# QMB Genome Biology Abstracts

## G1: Epigenetic reprogramming in mammalian development

Wolf Reik[1,2,3], Ferdinand von Meyenn[1], Melanie Eckersley-Maslin[1], Stephen Clark[1], Thomas Stubbs[1], Hisham Mohammed[1], Rebecca Berrens[1], Fatima Santos[1] & Wendy Dean[1]

[1]Epigenetics Programme, Babraham Institute, Cambridge CB22 3AT
[2]Centre for Trophoblast Research, University of Cambridge, CB2 3EG
[3]Wellcome Trust Sanger Institute, Cambridge CB10 1SA

Epigenetic information is relatively stable in somatic cells but is reprogrammed on a genome wide level in germ cells and early embryos. Epigenetic reprogramming appears to be conserved in mammals including humans. This reprogramming is essential for imprinting, and important for the return to naïve pluripotency including the generation of iPS cells, the erasure of epimutations, and perhaps for the control of transposons in the germ line. Following reprogramming, epigenetic marking occurs during lineage commitment in the embryo in order to ensure the stability of the differentiated state in adult tissues. Signalling and cell interactions that occur during these sensitive periods in development may have an impact on the epigenome with potentially long lasting effects. The epigenome changes in a potentially programmed fashion during the ageing process; this epigenetic ageing clock seems to be conserved in mammals.

Our recent work addresses the mechanisms and consequences of global epigenetic reprogramming in the germ line, and the role of passive and active mechanisms of DNA demethylation. Using single cell multi-epigenomics techniques, we are beginning to chart the epigenetic and transcriptional dynamics and heterogeneity during the exit from pluripotency, symmetry breaking, and initial cell fate decisions leading up to gastrulation. We are also interested in the potentially programmed degradation of epigenetic information during the ageing process and how this might be coordinated across tissues and individual cells.

## G2: Epigenome Remodelling in Cancer

Susan J Clark

Epigenetics Research Laboratory, Genomics and Epigenetics Division, Garvan Institute of Medical Research, Sydney, 2010, New South Wales, Australia.

**ABSTRACT**

A three-dimensional chromatin state underpins the structural and functional basis of the genome by bringing regulatory elements and genes into close spatial proximity to ensure proper, cell-type specific gene expression profiles. We performed Hi-C chromosome conformation capture and replication timing sequencing to investigate how three-dimensional chromatin organization in relation to the epigenome is disrupted in cancer. Our study provides new insights into the relationship between replication timing, long-range epigenetic deregulation and changes in higher-order chromatin interactions in cancer.

## G3: Links between DNA replication stress and telomere extension in cancer cells

Kartawinata, M.[1], Tong, A.S.[1], Stern, J.L.[1,2], Lamm-Shalem, N.[1], Cesare, A.J.[1], Bryan, T.M.[1]

[1]Children's Medical Research Institute, Westmead, NSW, Australia, [2]Current address: Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado USA

The ribonucleoprotein enzyme telomerase counteracts telomere shortening in germ cells, stem cells and a majority of cancers, allowing for unlimited cell division. Most normal cells do not express telomerase, so inhibition of telomerase is a promising avenue for development of specific anti-cancer treatments that should display few side effects, and be applicable to almost all cancers. Telomerase action can be inhibited in vivo by blocking the cellular pathways that direct telomerase to chromosome ends, i.e. telomerase 'recruitment' to telomeres. This process is highly regulated, but factors that regulate it are incompletely understood.

We have demonstrated that telomerase recruitment to telomeres in human cells is highly dependent on the DNA damage response kinases ATM and ATR[1]. ATR controls the cellular response to single-stranded DNA arising during replication fork stalling, and we also showed that replication stress increases the levels of telomerase at the telomere, in an ATR-dependent manner[1]. ATM phosphorylates the telomeric "shelterin" protein TRF1, leading to its dissociation from telomeres, that in turn leads to an increase in telomeric replication stress, and an increase in telomerase recruitment to telomeres. These observations represent a new understanding of the relationship between telomerase and the DNA damage response: both are key drivers of oncogenesis, and our data demonstrate that they are tightly linked.

We are currently investigating the mechanism of the link between DNA replication stalling and telomerase recruitment to telomeres. We have found that nuclear actin polymerisation plays a role in the restart of stalled replication forks, and is required for DNA replication. Polymerisation of actin and its import into the nucleus are also required for telomerase recruitment to telomeres. This implicates an involvement of the nuclear cytoskeleton in both replication of the genome by the DNA replication machinery, and in ensuring that the ends of the chromosomes are extended after replication has occurred.

1. Tong, A.S., Stern, J.L, Sfeir, A., Kartawinata, M., de Lange, T., Zhu, X.D. and Bryan, T.M. (2015) *ATM and ATR Signaling Regulate the Recruitment of Human Telomerase to Telomeres.* Cell Reports, 13(8):1633-46.

## G4: Cannabis use and epigenetics: EPIC insights into the human methylome

Osborne, A.J.,[1,2] Pearson, J.F.[2], Horwood, J.[3], Hore, T.[4], Macartney-Coxson, D.[5], Gemmell, N.J.[4], and Kennedy, M.K.[2]

[1]School of Biological Sciences, University of Canterbury, Christchurch, New Zealand [2]Department of Pathology, University of Otago, Christchurch, NZ, [3]Department of Psychological Medicine, University of Otago Christchurch, [4]Department of Anatomy, University of Otago, Dunedin School of Medicine, NZ, [5]Institute of Environmental Science and Research, Porirua, New Zealand,

Cannabis has been highly publicised of late, with controversy and debate surrounding legalisation and its application for medical purposes. These debates have emphasised the medicinal and therapeutic benefits of cannabis, but there is also strong evidence for negative psychosocial consequences of prolonged cannabis use. Much of the evidence on health effects derives from the work of the Christchurch Health and Development Study (CHDS) which has followed the lives of 1265 people since they were born in the 1970s. Many drugs impact on the pattern of epigenetic marks that control genome function, and while the effects of cannabis at a genomic level are not well understood, we do know that cigarette smokers, have distinct patterns of cytosine methylation compared with non-smokers and these "methylome" changes likely impact on the regulation of genes that underlie the effects of cigarettes.

Genome-wide analysis of methylation in blood of 96 CHDS subjects (48 cannabis smokers [24 of whom also used tobacco] vs. 48 controls) using the Illumina EPIC 850K array identified approximately 15,000 probes that were significantly differentially methylated at 5% differential methylation or greater (after adjustment for multiple testing) in tobacco + cannabis users compared to controls, with 99% of these sites hypermethylated. The most differentially methylated sites were in the AHRR and F2RL3 genes. Cannabis-only users displayed differential methylation at 4500 sites, with 99% of these hypomethylated. Approximately 1000 probes were hypermethylated in tobacco + cannabis users, but hypomethyated in cannabis only users, compared to controls; KEGG pathway analysis on these sites showed that these probes were most strongly represented in PI3K-Akt and oxytocin KEGG pathways.

These data suggest that cannabis has quite different methylome effects to tobacco smoking, and that there may be interaction effects which modify the methylome impacts when both cannabis and tobacco are smoked. These analyses may help to identify epigenome impacts of cannabis use, and point to gene regulatory processes that underlie both potentially positive as well as adverse consequences of its use.

## G5: Cohesin facilitates zygotic genome activation in zebrafish

Meier, M.[1], Grant, J.[1], Dowdle, A. E.[1], Thomas, A.[1], Gerton, J.[2], Collas, P.[3], O'Sullivan, J. M.[4,5], Horsfield, J. A.[1,5*1]

[1]Department of Pathology, University of Otago, Dunedin 9016, New Zealand. [2]Stowers Institute for Medical Research, 1000 E 50th Street, Kansas City, Missouri 64110, USA. [3]Department of Molecular Medicine, Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, Norway. [4]Liggins Institute, The University of Auckland, Private Bag 92019, Auckland, New Zealand. [5]Maurice Wilkins Centre for Molecular Biodiscovery, The University of Auckland, Private Bag 92019, Auckland, New Zealand.

At zygotic genome activation (ZGA), changes in chromatin structure are associated with new transcription immediately following the maternal-to-zygotic transition (MZT). The nuclear architectural proteins, cohesin and CCCTC-binding factor (CTCF), contribute to chromatin structure and gene regulation in a variety of cell types.

We show here that normal cohesin function is important for ZGA in zebrafish. Depletion of cohesin subunit Rad21 delays ZGA without affecting cell cycle progression. In contrast, CTCF depletion has little effect on ZGA whereas complete abrogation is lethal. Genome-wide profiling of Rad21 binding reveals a change in distribution from pericentromeric satellite DNA sequences and few locations including the *miR-430* locus (whose products are responsible for maternal transcript degradation), to gene regulatory elements as embryos progress through the MZT. After MZT, a subset of Rad21 binding occurs at genes dysregulated upon Rad21 depletion and overlaps pioneer factor Pou5f3, which activates early-expressed genes. Rad21 depletion disrupts the formation of nucleoli and RNA polymerase II foci, suggestive of global defects in chromosome architecture.

We propose that Rad21/cohesin redistribution to active areas of the genome is key to the establishment of chromosome organization and the embryonic developmental programme.

## G6: Dysregulated transcriptional response to differentiation signals in cohesin-mutant myeloid cells

Jisha Antony[1,2], Ian M. Morison[1] and Julia A. Horsfield[1,2]

[1]Department of Pathology, Dunedin School of Medicine, University of Otago, New Zealand
[2]Maurice Wilkins Centre for Molecular Biodiscovery, New Zealand

Cohesin is a multiunit protein complex known for its roles in cell division, DNA damage repair, and gene regulation. Insufficiency of cohesin function can lead to cell type-specific changes in gene transcription during development and in cancer. Recent large scale genome sequencing found that mutation in genes encoding cohesin subunits are key to the progression of myeloid leukemia[1].

The cohesin protein complex is made up of RAD21, SMC3, SMC1A/1B and STAG1/STAG2 subunits. Leukemogenic mutations in cohesin subunits are recurrent, mutually exclusive, and never lead to a complete loss of function. This suggests that wild type cohesin is important for normal hematopoietic development, and a reduction in dose contributes to leukemia[2].

We used the CRISPR-CAS9 gene editing system to modify the chronic myelogenous leukemia cell line, K562, to contain leukemia-associated mutations in either the RAD21 (E212*), SMC3 (R254*) or STAG2 (R614*) subunits of cohesin. STAG2 is on the X chromosome and hence mutation can lead to complete inactivation in males and mosaic in females (due to X inactivation)[3]. Strikingly homozygous STAG2 mutation altered cellular phenotype with the gain of adhesive properties. When K562 cells were induced to differentiate down the megakaryocyte lineage, cohesin mutant cells showed altered expression of differentiation markers depending on the dosage of cohesin reduction. Alteration in differentiation was accompanied by dysregulated transcription of lineage genes like RUNX1 and ERG. The results suggest that cohesin mutation alters transcription in K562 cells, changing their response to differentiation cues.

Identification of cohesin mutations offers the opportunity to develop new therapeutic strategies for leukemia. We found that heterozygous cohesin-mutant K562 cells were more sensitive to drugs targeting pathways predicted to interact with cohesin function.

