as correlation) can be.

Using data on absolute levels of yeast gene expression (i.e., mRNA copies per cell) over a 16-point time course experiment, we demonstrate why the concept of differential expression is challenging to interpret with relative abundances, and why correlation is an inappropriate measure of association for this kind of data.

We present a simple, well-principled measure of association that can be used with confidence on relative abundances: proportionality. We give a straightforward graphical presentation of a statistic that can be used to assess the degree of proportionality between two sets of values, and then show how that can be plugged into analysis strategies that are familiar in molecular bioscience, including networks of association and clustered heatmaps.

This approach gives bioscientists the means to analyse (and reanalyse) data generated in transcriptomics, proteomics, metagenomics and elsewhere, secure in the knowledge that the inferences made from the data are not artefacts of its relative nature.

POSTER 42

EXPLORING DIFFERENCES IN CHROMATIN ORGANIZATION WITH HI-C DATA

Aaron Lun, Gordon Smyth

ABSTRACT

The organization of chromatin in the nucleus plays an important role in the regulation of eukaryotic transcription. Genes are activated by enhancer interactions and silenced by packaging into heterochromatin. A new technique called Hi-C (chromatin conformation capture with high throughput sequencing) allows examination of chromatin structure by capturing all pairwise interactions between genomic loci. The data obtained from this procedure can provide some insights into the molecular mechanisms used to control gene expression. Here I'll present an overview of some strategies for alignment, statistical analysis and visualization of Hi-C data, using some publicly available datasets for demonstration.

POSTER 43

IMPROVED GENOTYPE CALLING FOR RARE VARIANTS

Ruijie Liu, Zhiyin Dai, Rafael A. Irizarry, Matthew E. Ritchie

ABSTRACT

Over the past decade, SNP genotyping microarrays have revolutionized the study of complex disease. The current range of commercially available genotyping products contain extensive catalogues of rare variants (those with frequency less than 5% in a population). Existing SNP calling methods have difficulty with rare variants, as the models rely on each genotype having a minimum number of observations to ensure accurate calls. We have developed a new unsupervised method for converting raw intensities into genotype calls that aims to overcome this issue. Our method allows a variable number of clusters (1, 2 or 3) for each SNP that is predicted using the available data, and offers improved performance over current approaches. We test our method against four competing genotyping algorithms on several Illumina data sets that include samples from the HapMap project where the true genotypes are known in advance.

POSTER 44

NEXT-GENERATION READ ALIGNMENT

Wei Shi

ABSTRACT

The conventional "seed-and-extend" read mapping paradigm, adopted by most of the existing read aligners, has been challenged in the past few years, with the rapid evolution of sequencing technologies. The advance of sequencing technologies features rapid increase of read length. However, the seed-and-extend paradigm has a very limited scalability in dealing with read length increase because of the inefficiency of its extension step. Also, it has much difficulty in mapping reads spanning a large genomic region, such as exon-spanning reads.

In this talk, I will describe a novel read mapping paradigm called "seed-and-vote". It operates by extracting from each read many short equi-spaced seeds (16bp mers), mapping them to the reference genome by a highly efficient hashing algorithm, and finally determining the read mapping location via voting of the selected seeds. No mismatches are allowed in seed mapping to ensure a highly accurate read mapping is achieved.

This paradigm largely avoids the extremely computing-intensive "extend" step of the seed-and-extend paradigm by seed voting, achieving a highly efficient read mapping and also a high scalability with read length increase. It uses perfectly matched flanking seeds to accurately detect indels and it is particularly powerful in mapping exon-spanning reads, by utilising many short seeds to locate the exons spanned by such reads. The seed-and-vote paradigm is implemented in the *Subread* aligner. *Subread* is found to be an order of magnitude faster on average, and more accurate, than popular aligners such as Bowtie2, BWA, Novoalign and TopHat2.

POSTER 45

A SYSTEMATIC EVALUATION OF OBSERVED ALLELIC SPECIFIC EXPRESSION ACROSS TWO HUMAN TISSUES

David Wood, Katia Nones, Anita Steptoe, Angelika Christ, Ivon Harliwong, Felicity Newell, Tim Bruxner, David Miller, Nicole Cloonan, Sean Grimmond

ABSTRACT

Allelic Specific Expression (ASE) is a process where RNA transcribed from parental alleles differs in abundance or form. ASE includes near-complete suppression of expression, as in X-inactivation and imprinting, or imbalanced expression, through polymorphisms modulating transcription factor binding affinity or splicing patterns. Studying ASE provides a mechanism to identify expression related disease variants, understand imprinting mechanisms, and link genotype to phenotype.

ASE is typically quantified by calculating the proportion of the reference or variant allele at expressed heterozygous SNPs, or the probability of this observed ratio. This approach, if used in isolation, is subject to increased false positive rates caused by alignment bias, alignment error and sampling issues. In this study, we have performed an evaluation of observed ASE in two human tissue samples (brain and liver) from different individuals, and developed a pipeline to robustly detect ASE at individual SNPs, and across entire genes and regions. We detect ASE by integrating the results of multiple RNA-seq libraries with matched genotyping array results, whole genome sequencing data, imputed SNPs, and computationally phased haplotypes. We systematically evaluate SNPs for bioinformatic bias, biological ASE causes such as copy number variation, known imprinting status or eQTL status, and verify observed events using independent cDNA and gDNA amplicon sequencing. Our results identify known ASE regions, uncover several novel ASE genes, suggest that ASE is not as widespread as previously thought, and demonstrate the importance of statistical and bioinformatic rigour and independent sequencing verification when testing for ASE.

POSTER 46

FINDING MSC GENE EXPRESSION SIGNATURE REGARDLESS OF THE TISSUE OF ORIGIN

Florian Rohart, Celena Heazlewood, Kim-Anh Le Cao, Christine Wells

ABSTRACT

Gene expression signatures can be found in most experimental series, but signatures which are reproducible across many experimental series are difficult to find. This is partly because of the technical barriers preventing large-scale meta-analysis of different expression experiments generated across different laboratories, or on different experimental platforms. We have developed a new gene-signature tool which works across multiple expression series to identify a robust mesenchymal cell signature. The tool has been implemented using 680 samples of different cell types from the large-scale Stemformatics stem cell expression database (www.stemformatics.org). Our approach integrates the YuGene cross-platform normalisation method, which is effective at reducing experimental batch effects and integrating data from different commercial platforms, while retaining genuine biological variation between samples. A sparse partial least-squares –differential analysis (sPLS-DA) method was used to identify an accurate classifier of mesenchymal cell types. A stability analysis provided an average of more than 96% classification accuracy as well as a highly stable signature consisting of 10 genes. Our approach highlights the robustness of biological signatures when experimental variables such as platform or batch can be reduced. The software is available in R/Cran, and can be implemented from mixomics (**http://mixology.qfab.org/**) and the stem cell signature can be visualised at Stemformatics.

POSTER 47

A WORKFLOW TO INCREASE VERIFICATION RATE OF CHROMOSOMAL STRUCTURAL REARRANGEMENTS USING HIGH THROUGHPUT NEXT GENERATION SEQUENCING

Quek K, Nones K, Patch A-M, Cloonan N, Newell F, Fink L, Christ A, Miller D, Bruxner T, APGI, Waddell N, Biankin AV, Pearson J, Grimmond S

ABSTRACT

Somatic rearrangements are commonly found in human cancer genomes and contribute to the progression and maintenance of cancers. As part of the International Cancer Genome Consortium (ICGC), QCMG is sequencing 350 pancreatic, 150 ovarian primary tumours and matched DNA to identify underlying genetic changes in the cancers. Here we describe a high throughput workflow for the verification of somatic rearrangements.

Conventionally, the verification workflow comprises many manual steps and capillary Sanger sequencing which is time consuming and labour intensive especially when verifying a large number of rearrangements across cancer patients. To increase the verification throughput, we proposed a workflow which utilises in-house bioinformatics tools to link the laboratory processes. Genomic locations of somatic rearrangements are identified by an in-house tool, qSV (using discordant mapped read pairs approach) and primers are automatically designed by the qPrimer tool for amplicons up to 3kb for each rearrangement. Visual analysis of the PCR products confirms and verifies the somatic nature of the variant and single band products are pooled and fragmented for Ion Torrent or MiSEQ sequencing. The resulting short sequencing reads are assembled into long consensus contigs by qAssemble and an automated BLAT performed to resolve the breakpoints to base level.

To evaluate the reliability and efficiency of this workflow, we sequenced somatic rearrangements using both capillary and Ion Torrent sequencing. The comparison of these methods revealed that whilst both methods were able to resolve 98% of the rearrangement breakpoints the automated high throughput workflow will process hundreds of events within 5 days as compared to the conventional workflow which takes weeks.

In conclusion, we demonstrate an efficient next generation sequencing based method to verify somatic rearrangements and identify their breakpoints. The workflow is reproducible and comparable to the conventional low throughput method.

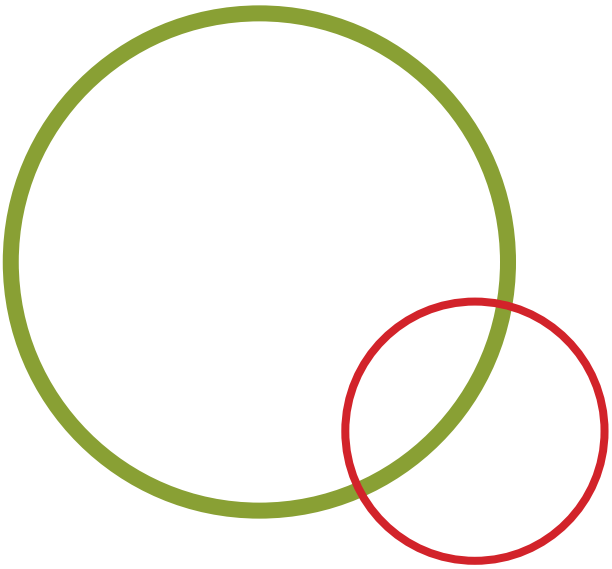
POSTER 48

MILKING THE PIGEON

Tamsyn Crowley, Meagan Gillespie, Volker Haring, Paul Monaghan, John Donald, Rob Moore, Kevin Nicholas

ABSTRACT

The pigeon is one of only a few birds that produce a nutrient substance 'crop milk' to feed their young. This nutrient substance is produced in the crop by both male and female birds and has been shown to have functional similarities with mammalian milk. As with mammalian milk, crop milk is essential for squab growth, providing both nutritional and immune benefits. We have spent the last few years studying this interesting biological phenomenon employing many different tools, including bioinformatics. Until recently there was little genomic information available, hence we have utilised bioinformatics and experimental biology in order to gain an insight into the production and benefits of pigeon crop milk.



POSTER 49

PHOSPHORYLATION SITE PREDICTION

Ralph Patrick, Kim-Anh Le Cao, Bostjan Kobe, Mikael Boden

ABSTRACT

The determinants of kinase-substrate binding can be found both in the substrate sequence, and the surrounding cellular context. Cell cycle progression, interactions with mediating proteins and even prior phosphorylation events are necessary for kinases to maintain substrate specificity. While much work has focussed on the use of sequence based methods to predict phosphorylation sites, there has been very little work invested into the application of systems biology to understanding phosphorylation. However, the generality of many kinase binding motifs means that sequence methods for predicting kinase binding sites are susceptible to high false-positive rates.

We present here a systems biology model that takes into account protein-protein interaction information, and protein abundance data across the cell cycle to predict kinase substrates for 19 kinases. The model shows high accuracy for substrate prediction (with an average AUC of 0.84) across the 19 kinases tested. When using the model to complement sequence based kinase-specific phosphorylation site prediction, we find that the additional information can increase the prediction accuracy of sequence based methods by over ten-fold at low false positive levels. Our results demonstrate that a systems approach to phosphorylation can account for the short-falls in sequence information and provide a robust description of the cellular events that regulate kinase binding.

APPLIED GENOMICS

POSTER 50

STEMFORMATICS: STEM CELL GENE EXPRESSION

Rowland Mosbergen, Othmar Korn, Jarny Choi, Nick Seidenman, Nick Matigian, Christine Wells

ABSTRACT

Stemformatics.org was built as a collaboration platform for Stem Cells Australia, for the analysis, visualisation and sharing of multiple omic datasets. It was designed for stem cell researchers with minimal experience in bioinformatics, so needed to provide an intuitive, gene-centric search interface, transparent access to primary references, and high-quality graphics. Analyses were implemented from the most common questions asked of the bioinformatics team. It was developed from the recognition of the value of published gene expression signatures, and acknowledged the difficulty that many stem cell researchers had to identify, search, compare, extract or analyse this data in generic array repositories.

With 60 public and 50 private datasets, there are a number of take home messages that our experiences and thoughts in dealing with end users have provided. We have a simplified, semi-automated data import pipeline that incorporates QC metrics from statistical and empirical perspectives, and which requires minimal manipulation of the primary data by the end user. We provide users with easy to use functionality to answer the most commonly asked questions, use interactive graphics to show complex data in multiple ways including co-visualisation of different datasets, and identify that users want rapid return of relevant information without having to navigate through too many areas. We work closely with collaborators to tailor specific tasks to their needs, and provide options for securely sharing datasets, genelists and analyses between linked collaborative groups.

POSTER 51

POST-LIGHT SEQUENCING: PROGRESS AND PROSPECTS

Lauren Bragg, Glenn Stone, Margaret Butler, Philip Hugenholtz, Gene Tyson

ABSTRACT

In early 2011, the Ion Torrent Personal Genome Machine (PGM) debuted as the first commercial sequencing technology to detect sequencing reactions via changes in pH rather than emitted light. By avoiding expensive photo-receptive equipment and artificial reagents the PGM is currently the only sequencer platform under $100K, and offers the cheapest sequencing run of the high-throughput platforms. Combined with rapid improvements in read-length and chip density, the PGM platform is a highly appealing technology for a range of applications, including de novo assembly, metagenomics, amplicon sequencing and transcriptomics. However, like all sequencers, there are caveats. In this presentation I will explore the biases, errors and other artifacts in PGM data that can influence your research outcomes.

POSTER 52

A VIRTUAL APPLIANCE APPROACH TO DELIVERING BIOINFORMATICS PIPELINES: AN EXAMPLE WITH TREVA A TARGETED RESEQUENCING ANALYSIS PIPELINE DEVELOPED AT PETER MAC

Jason Li, Maria Doyle, Isaam Saeed, Richard Tothill

ABSTRACT

Setting up a bioinformatics pipeline for analysis of high-throughput sequencing data (HTS) is a challenging and laborious activity. The need to set up and maintain individual software dependencies (such as MySQL and Perl/Python) is complicated by compatibility issues between software versions, packages, libraries and different operating system requirements. These issues are then confounded by the myriad of available software packages that address similar bioinformatics problems, of which some are based on preliminary research or are minimally tested, consequently affecting the robustness and reliability of a pipeline. Furthermore, a pipeline is often tied to the underlying

hardware and operating system, making it extremely difficult to share and deploy the pipeline at a different site. Many laboratories do not have the resources to invest in tackling these issues. This is especially true for clinical and pathology laboratories, which are increasingly adopting HTS technologies but unfortunately lack bioinformatics resource.

We propose the use of virtual machines (VM) to distribute bioinformatics pipelines, alleviating the need to configure and manage both hardware and software dependencies and requirements. We have built a VM called TREVA (Targeted REsequencing Virtual Appliance) to house our user-friendly whole-exome/targeted resequencing analysis pipelines developed at Peter Mac. The analyses that are supported include: single-nucleotide polymorphism and insertion/deletion variant calling, copy number analysis for targeted resequencing, cohort-based analyses such as pathway and significantly mutated genes analyses.

As a virtual appliance, a pipeline can be replicated and distributed with ease, addressing two critical factors of clinical bioinformatics.

POSTER 53

NORMALISATION AND DIFFERENTIAL ENRICHMENT ANALYSIS OF CHIP-SEQ DATA

Christoffer Flensburg, Alicia Oshlack

ABSTRACT

Large amounts of ChiP-seq data are being generated for studying histone modifications. Although new tools are constantly being developed many challenges still remain. In particular, epigenetic ChIP-seq data is often very noisy with a relatively large component of "background" DNA sequenced along with the true enriched signal. Sequencing input DNA is often used to correct for the background contamination, but an incorrectly used input, or an input of poor quality, can introduce as much new noise as it removes. We will discuss ways to assess the quality of an input, and how the input can be best used to improve the results. We also show methods for estimating the background levels in the absence of a good quality input. From this we are able to separate the enrichment signal from the background and show the importance of normalisation in this context; a library size normalisation is convenient, but can lead to incorrect conclusions. Using our signal estimates we show that we can perform differential enrichment analysis without the contamination of the background signal leading to higher quality performance.