**References**

1. Thota, S. & Viny, A.D. Genetic alterations of the cohesin complex genes in myeloid malignancies. **124**, 1790-8 (2014).
2. Fisher, J.B., McNulty, M., Burke, M.J., Crispino, J.D. & Rao, S. Cohesin Mutations in Myeloid Malignancies. *Trends in Cancer* **3**, 282-293.
3. Hill, V.K., Kim, J.-S. & Waldman, T. Cohesin mutations in human cancer. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1866**, 1-11 (2016).

## G7: Decoding the regulatory overlap between Obesity and Type-2 Diabetes using Hi-C genome structure and expression quantitative traits loci.

Tayaza Fadason[1], Cameron Ekblad[1], John R. Ingram[2], William Schierding[1], Justin M. O'Sullivan[1]

[1] Liggins Institute, The University of Auckland, Auckland, New Zealand

[2] The New Zealand Institute of Plant and Food Research, Mount Albert, Auckland, New Zealand

The mechanisms that underlie the association between obesity and type-2 diabetes have yet to be fully understood. Here we interpreted the combined impacts of diabetes and obesity associated single nucleotide polymorphisms (SNPs) by integrating data on the genes with which they physically interact and the functional (*i.e*. expression quantitative trait loci [eQTL]) outcomes associated with these interactions. We identified enrichment for spatially regulated genes involved in lipid metabolism in adipose, skeletal muscle, and pancreas (p-value = $1.57 \times 10^{-2}$). The spatial eQTL SNP-gene interactions occur in a tissue and disease specific manner. For example, obesity associated eQTL SNP-gene interactions occurred most frequently in the thyroid (23.16%) and tibial nerve (17.90%) while those associated with type-2 diabetes occurred most frequently in the thyroid (22.16%) and subcutaneous adipose (19.32%). Our results are consistent with differential regulation of genes by spatially connected regions that are marked by disease associated SNPs in tissues involved in regulating energy homoestasis and adiposity. Investigating these putative spatial SNP-gene interactions may shed more light on the development of obesity and type 2-diabetes.

## G8: Epigenetic genome control by non-coding RNAs and RNA processing factors

**Shiv Grewal**
Laboratory of Biochemistry and Molecular Biology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA
Email: grewals@mail.nih.gov

Genome-wide sequencing of the human genome and several model organisms has greatly facilitated research into genome structure and function. Much attention is now focused on understanding how the functional organization of the genome supports translation of genetic information into formation of different cell types during development, and how it contributes to disease progression. We have shown previously that distinct histone methylation patterns organize chromosomes into “open” euchromatin (H3K4me) and “closed” heterochromatin (H3K9me) domains, to modulate use of the genome. Our work has also revealed the roles for non-coding RNAs (ncRNAs) and RNA processing factors in the assembly of heterochromatin domains. In particular, we find that ncRNA, RNA processing factors and heterochromatin machinery are part of an adaptive cellular mechanisms that can reprogram the genome in response to different growth conditions. Our most recent work about connections between ncRNAs and heterochromatin assembly, which is critical for epigenetic gene regulation and genome stability, will be discussed.

Sponsored by the Maurice and Phyllis Paykel trust.

## G9: Novel lncRNAs, CUPID1 and CUPID2, modulate breast cancer risk by modulate response to DNA damage.

**Juliet French**

Qimr Berghofer Medical Research Institute

GWAS have identified 170 loci associated with breast cancer. Fine-mapping of these loci has shown that the putative causative variants at these loci are rarely in gene coding regions. Instead, that the majority of these variants fall in regulatory elements in DNA that often lie at distant physical sites to the genes that they regulate. One of the strongest breast cancer associations identified to date via GWAS is with SNP rs614367 at the 11q13 locus. To determine the candidate causal variant(s) underlying this association, we analysed 4,405 variants in 220,000 cases and controls from the Breast Cancer Association Consortium (BCAC) and identified two independent association signals for estrogen-receptor-positive tumors that map to a distal enhancer. We have previously shown that this enhancer regulates the *CCND1* gene. In addition to *CCND1*, this enhancer regulates two novel estrogen-regulated long noncoding RNAs, *CUPID1* and *CUPID2* that we identified by RNA CaptureSeq. We provide evidence that the risk-associated SNPs are associated with allelic imbalance of *CUPID1* and reduced chromatin looping between the enhancer and the CUPID1/2 bidirectional promoter. We further show that *CUPID1* and *CUPID2* are predominantly expressed in hormone receptor-positive breast tumors and play a role in modulating double strand break (DSB) repair pathway choice. These data reveal a novel mechanism for the involvement of this region in breast cancer.

## G10: miR-200 and the Quaking RNA binding protein control a large alternative splicing network and cell plasticity in epithelial cells.

Pillman, K.A.[1,2,3,4], Phillips, C.A.[1], Toubia, J.[1,3,4], Roslan, S. [1], Dredge K.[1,2], Bert A.G.[1,2], Neumann, D.P. [1], Li, X. [1], Conn S.J.[1,2], Liu, D. [1], Lumb, R. [1], Lawrence, D.M.[1,3,4], Schreiber, A.W.[1,3,4], Hollier B.G.[5], Khew-Goodall Y.[1,2], Selth L.A.[4,6], Goodall, G.J.[1,2,4] and Gregory, P.A.[1,2,4]

[1]Centre for Cancer Biology, Australia, [2]The University of South Australia, Australia, [3]ACRF Cancer Genomics Facility, Australia, [4]The University of Adelaide, Australia, [5]Australian Prostate Cancer Research Centre, Queensland University of Technology, Australia

[6]Dame Roma Mitchell Cancer Research Laboratories, Australia

Epithelial cell plasticity is required for cell differentiation during embryogenesis and contributes to metastatic cancer progression. The miR-200 family are critical gatekeepers of the epithelial state and are down-regulated as cells undergo epithelial-mesenchymal transition (EMT), leading to induction of the expression of hundreds of target genes. To determine which miR-200 targets were most relevant to cancer progression, we correlated, ranked and compared candidates for many experimental and clinical datasets, revealing the RNA binding protein Quaking (QKI) was consistently highly correlated with miR-200 activity.

Using deep RNA-sequencing and HITS-CLIP, we investigated the role of QKI in regulating alternative splicing during EMT. We determined that the miR-200c-regulated QKI controls the changes in a vast array of alternative splicing events during EMT. Furthermore, modulation of QKI expression results in minimal changes to gene expression but altered epithelial cell plasticity and phenotypic properties, indicating this can be primarily driven by alternative splicing. By HITS-CLIP, we determined that hundreds of these splicing events are regulated directly by QKI, with a convergence on genes within the actin cytoskeletal network. QKI levels also strongly correlate with levels of these splicing events in breast cancer cell lines and patient tumours. Although QKI is generally reduced in tumours relative to normal tissue, it is elevated in basal-like breast cancers and plays pleiotropic roles, both inhibiting tumour growth and driving cancer cell invasion.

Our findings demonstrate the existence of a miR-200-Quaking pathway that globally control alternative splicing and impacts on cancer-associated epithelial cell plasticity.

## G11: Developmental plasticity and its impacts on the genome

Elizabeth Duncan[1], Megan Leask[2] and Peter K. Dearden[3]

[1] School of Biology, University of Leeds [2] Genetics Otago, University of Otago

Polyphenisms and plastic events involve coordinated changes in gene expression across the genome. We have studied the activation of worker bee ovaries as a tractable example of a plastic event. Worker bee reproduction is repressed by the presence of a queen bee, and a pheromone she produces, Queen mandibular pheromone (QMP). Worker bees in the absence of QMP, can activate the ovaries and lay eggs. We have discovered the cell signalling pathway that links QMP to repression of reproduction, and examined epigenetic changes that are triggered by the loss of QMP leading to plastic changes in the ovary. In this talk, I will present evidence on the nature of QMP repression and its impact on the honeybee genome.

# G12: A nanopore sequencing-based multiplex assay for pharmacogenetics of clopidogrel and warfarin

Liau, Y[1], Cree, S.M.[1], Maggo, S.[1], Miller, A.L.[1], Pearson, J[2], Kennedy, M.A.[1]
[1]Department of Pathology, University of Otago, Christchurch, NZ, [2]Biostatistics and Computational Biology Unit, University of Otago, Christchurch, New Zealand

Genetic variations in several genes have been associated with response to clopidogrel and warfarin. Currently available assays either include only a small number of variants or require high capital instruments. Using a handheld nanopore-based sequencing device, minION (Oxford Nanopore technologies), we developed a multiplex assay to detect 26 Single Nucleotide Polymorphisms in seven pharmacogenes related to these drugs, selected based on Clinical Pharmacogenetics Implementation Consortium data and recent publications.

In total, 15 primer pairs (167 – 240 bps) covering 26 SNPs in *CYP2C19*, *ABCB1*, *B4GALT2*, *CYP2C9*, *VKORC1*, *GGCX*, and *CYP4F2* were designed using mfeprimer v3.0[1], and used in a one-tube multiplex PCR. After sample barcoding, all amplicons were sequenced by MinION using the 2D sequencing kit (SQK-LSK208) and R9.4 flowcell (Flo-MIN106) from Oxford Nanopore Technologies. Sample barcoding kit enabled a maximum of 96 samples in one single sequencing run. For assay validation, we included 12 reference samples. Albacore v1.0.8 was used for base-calling, Poretools[2] was used to write the fasta/fastq files and MarginAlign[3] (with the bwa mem[4] aligner) was used to align to reference and call the genotypes. Results are validated by Sanger sequencing.

All variants were accurately detected in all samples, except for one SNP, the rs7294 of *VKORC1* gene where 2 out of 5 heterozygous samples gave uneven allele proportion (< 25 % of alternate allele), thus complicating SNP calling. Other aligner including nanopolish[5] and graphmap[6] failed to improve the results for this particular SNP, suggesting that the bias originated from the sequencing or base-calling, rather than the alignment.

Here we present an assay to simultaneously detect multiple pharmacogene-related SNPs in numerous samples. The cost-effective feature of multiplexing and simple sequencing instrument sets this as a promising tool for practical pharmacogenetics testing, although sequencing quality improvement is needed.

References

1. http://mfeprimer.igenetech.com/
2. Nicholas J. Loman, Aaron R. Quinlan; Poretools: a toolkit for analyzing nanopore sequence data. *Bioinformatics* 2014; 30 (23): 3399-3401. doi: 10.1093/bioinformatics/btu555
3. Jain M1, Fiddes IT1, Miga KH1, Olsen HE1, Paten B1, Akeson M1. *Improved data analysis for the MinION nanopore sequencer*. "Nat Methods. 2015 Apr;12(4):351-6. doi: 10.1038/nmeth.3290. Epub 2015 Feb 16.
4. Li H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v1 [q-bio.GN]
5. Loman NJ, Quick J, Simpson JT. A complete bacterial genome assembled de novo using only nanopore sequencing data. Nat Meth. 2015;12(8):733-5.
6. Sović I, Šikić M, Wilm A, Fenlon SN, Chen S, Nagarajan N. Fast and sensitive mapping of nanopore sequencing reads with GraphMap. 2016;7:11307.