POSTER 54

THE VALUE OF CONTROLS IN PEAK CALLING FROM CHIP-SEQ EXPERIMENTS

Fabian Buske, Phillippa Taberlay, Susan Clark

ABSTRACT

ChIP-seq is the method of choice for interrogating the DNA occupancy of proteins involved in gene regulation. Reliable assessment of ChIP enrichment (peak calling) requires sequencing of a matched control library (e.g. input DNA) to compensate for biases (copy-number alterations, sequence content, chromatin structure, antibody quality). However, for economic reasons matched control libraries are often sequenced at lower depth than the ChIP enriched sample or are not sequenced at all. It is therefore important to address if input sequenced controls are required for the accurate interpretation of ChIP-seq data and if so should input data be from matched control libraries or will unmatched input libraries suffice.

We investigated the effect of input sequencing controls on peak calling by contrasting matched controls with libraries generated from unmatched biological replicates or obtained from ENCODE project using the same cell lines. We considered the peaks generated from matched controls as the gold standard and assessed the accuracy of unmatched controls from the same cell lines to call the equivalent enriched regions with at least 50% overlap. We observe for all three interrogated histone marks (H3K9K14ac, H3K4me1, H3K27ac) that high accuracy can be achieved with unmatched input controls depending on the sequencing depth of the control library (Peakranger accuracy of 0.99 vs 0.97 vs 0.79 using 21, 13 or 5.5 mil. mapped reads, respectively).

Furthermore, investigating the base pair overlap of the enriched regions, we observe that the algorithm of choice has a greater impact than utilizing an unmatched control (average Jaccard similarity coefficient of 0.24 between Peakranger, Homer and Chromablocks using matched controls and 0.76 between matched versus unmatched control libraries using the same algorithm).

We therefore conclude that it is reasonable to use an unmatched control even from public data if there is high sequencing coverage. This has important ramifications in the processing of ChIP data using different antibodies to interrogate the same cell type.

POSTER 55

THE TRANSCRIPT AND THE NOISE: COULD PERSONALIZED TRANSCRIPTOMICS BEAT THE DRUG CHEATS?

Sarah Hausner, Aaron Chuah, Christopher Gore, Hardip Patel, Simon Easteal

ABSTRACT

Until the present day, systematic illicit performance enhancement and the lack of adequate detection tools remain organised sport's greatest challenge. A seemingly endless influx of novel substances and techniques make traditional testing methods somewhat ineffective and pose a tremendous challenge to governmental agencies, sporting authorities and testing facilities alike. In order to facilitate sufficient disclosure of illicit performance enhancement a holistic personalized strategy is required. The World Anti-Doping Agency's ADAMS and Athlete Biological Passport systems enable orchestration of location data and routine test results. Longitudinal information, however, is limited to standard blood parameters, which have been subject to manipulation in the past and indicate non-physiological changes within cells just to a limited extent.

We suggest global gene expression in peripheral blood as a diagnostic test with greater sensitivity that could indicate illicit performance enhancement without a priori assumptions and might become a powerful extension of the existing monitoring system. In order to provide a general overview of peripheral blood gene expression in highly trained subjects which subsequent research can build on, we examined global gene expression in 49 Australian elite athletes upon normal variation and preserved transcriptome properties. We used a generalized linear model to examine samples upon systematic differences associated with gender, training mode, fasting, circadian rhythm or technical bias. We further determined network properties of gene co-expression through weighted correlation network analysis in order to determine whether, at this stage, global gene expression has the potential to become the key player of an effective, holistic detection strategy.

EPIGENETICS

POSTER 56

DE NOVO IDENTIFICATION OF DIFFERENTIALLY METHYLATED REGIONS IN THE HUMAN GENOME

Tim Peters, Mike Buckley, Susan Clark, Peter Molloy

ABSTRACT

The Illumina® HM450K array interrogates the human methylome by measuring methylation signals at approximately half a million CpG sites of biological interest. However, identifying the most differentially methylated (DM) probes alone, even with annotation, is of fairly limited use. What is more useful is identifying regions of DM: clusters of CpG sites whose DM signals correspond with loci of particular biological functionality. A principled agglomeration of DM CpG sites, informed by consecutivity, annotation, and relative genomic position, along with a robust measure of differential methylation itself, is needed to properly extract these regions. Methods such as bump hunting (Jaffe et al. 2012) and QDMR (Zhang et al. 2011) attempt to do this, but suffer from arbitrary breakpoints and operational issues. We present a less parameterised method that fits CpGs of interest to a weighted probability density function with kernel estimation, which is able to rank the most differentially methylated regions, across the whole genome, based on the density of the DM signal at any given point. This method mitigates the bias in the design of the array probeset, and provides a significance test for each region via comparison to a null. It is also able to detect regions of high magnitude and variability of methylation in unlabelled data, and has scope for integration into existing visualisation tools and statistical analysis software packages.

1. Preventative Health Flagship, CSIRO, North Ryde, NSW, Australia, 2113

2. The Garvan Institute of Medical Research, Darlinghurst, NSW, Australia 2010

3. Computational Informatics, CSIRO, North Ryde, NSW, Australia, 2113

POSTER 57

GLOBAL DNA METHYLATION PROFILING REVEALS SIGNIFICANT DIFFERENTIAL METHYLATION OF SUBCUTANEOUS AND OMENTAL ADIPOSE TISSUE BEFORE AND AFTER WEIGHT LOSS.

Donia Macartney-Coxson, Miles Benton, Alice Johnstone, David Eccles, Brennan Harmon, Mark Hayes, Rod Lea, Lyn Griffiths, Eric Hoffman, Richard Stubbs

ABSTRACT

The role of epigenetics in obesity is increasingly recognised. Adipose tissue plays a key role in the metabolic dysfunction seen in obesity, and gastric bypass provides a model in which to investigate obesity and weight-loss in humans. Therefore we investigated one epigenetic mechanism, DNA methylation, in adipose tissue from 15 females before and after gastric bypass and significant weight loss (>27% initial body weight). Genome wide DNA methylation profiling of 485,577 individual cytosines in matched (before and after weight-loss) subcutaneous and omental adipose tissue samples was performed using the Illumina 450K platform. Using a paired analysis we observed significant differential methylation (Bonferonni correction p<1x10-7) at 15 and 3601 CpG sites in omental and subcutaneous abdominal tissue respectively. A greater proportion of CpG sites were relatively hypermethylated before weight-loss.

Robustly identified CpG sites mapped within obesity associated genes, and homeobox gene clusters. We also observed differential methylation at ≥ 6 CpG sites within genes with known or putative roles in obesity and related traits: *PRKCZ*, *PRDM16*, *FOXP2*, *THBS1* and the type two diabetes associated gene *KCNQ1*. In addition, suggestive correlations between differential methylation and change in clinical trait before and after weight-loss were observed.

To our knowledge this is the first study to report global DNA methylation profiling of adipose tissue before and after the major weight-loss observed with gastric bypass. It provides a strong basis for future work and offers additional and compelling evidence for the role of epigenetics, and adipose tissue DNA methylation in particular, in obesity.

The project was supported by ESR core funding, a pilot grant from the National Center for Medical Rehabilitation Research At Children's National Medical Centre (5R24HD050846-08): NCMRR-DC Core Molecular and Functional Outcome Measures in Rehabilitation Medicine and The Wellington Medical Research Foundation Inc. Miles Benton is supported by a Corbett Postgraduate Research grant.

POSTER 58

CORRECTING FOR UNWANTED VARIATION IN 450K METHYLATION ARRAY STUDIES

Jovana Maksimovic, Johann Gagnon-Bartsch, Terry Speed, Alicia Oshlack

ABSTRACT

DNA methylation is the most widely studied epigenetic mark; it is known to be essential to normal development and is frequently disrupted in disease. The Illumina Infinium HumanMethylation450 (450k) BeadChips are being taken up at a rapid rate due to their affordability, genomic coverage and technical reproducibility. As with microarray expression experiments, 450k array studies are subject to unwanted technical variation such as batch effects and other, often unknown, sources of variation. We are currently investigating the utility of a suite of new methods for removing unwanted variation (RUV) in 450k differential methylation studies. These methods extend the RUV2 approach originally developed by Gagnon-Bartsch & Speed (2012) for microarray differential expression studies. The RUV methods use factor analysis of designated negative control probes to identify components of unwanted variation that are subsequently adjusted for using the linear modelling framework. We have assessed the performance of the RUV methods with various types of negative controls, using several 450k datasets. We show that the RUV methods consistently perform better than competing approaches. We also demonstrate that the RUV methods are able to correct for unwanted variation in cases where there is significant confounding with the factor of interest, which is often difficult to achieve using standard approaches.

Gagnon-Bartsch JA, Speed TP, *Using control genes to correct for unwanted variation in microarray data,* Biostatistics. 2012 Jul;13(3):539-52.