## G13: Genetic variation in genes of the GH/IGF-1 axis in human development and disease

Jain, L[1], Fadason T [1], Vickers, MH[1], O'Sullivan, JM [1], Perry, JK [1]
[1]Liggins Institute, University of Auckland, Auckland, NZ.

The spatial organisation of chromosomes inside the nucleus of a cell is linked to gene regulation and function. Genetic variation at loci can disrupt this unique arrangement of chromosomes in space, potentially leading to changes in gene expression that contribute to disease states. However, how or if alterations in 3D genome structure contributes to disease is still not well understood.
Genes related to the growth hormone/insulin-like growth factor-1 (GH/IGF-1) axis are crucial for normal growth and metabolism. Physiological functions mediated by this axis range from bone and muscle growth, reproduction, insulin synthesis and sensitivity, lipolysis and hepatic metabolism through to immune functions and neurogenesis. Compromised GH signalling and genetic variation in these genes have been implicated in disorders such as diabetes and cancer.
We hypothesise that polymorphisms which occur within the GH/IGF-1 axis genes have the potential to impact on disease phenotypes by altering or disrupting the spatial organisation of the genome.
We have collated a database of SNPs associated with GH/IGF-1 axis identified through genome-wide association studies (GWAS), online databases and published literature. We have used a computational pipeline that incorporates GWAS data, expression Quantitative Trait Loci (eQTL) associations, and spatial chromatin connections, to identify SNPs located in intron 1 of the *IGF2BP2* gene that exhibit significant spatial connectivity and are associated with changes in gene expression in cis and trans. Currently we are determining the functional effects of these polymorphisms using CRISPR-Cas9 genome editing and relevant enhancer analysis using appropriate *in vitro* cell models.

## G14: Our genomes: the interaction of genome and metagenomics paves the way towards personalized medicine

Jingyuan Fu[1,2,*], Alexander Kurilshikov[1], Marc Jan Bonder[1], Daria V. Zhernakova[1], Trang Le[1], LifeLines cohort, UMCG-microbiome group, TIFN group, Marten H. Hofker[2,†], Cisca Wijmenga[1], Alexandra Zhernakova[1]

[1] University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, the Netherlands.
[2] University of Groningen, University Medical Center Groningen, Department of Pediatrics, Groningen, the Netherlands.

[*]E-mail: j.fu@umcg.nl

The human gut microbiome plays a major role in the production of vitamins, enzymes, and other compounds that regulate our metabolic and immune systems. It can be considered as “the second human genome”, showing a remarkable interplay between host and microbes that has a major impact on our physiology. To study these complex interactions, we have built up LifeLines-DEEP, a multi-omics biobank that is part of LifeLines, a large population cohort study in the northern Netherlands. LifeLines-DEEP consists of 1,500 individuals (42% males, age range 18-81 years) for whom we have dietary, genetic, gut microbiome, immune and metabolic profiles. Our LifeLines-DEEP data show that the gut microbiome can be determined by host genetic factors as well as many other exogenous and intrinsic factors. Moreover, our study has revealed enormous additive and interaction effects of the human genome and metagenome on the host’s immune system, metabolism, circulating proteins, and diseases, thereby emphasizing the importance of both genome and metagenome in the development of personalized medicine.

## G15: Gut Microbes in Infants of Obese Mothers Promotes Hepatic Inflammation, Macrophage Polarization, and Fatty Liver in Germ Free Mice

Friedman, J.E.1

1Department of Pediatrics, Biochemistry & Molecular Genetics, University of Colorado Anschutz Medical Campus, Aurora, CO

Our central hypothesis is that an imbalance in gut microbes in infants born to OB mothers triggers remodeling of pro-inflammatory liver macrophage Mϕ leading to increased risk for obesity and Non Alcoholic Fatty Liver Disease (NAFLD), a disease that affects up to 38% of obese children. To test the hypothesis, we used Germ Free (GF) mice colonized with human infant gut microbiota Inf-Ob or Inf-NW stool at 2 wks and 4 mo. of age. GFM were colonized for 21d and sacrificed, or subjected to a 6wk Western Style Diet (WSD-45% fat, 30% Sucrose, 0.2% Cholesterol). Short-Chain Fatty Acids, Acetate, Propionate and Butyrate in the cecum were elevated in the Inf-Ob mice ($P<0.001$) consistent with previous findings in the human infant. 2wk Inf-Ob MB increased pro-inflammatory hepatic gene expression (Tnf, Infb1, Xbp1-spliced mRNA levels, $P<0.05$), total liver Mϕ ($P<0.05$) and demonstrated inflammation on H&E staining as well as increased collagen staining in the portal region. 4mo Inf-Ob MB increased resident, infiltrating and total Mϕ ($P<0.05$) and resident Mϕ had increased pro-inflammatory cytokine expression (Tnf, Il1β $P<0.05$). Bone Marrow Derived Macrophages (BMDM) from 2wk Inf-OB mice showed a pro-inflammatory shift with a significantly higher Il1β:Il10 gene expression ratio following LPS treatment ($P<0.05$). Compared with Inf-NW MB, primary, secondary and conjugated BAs were reduced in 2wk Inf-Ob mice ($P<0.05$). GF mice exposed to a 6wk WSD gained significantly more weight and fat mass ($P<0.01$), and triggered further histological evidence of NAFLD in Inf-Ob mice only. These findings suggest that infant MB dysbiosis induces liver inflammation, increases recruitment of alters BMDMϕ to the liver with evidence for early fibrosis, indicating that immune programming by MB. Thus, early dysbiosis may prime innate immune pathways in Mϕ that augment subsequent pro-inflammatory Mϕ activation and predispose infants of OB mothers to weight gain and NAFLD.

## G16: Stratification of TCGA melanoma patients according to TIL and PD-L1 expression using RNA-seq data.

Ahn, A..[1], Euan, RJ.[1], Chatterjee, A.[1], Eccles M.[1]

[1]Department of Pathology, University of Otago, Dunedin, NZ, [2]

Immune checkpoint inhibitors have revolutionised the treatment of melanoma. However only a subset of patients respond to treatment. The tumour microenvironment, namely the interaction between immune cells and tumour cells plays, a crucial role in the treatment outcome of immunotherapy. In order to better predict patient responses from immunotherapy a tumour stratification framework has been proposed based on the presence of Tumour Infiltrative Lymphocyte (TIL) and PD-L1 expression. Here we used RNA-seq data to categorise 469 melanoma patients in TCGA according to the presence of CD8+ TIL's and PD-L1 mRNA expression. A novel cell-type enrichment webtool called *xCell* was used to determine the abundance of CD8+ TIL to differentiate the TIL+ and TIL- group. Moreover PD-L1 median expression level was used to segregate the PD-L1+ and PD-L1- groups. Accordingly patients were categorised into 4 groups: $TIL^+/PD\text{-}L1^+$ (44.3%) , $TIL^-/PD\text{-}L1^+$ (5.5%), $TIL^+/PD\text{-}L1^-$(18.3%) and $TIL^-/PD\text{-}L1^-$ (31.8%). Two independent deconvolution methods (*CiberSort* and *MCPcounter*) were used to validate the grouping of $TIL^+$ and $TIL^-$. As previously shown in literature, the $TIL^+/PD\text{-}L1^+$ patient group had a favourable prognosis with a longer overall survival. In addition, the $TIL^+/PD\text{-}L1^+$ patient group had an upregulation of gene-signatures related to numerous immune pathways including the viral defensive pathway and cytokine expression whereas the TIL-/PD-L1- group had a downregulation of these immune-signatures. Overall, we demonstrate a method to stratify TCGA melanoma patients according to the presence of TIL and PD-L1 expression using RNA-seq data.

## G17: Sex specific RNA splicing in mammals

Rowena Lavery [1] , Massilva Rahmoun [2] , Sabine Laurent - Chaballier [3] , Nicolas Bellora [4] , Gayle K Philip [5] , Moïra Rossito [2] , Aleisha Symon [1] , Eric Pailhoux [6] , Florence Cammas [3] , Jessica Chung [5] , Stefan Bagheri - Fam [1] , Mark Murphy [7] , Vivian Bardwell [7] , David Zarkower [7] , Brigitte Boizet - Bonhoure [2] , Philippe Clair [8] , Francis Poulat [2] , Vincent R Harley [1]

1. Hudson Institute of Medical Research, Melbourne, VIC, Australia
2. Institute of Human Genetics, University of Montpellier, Montpellier, France
3. Institut de Recherche en Cancérologie de Montpellier, INSERM U896, Montpellier, France
4. Laboratorio de Microbiología Aplicada y Biotecnología, Universidad Nacional del Comahue, Bariloche, Argentina
5. VLSCI, LAB, Melbourne, VIC, Australia
6. INRA Biologie du Développement et Reproduction, Jouy - en - Josas, France
7. Department of Genetics, Cell Biology and Development, University of Minnesota, Minneapolis, USA
8. Université Montpellier 2, Montpellier GenomiX, Montpellier, France

In mammals, male fate is under the control of the master transcriptional regulator, SOX9: in its presence, somatic precursor cells of the embryonic gonads differentiate into Sertoli cells, the central organizers of testicular differentiation. Therefore, analyzing target genes of this transcription factor allows the elucidation of cellular commitment mechanisms at the genome level. With the use of ChIP-seq in murine and bovine wild-type testes combined with RNAseq from mouse testes lacking SOX9, we identified SOX9 target genes in the mammalian fetal gonad. SOX9 in fetal testes binds to a large set of genes conserved among mammals, including those with well-established roles in testis and ovary development. RNAseq analysis shows that testes and ovaries display sex-specific RNA splicing and that SOX9 mediates both target transcription and differential splicing of these target genes. Regions bound by SOX9 are predominantly 5’ proximal or intra-genic, and display specific genomic features we call "Sertoli Cell Signatures" or SCS. The SCS is conserved among mammals and comprises multiple binding motifs for the Sertoli reprogramming factors SOX9, GATA4 and DMRT1; indeed, independent DMRT1 ChIP-seq confirms the enrichment of the SCS. Bioinformatic analysis of SCS regions predicts novel regulatory mechanisms prompting functional validation, and examples will be discussed. Also, target genes are being validated by in vivo (XY KO or XX KI mice) and ex vivo (drug treatments of cultured gonads) approaches which will be described. This work is unravelling the regulatory networks during gonadal development and identifying novel sex determining genes, whose functions could be altered in patients with Disorders of Sex development, most of whose forms remain unexplained genetically.

## G18: Noncoding exons are universally alternatively spliced and are likely the modular functional unit of regulatory RNA

John S. Mattick, Ira W. Deveson, Martin A. Smith, Luis R. Arriola-Martinez and Tim R. Mercer

*Garvan Institute of Medical Research, Sydney*

Many if not most long noncoding RNAs (lncRNAs) are expressed in highly restricted patterns, commensurate with their emerging role as epigenetic guide molecules. However the restricted expression of lncRNAs has resulted in low sequence coverage in transcriptional profiling of complex tissues and consequent difficulties in building accurate transcript models. We have used RNA CaptureSeq to obtain high-depth, high-resolution coverage of transcription from human ch21 and its syntenic regions in mouse, which has revealed, surprisingly, that noncoding exons in both mRNAs and lncRNAs are universally alternatively spliced. In parallel, we have found that evolutionarily conserved RNA structures, of which there are thousands of families, are overwhelmingly confined within exons, with few crossing exon-exon boundaries. Both observations imply, and together force the conclusion, that the unit of structure-function in regulatory RNAs is the exon, combinations of which can create an enormous diversity of isoforms presumably required to organise a precise developmental ontogeny. This modular structure also provides enormous flexibility for adaptive radiation, and a firm basis for parsing the evolutionary history and the functional architecture of regulatory RNAs. It further suggests - given our previous observations that exons are preferentially located in nucleosomes and that alternatively spliced exons are pre-organised in transcription-splicing complexes - that exons are the primary unit of epigenetic regulation.

*Sponsored by ESR.*

## G19: Immune cell migration and genome organisation

Jacobson, E.[1], Perry, J. K.[1], Vickers, M.[1], O'Sullivan, J. M.[1]

[1]Liggins Institute, University of Auckland, Auckland, NZ

Neutrophils are the most common immune cell in the blood, and the first line of defence against pathogens. After differentiating in the bone marrow they must squeeze through tiny gaps in blood vessel walls to reach the bloodstream. During this migration into the blood, the nucleus is first deformed and then exposed to fluid shear stress – which must affect the shape of the nucleus. DNA is tightly packed into the nucleus, and its arrangement of loops, domains, and compartments contributes to gene regulation. As the nucleus deforms during migration, is DNA reorganised, and does this represent the final maturation step for neutrophils?

I have used retinoic acid to differentiate a leukemia cell line (HL60) into neutrophil-like cells. I then developed a continuous flow transwell assay to allow continuous collection of cells that have migrated through 5 or 14um diameter pores. I characterised the proportion of differentiated, dead, and apoptotic cells before and after migration. Having optimised this protocol, I collected cells for both genome organisation (Hi-C) and transcriptome (RNA-seq) analysis. After migration, cells were fixed in formaldehyde or preserved in Trizol. The former crosslinks the DNA, the first step in making a Hi-C library – a genome wide chromatin capture technique. The latter preserves RNA for extraction and RNA sequencing. Hi-C and RNA-seq libraries have been sequenced in triplicate for each condition (no migration, 5um pores, 14um pores). I will present the results of this data.

## G20: Analysis of the epigenetic regulation and three-dimensional organisation of ribosomal genes during malignant transformation reveals new targets for cancer therapy.

Jeannine Diesch[1], Elaine Sanij[1], Austen Ganley[2], Justin M. O'Sullivan[3], Gretchen Poortinga[1] and Ross D. Hannan[1,4,5,6]

*[1]Research Division Peter MacCallum Cancer Centre, Melbourne, Australia; [2]Massey University, New Zealand; [3]Liggins Institute, University of Auckland, New Zealand;*
*[4]Department of Biochemistry and Molecular Biology, University of Melbourne, Australia; [5]Department of Biochemistry and Molecular Biology, Monash University, Australia; [6]School of Biomedical Sciences, University of Queensland, Australia*

The ribosomal genes (rDNA) are transcribed by RNA polymerase I (Pol I) in the nucleoli from over 200 rDNA repeats per haploid mammalian genome. rDNA transcription is a key determinant of cellular growth and its dysregulation is a consistent feature of human cancers. Indeed, work from our laboratory has demonstrated that inhibition of Pol I can selectively kill tumour cells in vivo (Bywater et al, Cancer Cell 2012; Bywter et al, Nature Reviews Cancer 2013). Remarkably, despite rDNA transcription being limiting for growth, over 50% of the rRNA genes are transcriptionally silent at any one time and are reported to play an important role in regulating rDNA stability and general heterochromatin. Our studies utilising the transgenic Eμ-*Myc* mouse model of B-cell lymphoma demonstrate that previously silent rDNA repeats are reactivated as B-cells progress towards malignancy, which correlates with hyperactivation of Pol I and increased rDNA transcription. Critically our data demonstrate that reactivation of the rDNA repeats and subsequent nucleolar reorganisation is essential for tumour cell survival independent of rDNA transcription. One hypothesis we are testing is that nucleolar reorganization may drive changes in
global genome organization, resulting in gene expression changes that underlie tumour cell dependence
In order to assess epigenetic changes at the rDNA loci and to identify long distance rDNA interactions occurring during the transition to malignancy, we are utilising ChIP-seq and chromosome conformation capture sequencing (4C-seq) techniques at different stages of lymphoma progression in vivo. We will discuss our results which show that transition from premalignancy to malignancy is associated with robust changes in rDNA chromatin in Eμ-*Myc* B-cells as well as reorganization of rDNA-genome interactions with a significant increase in these interactions being detected in malignant cells.
Our data provide evidence of a role for rDNA in cancer progression beyond ribosome biogenesis and offer novel insights into the spatial and transcriptional dynamics of the rDNA-associated genome during malignant transformation.

## G21: Epigenetic driver of melanoma metastasis and a clue for novel methylation mediated gene activation mechanism?

Chatterjee, A.[1*] , Stockwell, P.A.[2], Leichter, A.L.[1], Ahn, A.[1], Rodger, E.J[1], Eccles, M.R.[1]

[1]Department of Pathology, University of Otago, Dunedin School of Medicine, NZ, [2]Department of Biochemistry, University of Otago, School of Biomedical Sciences. *for correspondence (aniruddha.chatterjee@otago.ac.nz)

Metastasis (i.e., spread of primary cancers to distant organs) is responsible for ~90% of cancer-related deaths. Remarkable progress has been made in the last decade to document genetic events that cause primary cancers. However, a big unsolved question is how a primary cancer cell becomes metastatic and what are the molecular events that underpin this process. Epigenetic alterations are increasingly implicated in metastasis, whereas very few genetic mutations have been identied as drivers of cancer metastasis. Recently, we mapped genome-wide DNA methylation patterns in three cutaneous primary and metastatic melanoma cell line pairs and identified Early B Cell Factor 3 (*EBF3*) as a potential metastasis-related epigenetic driver in melanoma [1].

Promoter DNA methylation has been strongly established as a gene silencing mechanism. However, surprsingly, we discovered high promoter methylation of *EBF3* facilitates mRNA expression in melanoma cell lines [1]. We confirmed this *EBF3* methylation-expression relationship in two melanoma patient cohorts. Importantly, when the melanoma cells were treated with methylation inhibitor drugs, *EBF3* mRNA levels decreased [1], raising the interesting possibility that DNA methylation has a causal role in gene activation. In an analysis of melanoma patients (n = 458), we identified several genes that showed high promoter methylation associated with high expression and, for three genes (*GATA4, HOXD12, MGMT*), we confirmed that the activated transcripts were derived from the methylated promoter [2]. Recent work from other groups has provided further support in favour of this phenomenon. This combined body of evidence challenges the tightly held view that DNA methylation obligatorily silences promoters. We are aiming to use targeted epigenomic editing tools and a well-characterised melanoma model to connect methylation with chromatin biology and 3D genome organisation to establish causation for DNA methylation-mediated gene activation in the mammalian genome. In this talk, I aim to discuss the identification of *EBF3* as an epigenetic driver and then discuss some current models to understand potential methylation mediated gene activation mechanisms.

1. Chatterjee A, Stockwell PA, Ahn A, Rodger EJ, Leichter AL, Eccles MR: **Genome-wide methylation sequencing of paired primary and metastatic cell lines identifies common DNA methylation changes and a role for EBF3 as a candidate epigenetic driver of melanoma metastasis**. *Oncotarget* 2016.
2. Chatterjee A, Stockwell PA, Rodger EJ, Parry MF, Eccles MR: **scan_tcga tools for integrated epigenomic and transcriptomic analysis of tumor subgroups**. *Epigenomics* 2016.

## G22: In silico detection of candidate synthetic lethal gene pairs using cancer genomics data

S Thomas Kelly, Cassie Stylianou, Augustine Chen, Parry J Guilford and Michael A Black

Department of Biochemistry and Te Aho Matatu: Centre for Translational Cancer Research
University of Otago, Dunedin, NZ

Synthetic lethality is an emerging strategy for identifying novel drug targets in cancer. Current methods for identifying synthetic lethal interactions are heavily dependent on cell line studies, however additional utility can be gained through the use of existing genome-wide data from human tumours. These computational approaches, which use publicly available, whole tumour genomics data, have the potential to identify robust new drug targets, particularly when used as a complement to in vitro screening. Here we describe a novel computational technique for the identification of synthetic lethal interactions in whole tumours using the tumour suppressor gene E-cadherin (CDH1) as an example. Application of this approach will be demonstrated through the use of the SLIPT methodology to identify potential synthetic lethal partners of CDH1 in RNA-Seq data from The Cancer Genome Atlas (TCGA) breast and gastric cancer projects.

## G23: Epigenetic regulation of genes associated with breast cancer risk and progression

Dunbier AK[1,2], van Huffel S[1,2], Day R[1,2] and Horsfield J[2,3]

[1]Department of Biochemistry, University of Otago, Dunedin, NZ, [2]Centre for Translational Cancer Research, University of Otago, NZ, [3]Department of Pathology, University of Otago, Dunedin, NZ

More than three quarters of women diagnosed with breast cancer are affected by tumours that express oestrogen receptor α (ER). ER binds oestrogenic ligands, predominantly oestradiol, which are the key mitogens for ER+ve breast cancer and ER expression is a powerful positive predictor for response to anti-oestrogen therapy including tamoxifen and aromatase inhibitors. Single nucleotide polymorphisms within the region immediately upstream of the gene encoding the oestrogen receptor, *ESR1*, a gene on 6q25.1 encoding the estrogen receptor α (ER), have been found to be associated with increased risk of breast cancer[1,2]. Three genes within this region, *CCDC170*, *RMND1* and *ARMT1* are co-expressed with *ESR1* in ER+ve breast cancers and *CCDC170* expression is inversely associated with proliferation and recurrence free survival in breast cancer patients[3]. In addition, gene fusions between *ESR1* and *CCDC170* have been associated with aggressive and endocrine-resistant luminal B tumours[4].

This study aimed to investigate chromatin structure within the 6q25.1 region using targeted 3C (chromatin confirmation capture), 4Cseq (circularised chromatin confirmation capture analysis sequencing) and ATACseq (Assay for Transposase-Accessible Chromatin) to explore whether the spatial organisation of the chromatin might be associated with the regulation of genes and mutational patterns at this locus. Analyses revealed that the *ESR1* promoter region contacted a number of different upstream regions including the promoters of *RMND1* and *CCDC170.* A high frequency of interactions was observed in the intergenic region between *CCDC170* and *ESR1* where a number of SNPs associated with breast cancer risk have been found. In addition, contact between the 3' end of the *ESR1* gene and the 3' end of *CCDC170* suggests that there may be a regulatory interaction in this region. Overall, these data support the hypothesis that chromatin structure may drive the co-expression of the genes at this locus.

1. Turnbull C *et a*l. Nature Genetics 2010;42:504-7.
2. Zheng W et al. Nature Genetics 2009;41:324-8.
3. Dunbier A et al. PLoS Genetics 2011;7:e1001382.
4. Veeraraghavan J. et al., Nature Communications 2013;5:4577.