RNA/TRANSCRIPTOMICS

POSTER 59

ORCHESTRATED INTRON RETENTION REGULATES NORMAL GRANULOCYTE DIFFERENTIATION

William Ritchie, Justin Wong, John Rasko, Jeff Holst

ABSTRACT

Intron retention (IR) is widely recognized as a consequence of mis-splicing that leads to failed excision of intronic sequences from pre-messenger RNAs. Our bioinformatic analyses of transcriptomic and proteomic data of normal white blood cell differentiation reveal IR as a physiological mechanism. IR actively regulates the expression of nearly one hundred functionally related genes, including those that determine the nuclear shape that is unique to granulocytes. Retention of introns in specific genes is associated with downregulation of splicing factors and is conserved between human and mouse. IR led to reduced mRNA and protein levels by triggering the nonsense-mediated decay (NMD) pathway. In contrast to the prevalent view that NMD is limited to mRNAs encoding aberrant proteins, our data establish that IR coupled with NMD is a conserved mechanism in normal granulopoiesis. Physiological IR may provide an energetically favorable level of dynamic gene expression control prior to sustained gene translation. Our findings have been published on 1st August 2013 in Cell.

POSTER 60

DIFFERENTIAL EXPRESSION ANALYSIS OF MULTIFACTOR RNA-SEQ EXPERIMENTS

Yunshun Chen, Gordon Smyth

ABSTRACT

A flexible statistical framework is developed for the analysis of read counts from RNA-Seq gene expression studies. It provides the ability to analyse complex experiments involving multiple treatment conditions and blocking variables while still taking full account of biological variation. Biological variation between RNA samples is estimated separately from the technical variation associated with sequencing technologies. Novel empirical Bayes methods allow each gene to have its own specific variability, even when there are relatively few biological replicates from which to estimate such variability.

Parallel computational approaches are developed to make nonlinear model fitting faster and more reliable, making the application of GLMs to genomic data more convenient and practical. Simulations demonstrate the ability of adjusted profile likelihood estimators to return accurate estimators of biological variability in complex situations. When variation is gene-specific, empirical Bayes estimators provide an advantageous compromise between the extremes of assuming common dispersion or separate genewise dispersion.

The pipeline is implemented in the edgeR package of the Bioconductor project. The methods developed here can also be applied to count data arising from DNA-Seq applications, including ChIP-Seq for epigenetic marks and DNA methylation analyses.

POSTER 61

GENE-LEVEL DIFFERENTIAL EXPRESSION ANALYSIS FOR DE NOVO ASSEMBLED TRANSCRIPTOMES

Nadia Davidson, Alicia Oshlack

ABSTRACT

The advent of next-generation sequencing has made it possible to analyse the transcriptomes of non-model organisms via a de novo assembly of RNA-seq reads. In particular, this approach allows for differential expression analysis without the need for a reference genome or annotation. Although a number of studies have compared the relative merits of different transcriptome assembly programs, less attention has been focused on the procedure for performing a differential expression analysis, following transcriptome assembly.

We show that performing a differential expression analysis at the level of genes rather than transcripts improves the interpretability of results and increases statistical power. However, the procedure for going from a *de novo* assembled transcriptome to differential gene expression results is not obvious. Moreover, existing methods to cluster transcripts into genes are not optimized for *de novo* assembled transcript sequences. To address these issues, we have developed Corset, a method that uses hierarchical clustering to group *de novo* assembled transcripts

based on shared reads and expression. Corset also summarises read counts to gene clusters, offering a convenient single-step solution to go from assembled transcripts to gene-level count data, ready for statistical testing. We show that our method offers superior differential gene expression results in all cases we have tested. Corset is open source software and is available from **https://code.google.com/p/corset-project**.

POSTER 62

THE EMERGING PARADIGM OF REGULATED INTRON RETENTION IN MRNA: MAPPING THE PROGRAM OF RNA-BINDING PROTEINS WITH RNA-SEQ

Dan Andrews, Vicky Cho, Arleen Sanny, Andy Tan, Chris Goodnow

ABSTRACT

Retention of a subset of introns in spliced polyadenylated RNA is emerging as a frequent, unexplained finding from deep sequencing of messenger RNA in mammalian cells. Until recently the biological significance of intron retention has remained unclear, though recent studies have indicated that differential intron retention is widespread and strongly associated with alternatively spliced exons. We have identified that, at least in one context, differential intron retention is a regulated process that can be controlled by RNA-binding proteins.

We have found that the developmentally regulated RNA-binding protein, hnRNPLL, induces retention of specific introns by comparing mRNA deep sequencing datasets from mutant T cells with an inactivating Hnrpll mutation. This is physiologically important in wild-type cells – a closely similar pattern of differential intron can be observed between T and B cells, where B cells naturally express low levels of hnRNPLL, and regulates alternative splicing of CD45 isoforms, which are a key developmental marker in T cells. We found that detection of genes with differentially retained introns by integration of several computational methods provides a general strategy to identify alternative spliced exons in mammalian RNA-seq data.

As the importance of the role of RNA-binding proteins becomes better appreciated, it is increasingly of interest to map the binding targets of this diverse family of regulators. We demonstrate a novel mapping approach using RNA-seq that may be paired with growing libraries of mutant hnRNPs, such as the Missense Mutation Library being generated at the Australian Phenomics Facility (**http://databases.apf.edu.au/mutations/**).

POSTER 63

COMPARISON OF IN VITRO AND IN VIVO GENE EXPRESSION IN BOVINE OVARIAN FOLLICLES BY MICROARRAY ANALYSES

Nick Hatzirodos, Katja Hummitzsch, Helen Irving-Rodgers, Claire Glister, Philip Knight, Ray Rodgers

ABSTRACT

A mature ovarian follicle consists of an oocyte within a fluid filled antrum surrounded by estrogen secreting granulosa cells, and a thecal cell layer which supplies precursors for estrogen synthesis. Only relatively few follicles will eventually fully develop and release a mature oocyte or become atretic, and this process is under hormonal control by FSH and LH and other local and systemic factors which are less well understood. We first examined gene transcription in the granulosa cells and theca interna of small healthy (n = 10 each), small atretic (n = 5 each), and large healthy follicles. This data was compared with that of granulosa cells cultured with or without FSH (n=7) and TNF-′ (n=8), and thecal cells with or without LH (n=8) and BMP-6 (n=8) to mimic the effects of growth and atresia using Bovine Affymetrix Genome Arrays by ANOVA with Partek Genomics Suite software (v6.5). The DEGs from different comparisons between analogous in vitro and in vivo cell environments were uploaded to Ingenuity Pathway Analysis (IPA) for pathway and functional analysis. The main effect of culture across all granulosa and thecal cell arrays was determined to be the up-regulation of proinflammatory pathways, perhaps due to the length and type of culture. Genes of interest which were up-regulated in vivo in small follicle granulosa cells apart from steroid synthesis, include MEST and AMH, and in large follicle granulosa cells CSPG4, IGFBP4 and PLA2G1B.

Thank You to Our Generous Sponsors

The organising committee for the 2013 AMATA conference would like to take this opportunity to thank the following organisations for their valuable contribution, and ask all delegates to please consider these organisations for future business relationships.

Platinum Sponsor

QIAGEN is a global leader in sample and assay technologies for life sciences, applied testing, and molecular diagnostics. Its products are considered standards that comprise complete solutions from sample to result. QIAGEN offers the broadest portfolio of molecular diagnostic assays for infectious diseases, including the only test for human papillomavirus (HPV), which has both FDA and CE-mark approval. The company has developed a comprehensive portfolio of more than 500 proprietary, consumable products and instruments for sample collection, nucleic acid and protein handling, separation, purification, detection, and open and target-specific assays.

QIAGEN is also establishing a leading role in the next-generation sequencing (NGS) application with complete sample-to-result solutions for NGS from combining innovative instrumentation and consumables for clinical research and healthcare applications. The recent acquisition of Ingenuity Systems will further add a leading solution for analysis and interpretation of complex biological data."

www.QIAGEN.com

Silver Sponsor

Illumina is advancing human health by unlocking the power of the genome; from genome-wide discovery through to targeted validation and beyond.

At Illumina, our goal is to apply innovative technologies and revolutionary assays to the analysis of genetic variation and function, making studies possible that were not even imaginable just a few years ago.

These revolutionary tools for DNA, RNA and protein analysis are enabling rapid advances in disease research, drug development and the development of molecular tests in the clinic.

www.illumina.com

Silver Sponsor

Millennium Science is a specialist distribution company established in 1999 specifically to source, supply and support high technology solutions to the Australasian Life Science market. The company is well established as a leading distributor of instrumentation and reagents for the research market with recognition of performance and business practice excellence by the global companies represented, including Fluidigm.

Fluidigm develops and manufactures Integrated Fluidic Circuit (IFC) technology that significantly improves productivity in life science research. Fluidigm IFCs enable the simultaneous performance of thousands of sophisticated biochemical measurements in nanolitre volumes. These systems are used for applications ranging from single-cell genomics, targeted resequencing, gene expression and digital PCR.

www.mscience.com.au
www.fluidigm.com

Silver Sponsor

Roche Applied Science offers cutting-edge research solutions, from manual or automated nucleic acid extraction and qPCR instrumentation to bench top sequencing on our new GS Junior System, utilising proven 454 technology for long reads. The GS FLX + instrument can generate read lengths up to 700bp or more, which facilitates whole genome assembly as well as accurate species identification in metagenomics studies. NimbleGen SeqCap EZ catalogue and customsolutions completes our range of quality genomics tools.

www.roche-applied-science.com

Keynote Speaker Sponsor

Stem Cells Australia brings together experts in bioengineering, nanotechnology, stem cell biology, bioinformatics and clinical research to tackle the big questions in stem cell science. Through understanding more about mechanisms involved in stem cell regulation and differentiation, this Australian Research Council-funded Special Research Initiative aims to develop innovative biotech and therapeutic applications. Our network includes over 120 scientists from University of Melbourne, University of Queensland, Monash University, University of NSW, Walter and Eliza Hall Institute for Medical Research, Victor Chang Cardiac Research Institute, The Florey Institute of Neuroscience and Mental Health and CSIRO.

stemcellsaustralia.edu.au

Student Poster Prize Sponsor

Australian Genome Research Facility Ltd (AGRF), a not for profit company, is Australia's largest provider of genomics services and solutions. AGRF has laboratories in Brisbane, Sydney, Melbourne, Adelaide 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. Our range of services is unique in Australia and recognised internationally. AGRF is accredited by NATA to the ISO/IEC17025:2005 Quality Standard.

www.agrf.org.au

Conference Supporter

The Australian Bioinformatics Network aims to connect people to

- people • resources • opportunities

to increase the benefits Australian bioinformatics can deliver. While our focus is on Australia, we want to help science make an even more positive difference to the world. We hope the Australian Bioinformatics Network will help foster science links internationally and we welcome participation from anyone interested in bioinformatics. The Australian Bioinformatics Network is free and open to all, in Australia or the rest of the world. It is funded by CSIRO, EMBL Australia and Bioplatforms Australia.

AustralianBioinformatics.net

Conference Supporter

www.goldcoastconventions.com

Thank You to Our Trade Exhibitors

Trade Booth 1

Roche Applied Science offers cutting-edge research solutions, from manual or automated nucleic acid extraction and qPCR instrumentation to bench top sequencing on our new GS Junior System, utilising proven 454 technology for long reads. The GS FLX + instrument can generate read lengths up to 700bp or more, which facilitates whole genome assembly as well as accurate species identification in metagenomics studies. NimbleGen SeqCap EZ catalogue and custom solutions completes our range of quality genomics tools.

www.roche-applied-science.com

Trade Booth 2 & 3

QIAGEN is a global leader in sample and assay technologies for life sciences, applied testing and molecular diagnostics. Its products are considered standards that comprise complete solutions from sample to result. QIAGEN offers the broadest portfolio of molecular diagnostic assays for infectious diseases including the only test for human papillomavirus (HPV), which has both FDA and CE-approvals. The company has developed a comprehensive portfolio of more than 500 proprietary, consumable products and instruments for sample collection, nucleic acid and protein handling, separation, purification, detection, and open and target specific assays.

www.QIAGEN.com

Trade Booth 4

Genesearch is a privately held Australian biotechnology company founded in 1980 and home of the e•Freezer. Talk to us about products from these suppliers:

New England Biolabs; NEBNext® reagents; a series of highly pure reagents that facilitate sample preparation of DNA or RNA for downstream applications such as next generation sequencing and expression library construction

Single Cell Whole Genome Amplification kit; amplify total DNA from single cells or their DNA equivalent about 1 million-fold to produce 2–5 micrograms of amplified DNA in under 3 hours!

Genecodes Sequencher; DNA sequence assembly and analysis tools for next-generation and traditional (Sanger or Capillary Electrophoresis) DNA data sets.

For more information visit:

www.genesearch.com.au

Trade Booth 5

OnQ Software is dedicated in providing leading edge custom technology solutions with the aim of enabling end to end integration across your entire workplace using our suite of state of the art software, tools and highly trained experts who can tailor a solution for your business needs.

Our software solutions range from different laboratory LIMS Systems to suit general labs to next generation sequencing labs, Freezer and Biobanking, electronic notebooks, digital pathology and hospital healthcare solutions. OnQ Software has been in the industry for over 20 years enabling partners and customers to achieve their goals through delivering quality solutions and services. Our consultancy services include:

- Gap analysis
- Functional and non-functional specification
- Business impact analysis

www.onqsoft.com.au

Trade Booth 6

TrendBio
Partnerships in Science

TrendBio is a supplier of specialised instrumentation for the Life Sciences, with specialties in Genomics, Proteomics, Pre-Clinical Animal Imaging and Biologics Production.

- DNA and Chromatin Shearing
- FFPE DNA Extraction
- PCR
- Quantitative Gene Expression Analysis
- NGS Dx Kits
- Biomolecular Interaction Analysis
- Small Animal Imaging (Optical, PET/SPECT, CT, X-Ray, Optoacoustic)
- Protein Analysis (Western Blotting)

Phone: 1300 720 574

Email: **info@trendbio.com.au**

www.trendbio.com.au

Trade Booth 7

agrf
Australian Genome Research Facility Ltd

Australian Genome Research Facility Ltd (AGRF), a not for profit company, is Australia's largest provider of genomics services and solutions. AGRF has laboratories in Brisbane, Sydney, Melbourne, Adelaide 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.

Our range of services is unique in Australia and recognised internationally. AGRF is accredited by NATA to the ISO/IEC17025:2005 Quality Standard.

www.agrf.org.au

Trade Booth 8

SIGMA-ALDRICH®

Sigma-Aldrich is a leading Life Science and High Technology Company that supplies over 220,000 products for Biology, Analytical Science and Chemistry. Our comprehensive Life Science portfolios include kits and reagents for NGS, RNA-based TRC libraries and custom deconvolution services. For further information, visit our booth or:

www.sigma-aldrich.com

Trade Booth 9

SARSTEDT

Sarstedt Australia is part of the Sarstedt group of companies. Founded in 1961 with international headquarters located in Germany; the group employs almost 2,500 people and is a leader manufacturer of diagnostic, laboratory and medical products. Our range includes both equipment and consumables for use in pathology, laboratory, research, transfusion medicine, blood banking, hospital, pharmaceutical, forensic and industrial applications.

The Sarstedt Group is committed to research, development and quality. This together with innovation and diversification gives the organisation an outstanding reputation world-wide.

www.sarstedt.com

Trade Booth 10

QFAB

Established in 2007, QFAB focuses on delivering robust, high quality and relevant bioinformatics services for life science researchers to analyse and manage large-scale datasets.

Our support ranges from experimental design, data capture and mining through to NGS, proteomic and metabolomic analyses. We are also expert in cross-domain integration with clinical data.

QFAB can work on specific projects, provide time against an ongoing need or become embedded in your organisation. Flexible, rapid and researcher oriented, QFAB has built a reputation for supporting the production of high quality outcomes which deliver those high impact publications or patents faster.

www.qfab.org

Trade Booth 11

bio-strategy
delivering technology

Bio-Strategy is a team of highly qualified people, throughout Australia and New Zealand, who share a passion for Science.

Representing world leading suppliers, Bio-Strategy brings our customers a complete range of solutions, from basic laboratory instrumentation through to advanced technology, covering applications in the Research, Healthcare, Biotechnology and Industrial Markets.

Our focus and commitment to our customers and suppliers, is key to Bio-Strategy's excellence and success.

www.bio-strategy.com

Trade Booth 12

life technologies™

We are a global life sciences company, shaping discovery and improving life.
Like you, we are grounded in science and innovating for the future. That's why we offer the most trusted portfolio of high-quality, innovative solutions to support life-changing endeavours around the world.

We believe in the power of science and appreciate its rigorous discipline. That's what drives our passion for innovation, leading to transformative offerings that support endeavors throughout the world.

Our extensive range of products and services, from instruments to everyday lab essentials, ensures quality and performance for every lab, every application. Customers in more than 180 countries count on us in their quest to improve life in meaningful ways.