## G24: Untethering the genome: How loss of HP1α enhances malignant cell invasion

Hale, T.K.[1], Solomon, R.[1], Bond, S,D.[1], Mehta, S.[2], Braithwaite, A. [2], Wheeler, D.[3], Fitzsimons, H.L.[1]

[1]Institute of Fundamental Sciences, Massey University, Palmerston North, NZ, [2]Department of Pathology, University of Otago, Dunedin School of Medicine, NZ, [3]Nextgen Bioinformatic Services, Palmerston North, NZ.

Altered nuclear morphology and chromatin patterning is observed in neoplasia, with loss of heterochromatin associated with more aggressive malignancies. Members of the HP1 family modulate chromatin plasticity, with the HP1α paralog required to maintain the condensation of transcriptionally silent heterochromatin and to sequester it at the inner periphery of the nuclear envelope. While increased HP1α expression is proposed to aid the onset of malignancy, the acquisition of an invasive phenotype correlates with the loss of this architectural protein in many solid tumours including those of the thyroid, kidney, colon and breast.

By modulating the level of HP1α in breast cancer cell lines and a *Drosophila* model of epithelial cell invasion we have demonstrated that HP1α is a suppressor of invasion. Knock-down of HP1α, but not HP1β, in the poorly invasive breast cancer cell line MCF7 results in an increase in invasive potential that is accompanied by changes in nuclear morphology and malleability, the adhesive properties of these cells and their ability to resist a DNA damaging agent. These changes correspond with the differential expression of genes in pathways that regulate altered cytoskeleton, cell-cell and cell-matrix interactions. Together these data demonstrate that loss of HP1α results in an invasive phenotype that is an intermediary of the epithelial to mesenchymal transition spectrum. We propose that this phenotype is a consequence of the nuclear reorganisation that results from the untethering of the genome from the nuclear periphery upon loss of HP1α.

## G25: Inhibition of a K9/K36 Demethylase by an H3.3 Point Mutation Found in Paediatric Glioblastoma

Hsiao P J Voon[1], Maheshi Udugama[1], Linda Hi[1], Jeff Mann[2] and Lee Wong[1]

[1] Epigenetics and Chromatin Laboratory, Monash University, Melbourne, Australia.
[2] Monash Genome Modification Platform, Monash University, Melbourne, Australia.

Brain cancers are the leading cause of cancer-related mortality in children and young people. The lack of effective treatments for paediatric gliomas means these cancers are invariably fatal with a median survival of only 12 months. Sequencing of paediatric gliomas has identified two common substitution mutations (K27M and G34R) in genes encoding histone H3.3. The substitution of methionine for lysine at position 27 (H3.3 K27M) acts by recruiting and inhibiting PRC2, a histone lysine methyltransferase which mediates trimethylation of H3K27 (H3K27me3). We report that the H3.3 G34R mutation (glycine to arginine substitution) acts in an analogous manner by inhibiting KDM4, a histone lysine (K9/K36) demethylase. We introduced a single-copy H3.3 G34R targeted mutation in mouse ES cells and observed gains in H3K36me3 and H3K9me3 across the genome, in a manner which phenocopied Kdm4 triple-KO cells. *In vitro* and *in vivo* immunoprecipitation assays demonstrated that the H3.3 G34R mutant binds Kdm4 with high affinity while simultaneously inhibiting enzymatic activity.

These studies show that H3.3 G34R acts in a dominant negative fashion by inhibiting Kdm4, triggering genome-wide alterations to promote oncogenesis. We propose a general model where oncogenic histone mutations are analogous to chromatin modifications, and alter the genome through interactions with epigenetic erasers, writers and readers.

## G26: Visualising polymorphisms in the mitochondrial genome - a new tool in mitochondrial transfer analysis

M.J. McConnell[1,2] and M. Rowe[1,2]

[1]Centre for Biodiscovery, Victoria University of Wellington, New Zealand, [2]Malaghan Institute of Medical Research, Wellington, New Zealand.

The transfer of mitochondria between mammalian cells has been documented in multiple in vitro and in vivo systems, although what controls this process is not understood, although cell stress or mitochondrial injury may alter the incidence of mitochondrial transfer between cells. Literature published in the field of intercellular mitochondrial transfer relies heavily on fluorescent labelling and confocal microscopy to understand this process. Instead, we have taken a genetic approach to accurately detect and quantify inter-cellular mitochondrial transfer. Leveraging mitochondrial DNA SNPs between mitochondrial donor and recipient cells, a novel strategy for *in situ* molecular genotyping has been developed. SNP-specific 'padlock' oligonucleotides serve as a template for rolling circle amplification, exploiting the enhanced fidelity of a thermostable Taq ligase. This enables detection of potentially rare mtDNA variants derived from the donor cell within a recipient cell. In multiplexed reactions, detection of two (or more) variant mtDNAs present within a co-culture is possible, allowing bidirectional study of intercellular transfer of mitochondrial genomes. Application of this strategy allows us to trace and quantify the transfer of mitochondrial genomes, and provides robust and visually striking data to support our microscopic observation of this phenomenon.

## G27: The non-coding genome: from genome-wide association study signal to causal variant

Tony Merriman
University Of Otago

Recent genome wide association studies have revealed several dozen genetic variants associated with serum urate levels, but the precise molecular mechanisms by which they affect serum urate are unknown. Here we tested for functional linkage of the maximally-associated genetic variant *rs1967017* at the *PDZK1* locus to elevated *PDZK1* expression.

We performed expression quantitative trait locus (eQTL) and likelihood analyses followed by gene expression assays. Zebrafish were used to determine the ability of *rs1967017* to direct tissue-specific gene expression. Luciferase assays in HEK293 and HepG2 cells were used to measure the effect of *rs1967017* on transcription amplitude.

PAINTOR analysis revealed *rs1967017* to be the most likely causal variant ($P_{Posterior}$=0.93) and *rs1967017* to be an eQTL for *PDZK1* in the colon and small intestine ($P_{Posterior}$=0.93). The region harboring *rs1967017* was capable of directly driving green fluorescent protein expression in the kidney and intestine of zebrafish embryos, consistent with a conserved ability to confer tissue-specific expression. The urate-increasing T-allele introduces a binding site for the transcription factor HNF4A. siRNA depletion of HNF4A reduced endogenous *PDZK1* expression in HepG2 cells. Luciferase assays showed that the T-allele of *rs1967017* gains enhancer activity relative to the C-allele with T-allele enhancer activity abrogated by HNF4A depletion

Our data predict that the urate-raising T-allele of *rs1967017* enhances HNF4A binding to the PDZK1 promoter, thereby increasing *PDZK1* expression. Because PDZK1 is a scaffold protein for many ion channel transporters, its increased expression may contribute to reduced excretion of uric acid.

## G28: Tissue-specific epigenetic markers and disease discovery

Morison IM[1], Weeks RJ[1], Macaulay EC[1], Alsaleh AA[1], Almomani SN[1], Day RC[2]

[1]Department of Pathology, Dunedin School of Medicine, University of Otago; [2]Department of Biochemistry, University of Otago

Remarkable diversity is imposed upon our genomes by the epigenome. This is most remarkably demonstrated by the marked phenotypic differences between cell types, each of which can be distinguished by a unique DNA methylome. We can utilising these differences to identify cells or cell-free DNA and diagnose disease states or predispositions.

For example, the placental methylome is substantially different than that of blood cells and can thus be used to unequivocally identify placentally derived circulating maternal DNA. Thus, if disease-associated differences in methylation are large enough, there is the potential to diagnose placental disorders early in pregnancy.

The methylome of childhood acute lymphoblastic leukaemia is markedly different than that of any adult blood or marrow cell type. We hypothesise that the leukaemia methylome reflects, and thereby provides a means to identify the cell lineage from which it is derived. We have used specific methylation markers of childhood leukaemia to identify methylated DNA in neonatal blood samples, confirming the existence of a distinct cell type that is likely to be the leukaemia progenitor.

These studies are made possible by the use of targeted massively parallel methylation sequencing which will be described.

## G29: Histone deacetylase activity is regulated following induction of long-term potentiation *in vivo*.

Kyrke-Smith, M.[1,2], Abraham, W.C.[2] and Williams, J.M.[1]

[1]Department of Anatomy and [2]Department of Psychology, The Brain Health Research Centre, Brain Research New Zealand - Rangahau Roro Aotearoa, University of Otago, Dunedin, New Zealand

Enhancement of synaptic transmission, through processes such as long-term potentiation (LTP), enables selective incorporation of neurons into an engram. How these processes are maintained over the long term is not well understood, although new gene expression is critical for the persistence of both LTP and engrams. In earlier work, we discovered that LTP induced *in vivo* is associated with both rapid upregulation and subsequent downregulation of gene expression. Analysis of gene networks formed suggested that this temporal shift in gene expression is partly mediated by histone deacetylases, epigenetic inhibitors of gene expression. To explore this further, we investigated the activity of HDAC1 and HDAC2 following the induction of LTP in the dorsal dentate gyrus *in vivo.* We observed a hitherto unrecognized bilateral increase in HDAC1 and HDAC2 activity. This occurred within minutes following the high-frequency stimulation that is well-documented to both increase expression of immediate early genes and generate LTP, solely in the ipsilateral hemisphere. The bilateral increase in HDAC activity was not reflected in parallel changes in HDAC mRNA or protein levels. ChIP-Seq analysis not only confirmed the rapid bilateral alteration in HDAC1 and HDAC2 activity but showed that distinct networks of genes were altered in the ipsilateral and contralateral hemispheres. Taken together, these data suggest that heightened HDAC activity occurs at the same time as heightened gene expression. This may represent an important negative regulation of the rapidly induced plasticity transcriptome which may protect the establishing engram from disruption and/or maintain the plastic capability of neurons following learning.

## G30: Non-conflict theories for the evolution of genomic imprinting

Hamish G. Spencer

Department of Zoology, University of Otago, Dunedin, New Zealand

Theories focused on kinship and the genetic conflict it induces are widely considered to be the primary explanations for the evolution of genomic imprinting. However, there have appeared many competing ideas that do not involve kinship/conflict. These ideas are often overlooked because kinship/conflict is entrenched in the literature, especially outside of evolutionary biology. Here I provide a critical overview of these non-conflict theories, focusing on some alternatives that may also provide tenable explanations of the evolution of imprinting for at least some loci.

## G31: Epigenetic memory in vertebrates

Julian R Peat , Oscar Ortega-Recalde , Michael P Collins, Olga Kardailsky, Morgan T Jones, Tim Hore
Department of Anatomy, University of Otago, Dunedin, OTAGO, New Zealand

Epigenetic modification can provide a mechanism for cells to 'remember' early developmental decisions in the absence of the signals which first initiated them. Methylation of CG dinucleotides is amongst the most iconic of all epigenetic systems because there is a well defined mechanism by which it transmits molecular memory following cell division. We have recently surveyed CG methylation in a wide range of vertebrates (>25 species) and characterised the methylome of a cartilagenous fish, the elephant shark (Callorhinchus milii). Despite last sharing a common ancestor with humans ~465 Mya, elephant shark shares many epigenetic characteristics with human, including an association between promoter methylation and gene silencing. These findings position the elephant shark as a valuable model to explore the evolutionary history and function of vertebrate methylation, and stimulate further questions surrounding what drives high levels of epigenetic memory and how it is reprogrammed between generations.

## G32: The epigenetics of cardiovascular diseases and associated risk factors.

*A/Prof Greg Jones*
*Department of Surgical Sciences, Dunedin School of Medicine*

Cardiovascular diseases and associated disorders, such as diabetes, obesity and dyslipidaemia, are leading contributors to morbidity and mortality in our community. While both inherited and environmental risk factors are significantly associated with these conditions, our current lack of understanding of gene-environment interactions has limited our ability to identify new adjunctive risk tools.

This talk will discuss the use of epigenetics, specifically genome-wide DNA methylation profiling, too identify novel risk markers for these conditions. Our data suggests that this approach may not only yield significant advances in disease risk prediction, but may also identify novel therapeutic targets amenable to drug treatments. By combining genetic, epigenetic and demographic risk information for each study participant this research represents an important step towards the delivery of personalised medicine for the management of chronic disorders.

## G33: Finding a function for genetic variants associated with abdominal aortic aneurysm

Marsman, J. [1], Phillips, V.L.[1], Horsfield, J.A.[2], Jones, G.T.[1]

[1]Department of Surgical Sciences, University of Otago, Dunedin School of Medicine, NZ, [2]Department of Pathology, University of Otago, Dunedin School of Medicine, NZ.

Abdominal aortic aneurysm (AAA) affects 5-7% of the male population over the age of 65 and is a major cause of sudden death in elderly. While AAA has some overlapping genetic and environmental risk factors with atherosclerosis, there are substantial differences, and AAA-specific medication is lacking. A recent meta-analysis of genome-wide association studies has identified four novel genetic variants associated specifically with AAA[1]. These variants are located in non-coding DNA, and their function is unclear. We aim to identify which genes are regulated by AAA-associated genetic variants, and how these variants alter gene expression. Using an enhancer luciferase assay, we have identified variants that alter enhancer activity, indicating that they may dysregulate the expression of their target genes. Our ongoing work uses circular chromosome conformation capture to detect interactions anchored at these variants, with the aim of identifying which genes they contact. We think that the results from this study will uncover new biological mechanisms of action for AAA, which may lead to the development of AAA-specific medication.

1. Jones, G.T., Tromp, G., Kuivaniemi, H., Gretarsdottir, S., Baas, A.F. *et al* (2016). *Meta-Analysis of Genome-Wide Association Studies for Abdominal Aortic Aneurysm Identifies Four New Disease-Specific Risk Loci.* Circ Res **120**, 341-353.

## G34: The MinION nanopore sequencer – toy or tool?

Kennedy, M.A. [1], Cree, S.M.[1], Miller, A.L.[1], Pearson, J.[2], Maggo, S.[1], Liau, Y.[1], Ton, K.[1], de Jong, L., Walker, L.C. [1]

[1]Department of Pathology, University of Otago, Christchurch, NZ, [2]Biostatistics and Computational Biology Unit, University of Otago, Christchurch, New Zealand

The world's smallest and newest machine for sequencing DNA is the MinION device, developed by Oxford Nanopore Technologies Ltd (UK). This remarkable device is half the size of a cellphone, and on a single flow cell it can generate 10Gb or more of long-read sequence data in real-time. We have been working with the MinION in my research laboratory for three years. I will describe some of the projects in which we have used the MinION, mainly in the area of pharmacogenetics, as well as cDNA analysis, and provide perspectives on the advantages and challenges of working with the device.

# G35: New tools for diet analysis: nanopore sequencing of unamplified genomic DNA from stomach contents to detect with high resolution what rats are eating

William Pearman, Georgia Breckell, James Dale, Olin K. Silander, Nikki E. Freed
Institute of Natural and Mathematical Sciences, Massey University, Auckland 0745, New Zealand

**Abstract**

New Zealand is unique in its lack of native terrestrial mammalian predators. Invasive mammalian predators, notably rats, have had severely detrimental effects on the viability of native species. Identifying the specific diet components of such animals is critical for understanding which species are most impacted and for determining where conservation efforts should be targeted. Previously, a wide variety of methods have been applied to quantify diet components, ranging from visual inspection of gut contents, to stable isotope analysis, to time-lapse video[1-3]. More recently, DNA-based methods have been used, with diet components being identified through a metabarcoding approach[4-7]. Metabarcoding methods first require a PCR-based amplification step, using primer sets targeting conserved regions across many taxa. This method has significant drawbacks[8,9]; notably the PCR step introduces biases in which species are amplified and thus detected.

Better, faster, and easier methods for diet analysis could qualitatively change our knowledge about diets, and thus further our understanding of the animals being studied. We present results from nanopore-based DNA sequencing of native, unamplified DNA isolated from rat stomach contents to determine the diet of these invasive predators. This represents the first instance, to our knowledge, of diet analysis using such an approach. Our results indicate, with high resolution and depth, the diets of 24 rats from three different environments in New Zealand and provide a framework for using this approach for diet analysis of other species and/or sample types.

**References**

1. Major, H. L., Jones, I. L., Charette, M. R. & Diamond, A. W. *Variations in the diet of introduced Norway rats (Rattus norvegicus) inferred using stable isotope analysis*. Journal of Zoology **271,** 463–468 (2007)
2. Brown, K. P. & Brown, K. P. *Predation at nests of two New Zealand endemic passerines; implications for bird community restoration*. Pac. Conserv. Biol*.* **3,** 91–98 (1997).
3. Brown, K. P., Moller, H., Innes, J. & Jansen, P. I*dentifying predators at nests of small birds in a New Zealand forest*. Ibis **140,** 274–279 (1998).
4. Lopes, C. M. *et al. DNA metabarcoding diet analysis for species with parapatric vs sympatric distribution: a case study on subterranean rodents.* Journal of Heredity **114,** 525–536 (2015).
5. Iwanowicz, D. D. *et al. Metabarcoding of Fecal Samples to Determine Herbivore Diets: A Case Study of the Endangered Pacific Pocket Mouse*. PLoS ONE **11,** e0165366 (2016).
6. Soininen, E. M. *et al. Sources of variation in small rodent trophic niche: new insights from DNA metabarcoding and stable isotope analysis.* Isotopes Environ Health Stud **50,** 361–381 (2014).
7. Khanam, S., Howitt, R., Mushtaq, M. & Russell, J. C. *Diet analysis of small mammal pests: A comparison of molecular and microhistological methods.* Integrative Zoology **11,** 98–110 (2016).
8. Taylor, H. R. & Harris, W. E. *An emergent science on the brink of irrelevance: a review of the past 8 years of DNA barcoding.* Molecular Ecology Resources **12,** 377–388 (2012).
9. Tedersoo, L. *et al. Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi*. MycoKeys **10,** 1–43 (2015).

## G36: The DMAP package for DNA methylation analysis

Stockwell, P.A[1] & Chatterjee, A.[2]

[1]Department of Biochemistry, University of Otago, School of Biomedical Sciences,
[2]Department of Pathology, University of Otago, Dunedin School of Medicine, NZ.

The DMAP package [4] was developed to meet needs arising from a project on establishing base level CpG methylation for a cohort of normal individuals. The work posed special problems for managing the quantity of data involved and required extensive interactive planning to establish a pipeline for the processing [6]. A feature of the development was the need for close collaboration between the software development and the laboratory and conceptual sides of the work.

This paper will cover factors contributing to the package development, conceptual complications that were encountered and will examine briefly some of the results which have been achieved with it in human and non-human RRBS and WGBS projects [1-3,4-5]

1. Chatterjee A, Macaulay EC, Rodger EJ, Stockwell PA, Parry MF, Roberts HE, Slatter TL, Hung NA, Devenish CJ, Morison IM: **Placental Hypomethylation Is more Pronounced in Genomic Loci Devoid of Retroelements**. *G3 (Bethesda)* 2016.
2. Jones AM, Ferguson P, Gardner J, Rooker S, Sutton T, Ahn A, Chatterjee A, Bickley VM, Sarwar M, Emanuel P *et al*: **NRAS and EPHB6 mutation rates differ in metastatic melanomas of patients in the North Island versus South Island of New Zealand**. *Oncotarget* 2016.
3. Chatterjee A, Stockwell PA, Rodger EJ, Duncan EJ, Parry MF, Weeks RJ, Morison IM: **Genome-wide DNA methylation map of human neutrophils reveals widespread inter-individual epigenetic variation**. *Scientific reports* 2015, **5**:17328.
4. Stockwell PA, Chatterjee A, Rodger EJ, Morison IM: **DMAP: differential methylation analysis package for RRBS and WGBS data**. *Bioinformatics* 2014, **30**(13):1814-1822.
5. Chatterjee A, Ozaki Y, Stockwell PA, Horsfield JA, Morison IM, Nakagawa S: **Mapping the zebrafish brain methylome using reduced representation bisulfite sequencing**. *Epigenetics* 2013, **8**(9):979-989.
6. Chatterjee A, Stockwell PA, Rodger EJ, Morison IM: **Comparison of alignment software for genome-wide bisulphite sequence data**. *Nucl Acids Res* 2012, **40**(10):e79.

## G37: Integrated profiling of single-cell chromatin accessibility and transcriptome reveals regulatory heterogeneity

Longqi LIU[1,2,4*], Chuanyu LIU[1,2,3*], Liang WU[1,2*], Yue YUAN[1,2,3], Mingyue WANG[1,2,3], Fang CHEN[1,2,6], Zhouchun SHANG[1,2,5#], Xun XU[1,2#]

1 BGI-Shenzhen, Shenzhen, China
2 China National GeneBank, BGI-Shenzhen, Shenzhen, China
3 BGI Education Center, University of Chinese Academy of Sciences, Shenzhen, China
4 Harbin Institute of Technology Shenzhen Graduate School, Xili University Town, Shenzhen, China
5 Department of Regenerative Medicine, Tongji University School of Medicine, Shanghai, China
6 MGI Tech Co., Ltd. , Shenzhen, China

* These authors contributed equally to this work.
# Correspondence should be addressed to Z.S (shangzhouchun@genomics.cn) or X.X (xuxun@genomics.cn)

The advances of single cell sequencing technologies have greatly improved our understanding of heterogeneity in terms of genetic, epigenetic and transcriptional regulation within cell populations. We and other groups have developed single-cell whole genome, exome, methylome and transcriptome technologies and applied these approaches to analyzing the complexity of cell populations in tumorigenesis, developmental process and cellular reprogramming. Meanwhile, single-cell epigenome techniques including single cell ChIP-seq, ATAC-seq, DNase-seq and Hi-C, have been developed to decipher histone modifications, chromatin accessibility landscapes, and 3D chromatin contacts respectively in single cells. Integrative analysis of single-cell multimodal data is critical for accurate dissection of cell-to-cell variation within certain cell populations. Recent progress on measuring multi-omics in the same cells has enabled analysis of associations between different layers of regulation on gene expression. So far, the relationship between chromatin accessibility and gene expression has not been investigated at the resolution of one single cell. To address this question, we report scCAT-seq, a technique for simultaneous assay of chromatin accessibility and transcriptome within the same single cell. By applying scCAT-seq to different cancer cell types, we identified *trans*-factors as bridges linking accessibility variation of *cis*-regulatory elements to cell-type-specific gene expression across single cells. We further characterized subpopulations within cancer cells and uncovered the regulatory clues that drive transcriptional heterogeneity. Together, scCAT-seq is a promising tool for the joint analysis of multimodal data of single cells, which also offers the potential for clinical applications such as preimplantation screening and cancer diagnosis.