For more information visit **www.lifetechnologies.com** or, contact us on 1800 636 327

www.lifetechnologies.com

Trade Booth 13

Partek®
turning data into discovery®

Partek Incorporated is a leading provider of integrated genomics data analysis and visualization software. Trusted by scientists worldwide, its software suite—Partek® Flow™, Partek® Genomics Suite™, and Partek® Pathway™—uniquely supports all microarray and next-generation sequencing technologies in a single software package. Partek, turning data into discovery®.

www.partek.com

Trade Booth 15

Macrogen has over 16 years of experience in sequencing and analysis and is a provider of a complete pipeline of Next-Generation Sequencing (NGS) services – including library construction, sequencing, and bioinformatic analysis. Our quality-certified (ISO 9001:2000) facility is one of the top 5 largest sequencing facilities in the world, and it assures the client samples are prepared, sequenced, and analysed according to the highest standard operating procedures. We proudly serve the research community with the industry-leading data quality and services, and we provide full personal consultation throughout the sequencing process. Contact us today and speak to our expert now.

www.macrogen.co

Trade Booth 17

Integrated Sciences, established in 1983, supplies the Australian and New Zealand Diagnostics and Life Science markets. We pride ourselves in offering quality products with the highest levels of technical support and service. To bring you the best, we partner with premium suppliers, including Agilent Technologies who are a global leader in target enrichment for next-generation sequencing and genomic microarrays.

Products arising from Agilent's expertise in synthesizing custom complex mixtures of long oligonucleotides include SureSelect and Haloplex target enrichment systems, microarrays for genome-wide measurement of gene expression and cGH, as well as SureFISH, a highly specific and sensitive product line for oligonucleotide FISH.

www.integratedsci.com.au

Trade Booth 18

Illumina is advancing human health by unlocking the power of the genome; from genome-wide discovery through to targeted validation and beyond.

At Illumina, our goal is to apply innovative technologies and revolutionary assays to the analysis of genetic variation and function, making studies possible that were not even imaginable just a few years ago.

These revolutionary tools for DNA, RNA and protein analysis are enabling rapid advances in disease research, drug development and the development of molecular tests in the clinic.

www.illumina.com

Trade Booth 19 & 20

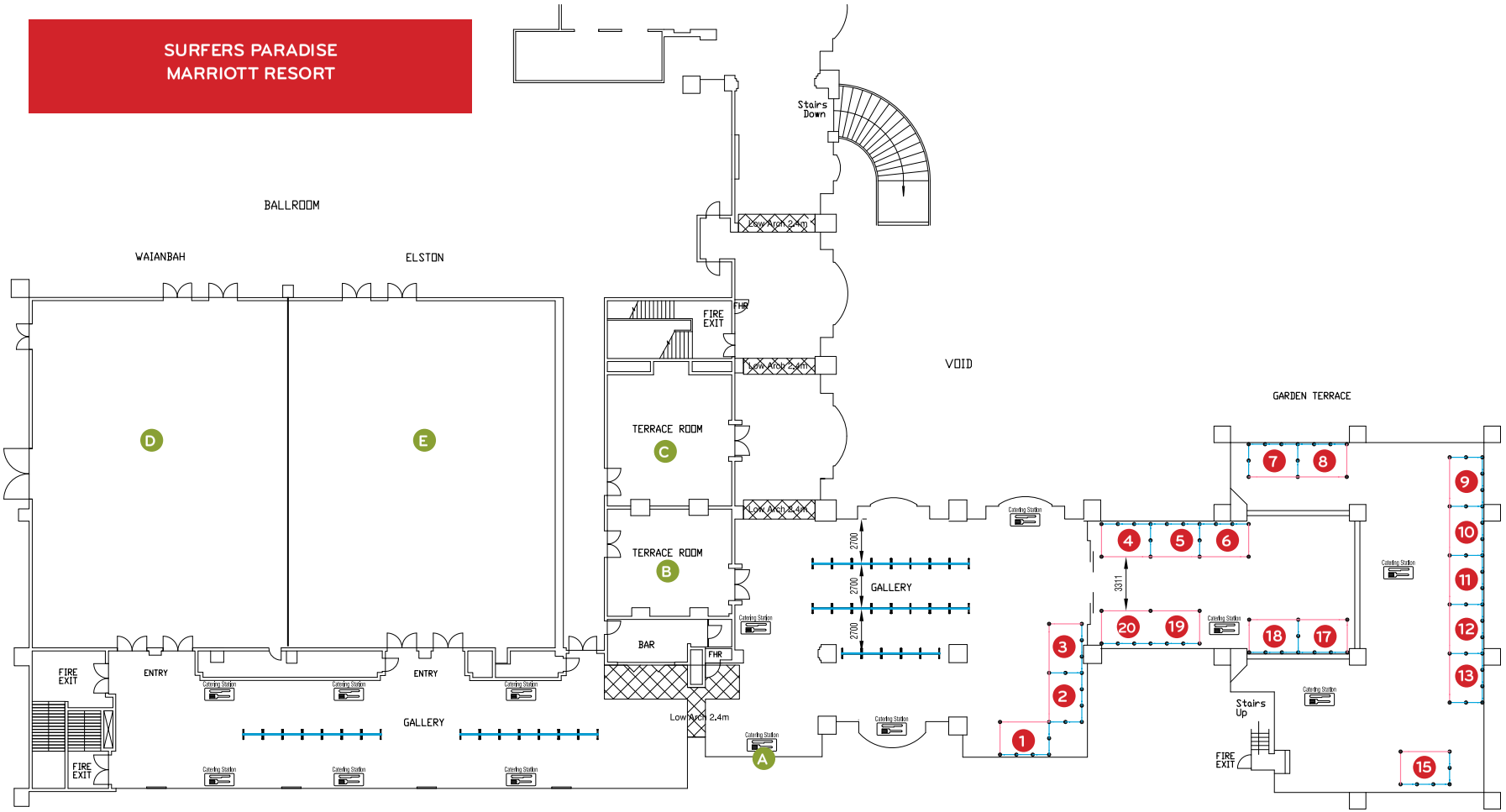
Millennium Science is a specialist distribution company established in 1999 specifically to source, supply and support high technology solutions to the Australasian Life Science market. The company is well established as a leading distributor of instrumentation and reagents for the research market with recognition of performance and business practice excellence by the global companies represented, including Fluidigm.

Fluidigm develops and manufactures Integrated Fluidic Circuit (IFC) technology that significantly improves productivity in life science research. Fluidigm IFCs enable the simultaneous performance of thousands of sophisticated biochemical measurements in nanolitre volumes. These systems are used for applications ranging from single-cell genomics, targeted resequencing, gene expression and digital PCR.

www.mscience.com.au

www.fluidigm.com

SURFERS PARADISE MARRIOTT RESORT



AMATA EXHIBITORS

1. Roche Diagnostics
2. 3. QIAGEN
4. Genesearch
5. OnQ Software
6. Trend Bio
7. Australian Genome Research Facility
8. Sigma-Aldrich
9. Sarstedt Australia
10. QFAB
11. Bio-Strategy
12. Life Technologies
13. Partek
15. Macrogren
17. Integrated Sciences
18. Illumina
19. 20. Millennium Science

GENERAL INFORMATION

- A AMATA Registration
- B AMATA Speakers Prep
- C AMATA Parents Room
- D AMATA Dinner
- E AMATA Plenary