**Key Words**

single cell, sequencing, epigenomics, chromatin accessibility, transcriptome

## GENOME BIOLOGY POSTER SESSION

**Poster summary**

| **Genome Biology Posters** | | |
|---|---|---|
| (G38) | **Scott Cohen**<br>Children's Medical Research Institute<br>*Electron microscopy of human telomerase* | **Trades Area** |
| (G39) | **Telfer, E.J**<br>Scion<br>*A preliminary assembly of the massive 25 Gb Pinus radiata megagenome.* | **Trades Area** |
| (G40) | **Hunter, F.W**<br>University of Auckland<br>*Development and application of CRISPR functional genomic capability at the University of Auckland* | **Trades Area** |
| (G41) | **Zhu, Z.Z.**<br>BGI<br>*Application of cell-free RNA sequencing on BGISEQ-500 platform* | **Trades Area** |
| (G42) | **Leichter, A.L.**<br>University of Otago<br>*Development of single-cell methodologies to interrogate DNA methylation patterns at a whole genome scale* | **Trades Area** |
| (G43) | **Chunyu Geng**<br>BGI<br>*Performance of Enzymatic Fragmentation WGS Library on BGISEQ-500 Platform* | **Trades Area** |
| (G44) | *Withdrawn* | |
| (G45) | **Issam Mayyas**<br>University of Otago<br>*Mechanisms of Active DNA Demethylation in somatic cells* | **Trades Area** |
| (G46) | **Suzan Momani**<br>University of Otago<br>*Predicting altered methylation patterns in early pre-eclampsia* | **Trades Area** |
| (G47) | **M. Safavi**<br>Environmental Protection Agency, New Zealand<br>*A journey of GMO regulation in New Zealand* | **Trades Area** |
| (G48) | **Eccles, M.R.**<br>University of Otago<br>*Epigenetic driver of melanoma metastasis identified by genome-wide analysis* | **Trades Area** |
| Guest | **Billy Apple**<br>**n=1** | **Trades Area** |

## G38: Electron microscopy of human telomerase

Rosalba Rothnagel[2], Tram Phan[3], George Lovrecz[3], Timothy Adams[3], Tracy Bryan[1], Michael Parker[4], Ben Hankamer[2], Scott Cohen[1*]

[1]Children's Medical Research Institute, Westmead NSW Australia

[2]Institute for Molecular Bioscience, University of Queensland, St. Lucia QLD Australia

[3]CSIRO Materials Science & Engineering, Parkville VIC Australia

[4]St. Vincent's Institute for Medical Research, Fitzroy VIC Australia

Telomeres, the repetitive DNA-protein complexes at the ends of linear chromosomes, shorten with each cycle of DNA replication, providing a counting mechanism to limit the number of times a cell can divide. Most cancer cells have activated the ribonucleoprotein enzyme telomerase to add telomeric DNA repeats and counteract telomere shortening, allowing for unlimited proliferation. Although inhibition of telomerase has been considered a promising approach to cancer therapy for more than two decades, its low cellular abundance (~50-100 copies/cell) and challenging biochemistry have stymied development of small-molecule inhibitors.

We reported the purification and composition of the core human telomerase enzyme complex, consisting of two molecules each of: i) the telomerase reverse transcriptase catalytic protein; ii) telomerase RNA; and iii) the RNA-binding protein dyskerin (*1*). Building on this knowledge we developed an over-expression system in suspension HEK-293T cells that yields ~400-fold greater activity over endogenous levels; this system is providing sufficient telomerase for electron microscopy studies. We have determined a low-resolution structure by negative-stain EM, revealing an elongated, bilobal structure, and are currently in the early stages of imaging with cryo-EM. Our long-term aim is to apply structure-guided design to the development of small-molecule telomerase inhibitors.

(*1*) Cohen SB, et al. (2007). *Protein Composition of Catalytically Active Human Telomerase from Immortal Cells* Science. 315: pp 1850-1853.

## G38: A preliminary assembly of the massive 25 Gb *Pinus radiata* megagenome.

Telfer, E.J.[1], Sturrock, S.[2], Winkworth, R.C.[3], Graham, N.J.[1], Macdonald, L.J.[2], Frickey, T.[2], Wilcox, P.L.[4], and Dungey, H.S.[1]

[1]Forest Genetics, Scion, Rotorua, NZ, [2] Forest Industry Informatics, Scion, Rotorua, NZ, [3]Institute of Fundamental Sciences, Massey University, Palmerston North, NZ, [4]Department of Mathematics and Statistics, University of Otago, Dunedin, NZ.

A well assembled and annotated reference genome can faciliate a range of research activities, including marker development for breeding and selection, functional genomics for understanding complex gene interactions and comparative evolutionary studies. Conifer genomes dwarf those of other tree species. For example, the *Pinus taeda* genome is 21 Gb in size compared to 420 and 640 Mb for *Populus trichocarpa* and *Eucalyptus grandis*, respectively; much of this difference reflects paleo-polyploidy resulting in diverse gene families groupings and a massive increase in the amount of repetitive DNA. Development of an annotated genome resource for *Pinus radiata* therefore represents a significant commitment. Over the last four years, we have generated 40 lanes of Illumina HiSeq pair-end short reads as well as 3-, 5- and 10-kb mate-pair libraries. However, even using the closely related *Pinus taeda* as a reference, the computational resources needed to assemble these data have outstripped those available in New Zealand. Recently, progress has been made using a newly acquired high-end server and the addition of Pacbio long read data. Here we present an update on our progress towards an annotated reference genomes for the nuclear and organellar compartments of *P. radiata*. We have assembled the first complete *P. radiata* chloroplast genome, including the resolution of a polymorphic 55 bp repeat region, and have initial *de novo* assemblies for both the nuclear and mitochondrial genomes.

## G39: Development and application of CRISPR functional genomic capability at the University of Auckland

Hunter, F.W.1,2, Tsai, P.3, Bohlander, S.K.2,3, Print, C.G.2,3, Wilson, W.R.1,2

1Auckland Cancer Society Research Centre, University of Auckland, Auckland, NZ, 2Maurice Wilkins Centre for Molecular Biodiscovery, NZ, 3Department of Molecular Medicine and Pathology, University of Auckland, Auckland, NZ.

Prokaryotic CRISPR-Cas systems have been adapted for a range of applications in eukaryotic model organisms, including genome editing and functional genomics. The latter entails the performance of whole-genome screens in complex libraries generated via CRISPR-mediated gene knockout, inhibition or activation, and offers incisive tools for investigating genetic determinants of virtually any cell-autonomous phenotype.

We report the establishment of an efficient, end-to-end pipeline for genome-scale CRISPR knockout screens at the University of Auckland. We optimised methods for the manufacture of ultracentrifuge-concentrated lentiviral vectors encoding GeCKO v2 library (Feng Zhang, Broad Institute). Improved protocols for PCR-amplification and next-generation Illumina sequencing were developed. Shiny server assets were developed to automate trimming, alignment and quality assessment of sequence data, to interface with third-party statistical analysis tools (RIGER, MAGeCK, HiTSelect and RSA) and to integrate genomic meta-data (pathway analysis, gene ontology, functional clustering). Complex gene knockout libraries have been generated in the KBM-7 (CML), NZM37 (melanoma), UT-SCC-74B, FaDu and UT-SCC-54C (all HNSCC), BT-474, MDA-MB-453 and MDA-MB-231 (all HER2-driven mammary carcinoma) and SiHa (cervical carcinoma) cell lines. The resulting libraries were structurally characterised by deep sequencing and functionally characterised by selecting with the chemotherapeutic agent 6-thioguanine, resulting in the enrichment of single guide (sg) RNA targeting known 6-thioguanine sensitivity genes (*MLH2*, *MSH6*, *NUDT5* and *HPRT1*). These GeCKO libraries have been deployed in a series of studies to investigate the mechanism of action and genetic determinants of sensitivity/resistance to a number of investigational or approved anti-cancer agents, with data pertaining to the evaluation of DNA-crosslinking agents to be presented, where known (e.g. *SLFN11*, *FANCG*) and potentially novel (e.g. *PDCD10*, *CABIN1*, *NSUN2*) determinants were identified.

An efficient, well-characterised and fully-staffed pipeline for CRISPR knockout screens is now available at the University of Auckland. The technology has broad potential applications to functional genomic research in New Zealand.

## G41: Application of cell-free RNA sequencing on BGISEQ-500 platform

Zhu, Z.Z.[1], An, D.[1], Yang, T.T.[1], Geng, C.Y.[1], Liang, X.M.[1], Chen, F.[1,2], Jiang, H.[1], Mu, F.[1]

[1]BGI-Shenzhen, Shenzhen 518083, China

[2] MGI Tech Co., Ltd. , Shenzhen, China

Cell-free RNA (cfRNA) is well known as extracellular RNA that present outside of the cells from which they were transcribed. In Homo sapiens, cfRNAs have been discovered in bodily fluids such as venous blood, saliva, and so on. Although the biological function is not fully understood, cfRNAs can play a role in a variety of biological processes[1]. Here we construct a high throughput sequencing method to detect cfRNA from plasma.

The cfRNA was isolated from human plasma and divided into 6 equal aliquots, each aliquot was 10ng and spiked in the ERCC as control. 3 aliquots were treated with MGIEasy™ rRNA Depletion Kit V1 to remove the rRNA and the remained 3 aliquots didn't remove the rRNA. Then the RNA sequencing libraries were prepared following the MGIEasy™ mRNA Library Prep Kit V2 with some modifications and sequenced on the BGISEQ-500 with single end (SE) 50bp. Raw data were filtered and normalized to 22M reads per sample. The reads were mapped to the reference sequence, including genome sequence, gene sequence, globin RNA sequence, lncRNA sequence, ERCC sequence and virus database, and the FPKM value were calculated based on the mapped reads. The rRNA-removed group have higher genome mapping rate (82.66% to 72.93%), gene mapping rate (38.4% to 28.9%) and gene detection number (16757 to 12075) than the other group. Above 90% of the detected genes are low expressed gene (FPKM <1). Meanwhile, the rRNA-removed group get a little lower globin RNA mapping rate (0.01% to 0.07%), lncRNA mapping rate (1.18% to 1.38%) and virus database mapping rate (0.2% to 0.27%) than the other group. For ERCC control, the two groups can obtain reasonable Spearman correlation coefficient (0.978 and 0.981) and Pearson correlation coefficient (0.953 and 0.935). In a word, the cfRNA can be well sequenced on BGISEQ-500 platform and the data showed that there is quite a bit genes free from cell with low expression.

Iguchi, H., Kosaka, N., Ochiya, T. (2010). *Secretory microRNAs as a versatile communication tool*. Communicative & integrative biology. 3 (5): 478–81.