Delegate List

Last Name	First Name	Title	Organisation	State	Country
Alexandrov	Kirill	Prof	The University of Queensland	QLD	AUSTRALIA
Andrews	Dan	Dr	Australian Phenomics Facility	ACT	AUSTRALIA
Andrew	Rose	Dr	University of British Columbia		CANADA
Anko	Minni	Dr	Walter & Eliza Hall Institute of Medical Research	VIC	AUSTRALIA
Bahlo	Melanie	Dr	Walter & Eliza Hall Institute of Medical Research	VIC	AUSTRALIA
Bailey	Timothy	Dr	The University of Queensland	QLD	AUSTRALIA
Baines	Katie	Dr	The University of Newcastle	NSW	AUSTRALIA
Barash	Mark	Mr	Bond University	QLD	AUSTRALIA
Bassam	Brant	Dr	Bio-Strategy		NEW ZEALAND
Batra	Jyotsna	Dr	Queensland University of Technology	QLD	AUSTRALIA
Bauer	Denis	Dr	CSIRO	NSW	AUSTRALIA
Bedo	Justin	Dr	NICTA	VIC	AUSTRALIA
Beesley	Jonathan	Dr	Queensland Institute of Medical Research	QLD	AUSTRALIA
Bellos	Evangelos	Mr	The University of Queensland	QLD	AUSTRALIA
Bragg	Lauren	Dr	CSIRO	QLD	AUSTRALIA
Budzinska	Magdalena	Ms	Centenary Institute	NSW	AUSTRALIA
Burdon	Kathryn	Dr	Flinders University	SA	AUSTRALIA
Buske	Fabian	Dr	Garvan Institute of Medical Research	NSW	AUSTRALIA
Butcher	Suzanne	Miss	Australian Institute For Bioengineering & Nanotechnology	QLD	AUSTRALIA
Cameron	Graham	Mr	EBI Australia	QLD	AUSTRALIA
Catchpoole	Dan	A/Prof	The Kids Research Institute	NSW	AUSTRALIA
Cazaly	Emma	Miss	Menzies Research Institute	TAS	AUSTRALIA
Chandrananda	Dineika	Ms	Walter & Eliza Hall Institute of Medical Research	VIC	AUSTRALIA
Charlesworth	Jac	Dr	University of Tasmania	TAS	AUSTRALIA
Chen	Yunshun	Mr	Walter & Eliza Hall Institute of Medical Research	VIC	AUSTRALIA
Cloonan	Nicole	Dr	Queensland Institute of Medical Research	QLD	AUSTRALIA
Coin	Lachlan	Dr	The University of Queensland	QLD	AUSTRALIA
Colgrave	Michelle	Dr	CSIRO	QLD	AUSTRALIA
Corbin	Vincent	Dr	Walter & Eliza Hall Institute of Medical Research	VIC	AUSTRALIA
Corley	Susan	Dr	The University of New South Wales	NSW	AUSTRALIA
Cowley	Mark	Dr	Garvan Institute of Medical Research	NSW	AUSTRALIA
Crowe	Mark	Dr	Queensland Facility for Advanced Bioinformatics	QLD	AUSTRALIA
Crowley	Tamsyn	Dr	Deakin University	VIC	AUSTRALIA
Davidson	Nadia	Dr	Murdoch Childrens Research Institute	VIC	AUSTRALIA

Last Name	First Name	Title	Organisation	State	Country
Dawes	Ian	Emeritus Prof	The University of New South Wales	NSW	AUSTRALIA
Desselle	Mathilde	Ms	Queensland Facility for Advanced Bioinformatics	QLD	AUSTRALIA
Devgan	Vikram	Dr	QIAGEN		USA
Dibbens	Leanne	A/Prof	University of South Australia	SA	AUSTRALIA
Diesch	Jeannine	Dr	Peter MacCallum Cancer Centre	VIC	AUSTRALIA
Dinger	Marcel	A/Prof	Garvan Institute of Medical Research	NSW	AUSTRALIA
Dollard	Ruth	Ms	Sarstedt Australia	SA	AUSTRALIA
Dumsday	Jeremy	Mr	Agilent Technologies	VIC	AUSTRALIA
Eades	Tamsin	Dr	Illumina	VIC	AUSTRALIA
Edwards	Stacey	Dr	QIMR	QLD	AUSTRALIA
Engelke	Ulrich	Dr	CSIRO	TAS	AUSTRALIA
Everest	Marie	MS	Australian Genome Research Facility	VIC	AUSTRALIA
Faiz	Fathimath	Mrs	University of Western Australia	WA	AUSTRALIA
Fan	Jian Bing	Dr	Illumina		USA
Filipovska	Aleksandra	Prof	Walter & Eliza Hall Institute of Medical Research	WA	AUSTRALIA
Fink	Lynn	Dr	Queensland Centre For Medical Genomics	QLD	AUSTRALIA
Flensburg	Christoffer	Dr	Murdoch Childrens Research Institute	VIC	AUSTRALIA
Flynn	Aidan	Mr	Peter MacCallum Cancer Centre	VIC	AUSTRALIA
Gabrielli	Brian	Dr	The University of Queensland	QLD	AUSTRALIA
Gandhi	Kaushal	Dr	Integrated Science	NSW	AUSTRALIA
Gardiner	Brooke	Dr	The University of Queensland	QLD	AUSTRALIA
Garred	Michelle	Ms	Illumina	VIC	AUSTRALIA
Ghazanfar	Shila	Miss	The University of Sydney	NSW	AUSTRALIA
Gloss	Brian	Dr	Garvan Institute of Medical Research	NSW	AUSTRALIA
Goldinger	Anita	Ms	The University of Queensland	QLD	AUSTRALIA
Gorse	Dominique	Mr	Queensland Facility for Advanced Bioinformatics	QLD	AUSTRALIA
Gould	Kate	Ms	Peter Maccallum Cancer Centre	VIC	AUSTRALIA
Greenfield	Paul	Mr	CSIRO	NSW	AUSTRALIA
Hancock	Timothy	Dr	NICTA	VIC	AUSTRALIA
Hatzirodos	Nick	Mr	The University of Adelaide	SA	AUSTRALIA
Hausner	Sarah	Ms	Australian National University	ACT	AUSTRALIA
Hayward	Nick	Dr	QIMR	QLD	AUSTRALIA
Heazlewood	Celena	Dr	Australian Institute For Bioengineering & Nanotechnology	QLD	AUSTRALIA
Ho	Joshua	Dr	Victor Chang Cardiac Research Institute	NSW	AUSTRALIA
Huang	Edward	Mr	The University of Queensland	QLD	AUSTRALIA
Humphreys	David	Dr	Victor Chang Cardiac Research Institute	NSW	AUSTRALIA

Last Name	First Name	Title	Organisation	State	Country
Jeddeloh	Jeff	Dr	Roche NimbleGen		USA
Jones	Nicole	Dr	QIAGEN	VIC	AUSTRALIA
Kainer	David	Mr	Australian National University	ACT	AUSTRALIA
Kasinadhuni	Naga	Mr	Australian Genome Research Facility	VIC	AUSTRALIA
Kassahn	Karin	Dr	SA Pathology	SA	AUSTRALIA
Keown	Paul	Prof	University of British Columbia & Vancouver General Hospital		CANADA
Kliese	Rachel	Mrs	Australian Genome Research Facility	VIC	AUSTRALIA
Kolle	Gabriel	Dr	Illumina	VIC	AUSTRALIA
Korn	Othmar	Mr	Australian Institute For Bioengineering & Nanotechnology	QLD	AUSTRALIA
Kulheim	Carsten	Dr	Australian National University	ACT	AUSTRALIA
Lacaze	Paul	Mr	Millenium Science	VIC	AUSTRALIA
Lai	John	Dr	Queensland University of Technology	QLD	AUSTRALIA
Larney	Christian	Mr	The University of Queensland	QLD	AUSTRALIA
Laugraud	Aurelie	Miss	AgResearch Limited		NEW ZEALAND
Laurie	Becky	Dr	Otago University		NEW ZEALAND
Le	Lien	Mrs	Australian Genome Research Facility	VIC	AUSTRALIA
Le Cao	Kim-Anh	Dr	Queensland Facility for Advanced Bioinformatics	QLD	AUSTRALIA
Legaie	Roxane	Miss	Queensland Facility for Advanced Bioinformatics	QLD	AUSTRALIA
Leong	Mee Lian	Ms	University f Malaysia		MALAYSIA
Li	Jason	Dr	Peter MacCallum Cancer Centre	VIC	AUSTRALIA
Li	Jun	Dr	QIMR	QLD	AUSTRALIA
Lin	Ruby	Dr	The University of New South Wales	NSW	AUSTRALIA
Liu	Ruijie	Ms	Walter & Eliza Hall Institute of Medical Research	VIC	AUSTRALIA
Loetsch	Claudia	Ms	Garvan Institute of Medical Research	NSW	AUSTRALIA
Loring	Jeanne	Prof	The Scripps Research Institute		USA
Lovell	David	Dr	CSIRO	ACT	AUSTRALIA
Lun	Aaron	Mr	Walter & Eliza Hall Institute of Medical Research	VIC	AUSTRALIA
Maag	Jesper	Mr	The Garvan Institute of Medical Research	NSW	AUSTRALIA
Macartney-Coxson	Donia	Dr	Institute of Environmental Science & Research		NEW ZEALAND
Maksimovic	Jovana	Dr	Murdoch Childrens Research Institute	VIC	AUSTRALIA
Makunin	Igor	Dr	The University of Queensland	QLD	AUSTRALIA
Marshall	Vikki	Ms	Australasian Genomics Technologies Association	QLD	AUSTRALIA
Mason	Elizabeth	Miss	The University of Queensland	QLD	AUSTRALIA
McAndrew	Damara	Dr	Genesearch	QLD	AUSTRALIA
McCabe	Mark	Dr	Garvan Institute of Medical Research	NSW	AUSTRALIA