## G42: Development of single-cell methodologies to interrogate DNA methylation patterns at a whole genome scale

Leichter, A.L.[1*], Day, R.C.[2], Stockwell, P.A.[2], Wilson, M.[3], Guilford, P.[2], Eccles, M.R.[1] & Chatterjee, A.[1*]

[1]Department of Pathology, University of Otago, Dunedin School of Medicine, NZ, [2]Department of Biochemistry, University of Otago, School of Biomedical Sciences, [3]Department of Microbiology and Immunology, University of Otago, School of Biomedical Sciences. *for correspondence

To date DNA methylation studies have been carried out on bulk cell populations which provide an averaged snapshot of methylation but fail to detect individual cell-specific or rare cell subpopulation changes. There is enormous potential of biomedical applications for single-cell epigenomic technologies. Using a single cell methylation (SCM) technique would enable profiling of circulating tumour cells, exfoliated cells and cancer stem cells. This would provide critical detection and dynamic measurement of epigenetic abnormalities during disease initiation, evolution, relapse and metastasis. However, currently in Australasia low cell number or single cell epigenomics analyses are not well established, and therefore robust SCM investigations are not yet possible.

Previously, we have established analytical tools and methods for large-scale methylation analysis in bulk cells[1-3]. We are now establishing an experimental and analytical approach for the interrogation of single-cell DNA methylation patterns, specifically using single-cell reduced representation bisulfite sequencing (sc-RRBS) and whole genome bisulfite sequencing (sc-WGBS)[4]. Several single cell libraries have been prepared, yet alignment to the human bisulfite converted genome remains low. We are currently developing unique individual protocols and troubleshooting steps to optimise the progressive steps of the sc-RRBS protocol including; lysis, MSPI digestion, end repair/dA-tailing, adapter ligation, bisulfite conversion and PCR amplification. Streamlining this protocol will open new avenues to perform DNA methylation assessment on rare cell populations, such as circulating tumour cells and provide critical information on the phenotype of individual cells to assist clinicians in prediction of the reoccurence or metastasis of disease.

1. Chatterjee A, Stockwell PA, Ahn A, Rodger EJ, Leichter AL, Eccles MR: **Genome-wide methylation sequencing of paired primary and metastatic cell lines identifies common DNA methylation changes and a role for EBF3 as a candidate epigenetic driver of melanoma metastasis**. *Oncotarget* 2016.
2. Stockwell PA, Chatterjee A, Rodger EJ, Morison IM: **DMAP: differential methylation analysis package for RRBS and WGBS data**. *Bioinformatics* 2014, **30**(13):1814-1822.
3. Chatterjee A, Stockwell PA, Rodger EJ, Morison IM: **Comparison of alignment software for genome-wide bisulphite sequence data**. *Nucl Acids Res* 2012, **40**(10):e79.
4. Guo H, Zhu P, Guo F, Li X, Wu X, Fan X, Wen L, Tang F: **Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing**. *Nat Protoc* 2015, **10**(5):645-659.

## G43: Withdrawn

## G44: Application of RNA sequencing on BGISEQ-500 platform

Yang, T.T.[1], An, D.[1], Zhu, Z.Z.[1], Geng, C.Y.[1], Liang, X.M.[1], Chen, F.[1], Jiang, H.[1], Mu, F.[1]

[1]BGI-Shenzhen, Shenzhen 518083, China

Currently, high-throughput sequencing (HTS) are massively applied to research and clinical domain, especial the application of RNA next generation sequencing accelerates the development of RNA research[1].In this study, we applied the new combinatorial probe-anchor synthesis (cPAS)-based BGISEQ-500 sequencing platform that combines DNA nanoball (DNB) nanoarrays with stepwise sequencing using polymerase. An important feature of this technique compared to the current sequencing systems (as Illumina, SOLiD, et al.) is in that no PCR is applied in preparing sequencing arrays.

We have two kits (MGIEasy$^{TM}$ mRNA Library Prep Kit V2 and MGIEasy$^{TM}$ rRNA Depletion Kit V1) to construct a RNA library, which are based on two methods, mRNA enrichment (oligo dT beads) and rRNA depletion (RNase H digestion). We started a library with 1μg UHRR standard RNA sample following the one of kits and sequenced on the BGISEQ-500 with paired end (PE) 100bp. Raw data were filtered and normalized to 8G-10G base pairs per sample and then we calculated FPKM values based on those mapped bases. With the transcriptome PE 100bp sequencing, for instance, compared to HiSeq, the BGISEQ-500 platform had a slightly better performance on the qPCR Spearman correlation coefficient (0.884 to 0.866), expressed gene number (19413 to 19181) and ratio of known junction (0.63 to 0.616). Meanwhile, the BGISEQ-500 platform got a little lower genome mapping rate (0.881 to 0.884), gene mapping rate (0.61 to 0.68) and qPCR Pearson correlation coefficient (0.872 to 0.883). The overlapped gene number ratio of the two platform is 93.19%. With different species sequencing, two platforms also had similar performance. As for Saccharomyces sequenced on the BGISEQ-500 and HiSeq, the genome mapping rate (0.97 to 0.969), gene mapping rate (0.74 to 0.76), expressed gene number (5763 to 5770) and all junction number (1387 to1184) were equivalent. All in all, the RNA sequencing application on BGISEQ-500 performs equally to that on Illumina platform.

1. Xian Adiconis, Diego Borges-Rivera, Rahul Satija, *et al* (2013). Comparative analysis of RNA sequencing methods for degraded or low-input samples. Nature Methods. 10(7):623-29.

## G45: Mechanisms of Active DNA Demethylation in somatic cells

Issam Mayyas[1], Rob Weeks[1], Mark Hampton[2], Ian Morison[1].

[1]Department of Pathology, University of Otago, Dunedin School of Medicine, NZ, [2]Department of Pathology, University of Otago, Christchurch, NZ.

While, every cell in an organism is genetically identical there are marked phenotypic differences between tissues and organs that are controlled by epigenetic modifications. The most stable epigenetic modification is methylation of cytosine. Many cancers show significant global loss of methylation. Our research investigates the mechanism of changes in DNA methylation as they occur during human life using a *barcoded hairpin-bisulphite sequencing technique*. We have observed rapid demethylation in cultured Jurkat cells (T cell leukaemia) implicating novel mechanisms of active demethylation that have not yet been recognised by researchers in the field. It is likely that the demethylation pathways that we are studying operate during the onset of cancer and the existence of molecules (ascorbate or transition metals) that alter TET activity may have implications for modification of this process. While there is a substantial amount to be done before making therapeutic or dietary recommendations, our results might provide a rationale for long-term intervention to alter an individual's epigenetic risk.

## G46: Predicting altered methylation patterns in early pre-eclampsia

Suzan Momani[1], Ian M. Morison[1], Noelyn A. Hung[1], Tania L. Slatter[1], Erin C. Macaulay[1]
[1] Department of Pathology, Dunedin School of Medicine, University of Otago, 270 Great King Street, Dunedin 9054, New Zealand

Pre-eclampsia is one of the most common adverse pregnancy conditions that complicates 5-10% of pregnancies worldwide, yet the underlying cause is unknown. The use of biomarkers to accurately identify women with an increased risk of developing pre-eclampsia would be a major step forward in antenatal care. Our lab previously performed genome-wide methylation sequencing to identify DNA methylation differences between dysfunctional pre-eclamptic placentas and matched healthy controls. Validation studies are currently being performed in an independent cohort to determine which of the identified methylation changes are representative features of pre-eclampsia. The early detection of methylation changes in pre-eclampsia would provide a feasible route to early diagnosis and potential intervention. We are currently determining whether our panel of candidate biomarkers for pre-eclampsia can be identified in the circulating cell-free placental DNA that comprises approximately 10% of a pregnant women's blood plasma. We are performing deep sequencing using the MiSeq platform to examine the methylation of placental DNA in maternal plasma. We will measure the methylation of our candidate biomarkers in plasma samples from early gestation (15 and 20 weeks) and determine whether their differential methylation can be reliably profiled in maternal plasma from women who later developed pre-eclampsia. Ultimately, we seek to identify a DNA methylation signature of pre-eclampsia in maternal blood plasma that can be used clinically to predict women who are at risk of developing this threatening condition of pregnancy.

## G47: A journey of GMO regulation in New Zealand

Safavi M.

*Environmental Protection Authority, New Zealand*

Achieving a predator-free New Zealand by 2050 is likely to require genetic based approaches such as CRISPR/Cas9 based gene drives, the Trojan female technique, species-specific toxins, and non-genetic based approaches such as novel bio control agents. Many of these approaches are likely to be regulated by the Hazardous Substances and New Organisms Act 1996 (HSNO Act). 'New organisms' are those organisms that were not present in New Zealand immediately before 29 July 1998, including any genetically modified organisms (GMO). Under the HSNO Act, any applications to import, develop, field test, or release any new organisms are assessed by the Environmental Protection Authority (EPA). Each application is considered on a case-by-case basis, reflecting the associated potential positive or adverse effects of the new organism on the environment. In this talk, I will take you through the regulatory journey of some previously-approved GMOs, the matters that were taken into account for their risk assessment, and the requirements for field testing or releasing any new organisms. The necessity to obtain social licence for new technologies will also be discussed, as public notification is a statutory requirement for the applications that are likely to be of significant public interest.

## G48: Epigenetic driver of melanoma metastasis identified by genome-wide analysis

Eccles, M.R.[1], Chatterjee, A.[1], Rodger, E.J.[1], Ahn, A.[1], Leichter, A.[1], Motwani, J.[1], Stockwell, P.[2], Parry, M.[3]

[1]Department of Pathology, Dunedin, University of Otago, New Zealand. [2]Department of Biochemistry, Dunedin, University of Otago, New Zealand. [3]Department of Mathematics & Statistics, Dunedin, University of Otago, New Zealand.

As cancer metastasis is a major cause of cancer-related mortality, understanding the causes and effects of genomic changes in this process will help in identifying metastasis-related treatment targets. We have focused on investigating genomic changes in melanoma metastasis, because melanoma is one of the most aggressive and invasive of human cancers. While epigenetic alterations are hypothesized to play a facilitatory role in metastasis, for example in the promotion of epithelial to mesenchyme transition (EMT), which is increasingly supported by accumulating evidence, it is not clear whether epigenetic changes are necessary or sufficient to drive metastasis. We have addressed this question by using an approach involving mapping genome-wide DNA methylation changes in order to identify shared methylation changes across a series of primary and metastatic matched melanoma samples. Using this approach we identified 75 common (10 hyper- and 65 hypomethylated) genomic regions associated with 68 genes showing significant methylation differences between metastatic and primary tumours, which were shared amongst all three unrelated metastatic melanoma cell lines. One epigenetic change involved elevated Early B Cell Factor 3 (EBF3) mRNA levels and concomitant promoter hypermethylation in the three metastatic melanoma cell lines. RNAi-mediated knockdown of EBF3 in melanoma cell lines demonstrated an oncogenic role for EBF3 expression in promoting aggressive melanoma behaviour. We investigated collections of additional melanomas, and identified similar significant promoter hypermethylation and significantly elevated EBF3 mRNA levels in metastatic versus primary melanomas in two publicly available independent melanoma cohorts (n=40 and 458 melanomas, respectively). Moreover EBF3 has recently been identified as a top hit in GWAS screening of melanoma susceptibility loci. Including ours, several studies now suggest that epigenetic alterations have the potential to be oncogenic “drivers” of cancer aggressiveness, and are commonly recruited during the process of metastatic progression in tumour cells.