Last Name	First Name	Title	Organisation	State	Country
McGrath	Annette	Dr	CSIRO	ACT	AUSTRALIA
McMahon	Jennifer	Dr	QIAGEN	VIC	AUSTRALIA
Miller	David	Mr	The University of Queensland	QLD	AUSTRALIA
Mir	Gisela	Dr	Peter MacCallum Cancer Institute	VIC	AUSTRALIA
Moore	Cath	Dr	QIAGEN	VIC	AUSTRALIA
Morris	Kevin	Dr	The University of New South Wales	NSW	AUSTRALIA
Nair	Radhika	Dr	Garvan Institute of Medical Research	NSW	AUSTRALIA
Nakasugi	Kenlee	Dr	The University of Sydney	NSW	AUSTRALIA
Ng	Milica	Ms	CSL	VIC	AUSTRALIA
Nones	Katia	Dr	The University of Queensland	QLD	AUSTRALIA
O'Connor	Timothy	Mr	The University of Queensland	QLD	AUSTRALIA
Oshlack	Alicia	Dr	Murdoch Childrens Research Institute	VIC	AUSTRALIA
Pal	Bhupinder	Dr	Walter & Eliza Hall Institute of Medical Research	VIC	AUSTRALIA
Papenfuss	Tony	A/Prof	Walter & Eliza Hall Institute of Medical Research	VIC	AUSTRALIA
Parks	Matthew	Dr	Griffith University	QLD	AUSTRALIA
Parnell	Grant	Mr	Westmead Millennium Institute	NSW	AUSTRALIA
Patel	Hardip	Dr	John Curtin School of Medical Research	ACT	AUSTRALIA
Patrick	Ralph	Mr	The University of Queensland	QLD	AUSTRALIA
Pattel	Dhru	Mr	OnQ Software	VIC	AUSTRALIA
Peters	Tim	Dr	CSIRO	NSW	AUSTRALIA
Phipson	Belinda	Dr	Murdoch Childrens Research Institute	VIC	AUSTRALIA
Powell	Joseph	Dr	The University of Queensland	QLD	AUSTRALIA
Print	Cristin	A/Prof	The University of Auckland		NEW ZEALAND
Quackenbush	John	Prof	Harvard School of Public Health		USA
Quek	Kelly	Miss	The University of Queensland	QLD	AUSTRALIA
Quinn	Michael	Dr	The University of Queensland	QLD	AUSTRALIA
Reinert	Knut	Prof	Freie Universität Berlin		GERMANY
Ricos	Michael	Dr	University of South Australia	SA	AUSTRALIA
Ritchie	William	Dr	Centenary Institute	NSW	AUSTRALIA
Ritchie	Matthew	Dr	Walter & Eliza Hall Institute of Medical Research	VIC	AUSTRALIA
Roberts	Aedan	Mr	The Children's Hospital at Westmead	NSW	AUSTRALIA
Roden	Daniel	Dr	Garvan Institute of Medical Research	NSW	AUSTRALIA
Rodriguez	Astrid	Miss	Griffith University	QLD	AUSTRALIA
Rohart	Florian	Dr	AIBN	QLD	AUSTRALIA
Rudd	Stephen	Dr	Queensland Facility for Advanced Bioinformatics	QLD	AUSTRALIA
Sadedin	Simon	Mr	Murdoch Childrens Research Institute	VIC	AUSTRALIA
Sandelin	Albin	A/Prof	The Bioinformatics Centre		DENMARK

Last Name	First Name	Title	Organisation	State	Country
Sanders	Katherine	Miss	Bond University	QLD	AUSTRALIA
Schreiber	Andreas	Dr	ACRF South Australian Cancer Genomics Facility	SA	AUSTRALIA
Schwaber	Jessica	Miss	AIBN	QLD	AUSTRALIA
Shackleton	Mark	Dr	Peter MacCallum Cancer Centre	VIC	AUSTRALIA
Shafiei	Farhael	Mr	Sigma-Aldrich	NSW	AUSTRALIA
Shi	Wei	Dr	Walter & Eliza Hall Institute of Medical Research	VIC	AUSTRALIA
Shi	Leming	Dr	Fudan University		CHINA
Sim	Chou Hung	Dr	Murdoch Childrens Research Institute	VIC	AUSTRALIA
Smyth	Gordon	Prof	Walter & Eliza Hall Institute of Medical Research	VIC	AUSTRALIA
Sundaravadanam	Yogi	Mr	Australian Genome Research Facility	VIC	AUSTRALIA
Swarbrick	Alex	Dr	Garvan Institute of Medical Research	NSW	AUSTRALIA
Taft	Ryan	Dr	The University of Queensland	QLD	AUSTRALIA
Tajouri	Lotti	A/Prof	Bond University	QLD	AUSTRALIA
Tang	Benjamin	Dr	Nepean Hospital/Sydney University	NSW	AUSTRALIA
Thompson	Erik	Prof	St. Vincent's Institute	VIC	AUSTRALIA
Thorne	Natalie	Dr	Walter & Eliza Hall Institute of Medical Research	VIC	AUSTRALIA
Tothill	Richard	Dr	Peter MacCallum Cancer Centre	VIC	AUSTRALIA
Tu	Thomas	Dr	Centenary Institute	NSW	AUSTRALIA
Van Der Hoek	Mark	Mr	SA Pathology	SA	AUSTRALIA
Vasic	Vivien	Ms	Monash Health Translation Precint	VIC	AUSTRALIA
Vasilevski	Oliver	Mr	Millennium Science		USA
Vincent	Benita	Dr	CSIRO	ACT	AUSTRALIA
Voss	Anne	Dr	Walter & Eliza Hall Institute of Medical Research	VIC	AUSTRALIA
Waddell	Nic	Dr	The University of Queensland	QLD	AUSTRALIA
Waltham	Mark	Dr	St Vincent's Institute of Medical Research	VIC	AUSTRALIA
Warburton	Trent	Mr	TrendBio Pty Ltd	VIC	AUSTRALIA
Wells	Christine	A/Prof	The University of Queensland	QLD	AUSTRALIA
Whang	Trystan	Mr	Macrogen Inc	NSW	AUSTRALIA
White	Steve	Mr	Sigma-Aldrich	NSW	AUSTRALIA
Wilson	Peter	Dr	The University of Queensland	QLD	AUSTRALIA
Wisnieswski	Karen	Ms	Sigma-Aldrich	NSW	AUSTRALIA
Wong	Justin	Dr	Centenary Institute	NSW	AUSTRALIA
Wong	Nick	Dr	Olivia Newton John Cancer And Wellness Centre	VIC	AUSTRALIA
Wong	Jason	Dr	The University of New South Wales	NSW	AUSTRALIA
Wong	Hong Soo	Ms	Griffith University	QLD	AUSTRALIA
Wood	David	Mr	Queensland Centre For Medical Genomics	QLD	AUSTRALIA
Yang	Jean	Dr	The University of Sydney	NSW	AUSTRALIA
Yarski	Mike	Dr	Millenium Science	VIC	AUSTRALIA

Australasian Genomic Technologies Association Conference

12–15 OCT 2014
Crown Promenade, Melbourne

SAVE THE DATE!

The 2014 AGTA conference (formerly known as AMATA) will cover themes that include cancer genomics, epigenomics, plant genomics, bioinformatics, transcriptomics, functional genomics and clinical sequencing.

We have invited an outstanding list of international and national speakers plus have engaged industry to showcase the best in new genomic technologies.

Please come and join us for some great networking and cutting edge science in the heart of the "most livable city" in the world.

CONFIRMED SPEAKERS

Dr Peter Campbell
Sanger Centre, Cambridge, UK

Prof Alexander (Zander) Myburg
University of Pretoria, South Africa

A/Prof Titus Brown
Michigan State University, USA

Prof Barbara Wold
Caltech, USA

Dr Iain Fraser
National Institutes of Health, USA

Dr Melissa Fullwood
Yale-NUS College, Singapore

HAVE A QUESTION? Contact **Leishman Associates**
Email: **brigitte@leishman-associates.com.au** Phone: **+61 3 6234 7844**

Leishman
it's all about people

THANK YOU TO OUR GENEROUS SPONSORS & EXHIBITORS

QIAGEN®

Roche

illumina®

millennium science

Fluidigm

STEMCELLS AUSTRALIA

Australian Bioinformatics Network

Gold Coast Convention Bureau

agrf Australian Genome Research Facility Ltd

QFAB

MACROGEN

Partek® turning data into discovery®

life technologies™

bio-strategy delivering technology

GENESEARCH

Agilent Technologies

SIGMA-ALDRICH®

OnQ software

TrendBio Partnerships in Science

SARSTEDT