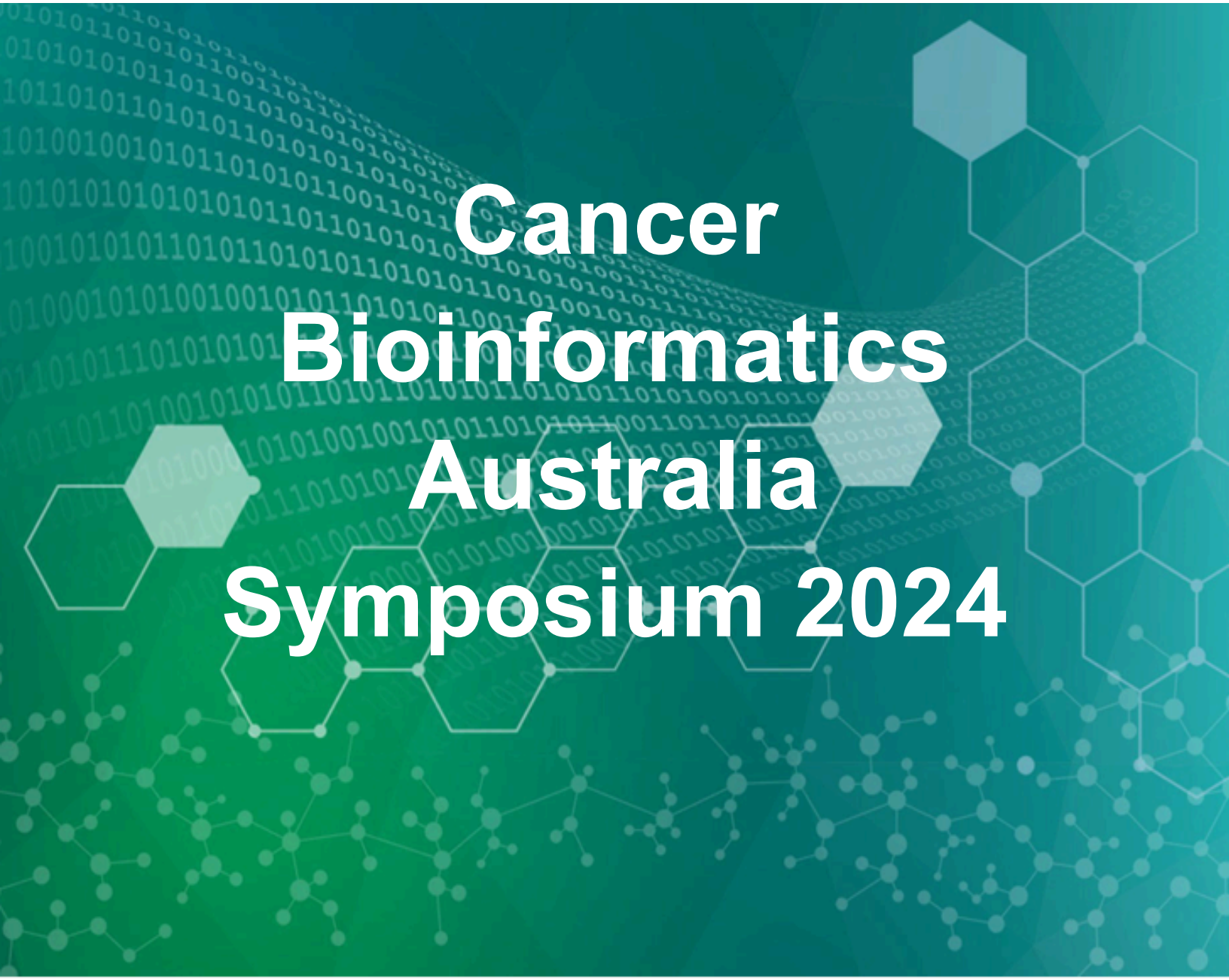


Cancer Bioinformatics Australia Symposium 2024

Peter MacCallum Cancer Centre
June 25th, 2024

Program Version 210624.1

Index

Cancer Bioinformatics Australia Symposium 2024

Acknowledgement of Country

The Peter MacCallum Cancer Centre acknowledges the Traditional Custodians of Country. We pay our respects to Elders past and present, who hold the healing traditions, lore, memories, and hopes of First Nations peoples. Peter Mac recognises the continuing connection to land, waters, and community, and the role these connections play in strengthening approaches to First Nations peoples' health and wellbeing. We honour the Wurundjeri Woi-Wurrung people, Dja Dja Wurrung people, and Boon Wurrung people of the Kulin Nations, on whose lands we meet and work each day, and acknowledge that sovereignty was never ceded.

Welcome by the organising committee

The wonderful and dedicated organising committee would like to welcome you to the Cancer Bioinformatics Australia symposium 2024.

Cancer Bioinformatics Australia is the new iteration of the Victorian Cancer Bioinformatics Symposium (VCBS) which was started in 2019. Our hope with rebranding the symposium this year is to include all people working in Cancer Bioinformatics across the country to build and foster the wonderful community we have. There is so much important work that needs to be done in this field and we can't make progress without bringing people together to share ideas.

We are delighted to host four outstanding invited speakers and have made room in our program for many people who submitted abstracts at a variety of career levels. We have had an outstanding response from the community and tickets for the event were sold out a month before the event. We wish to thank the support of our sponsors who have been eager to support the program. We are looking forward to a day of amazing science combined with connecting and networking with our peers.

Committee:

- Alicia Oshlack (Peter Mac)
- Andrew Lonsdale (Peter Mac)
- Chelsea Mayoh (Children's Cancer Institute)
- Claire Sun (Hudson Institute)
- David Goode (Peter Mac)
- Mark Cowley (Children's Cancer Institute)
- Michelle Meier (Peter Mac)
- Miriam Yeung (Peter Mac)
- Nenad Bartonicek (Peter Mac)
- Richard Lupat (Peter Mac)
- Saskia Freytag (WEHI)
- Sehrish Kanwal (University of Melbourne)
- Tu Nguyen-Dumont (Monash University)

We would like to thank our sponsors

GOLD

MONASH University

THE UNIVERSITY OF MELBOURNE

Peter Mac
Peter MacCallum Cancer Centre
Victoria Australia

SILVER

illumina®

BRONZE

agrf

MP CCC Monash Partners Comprehensive Cancer Consortium

Novogene

PATRON

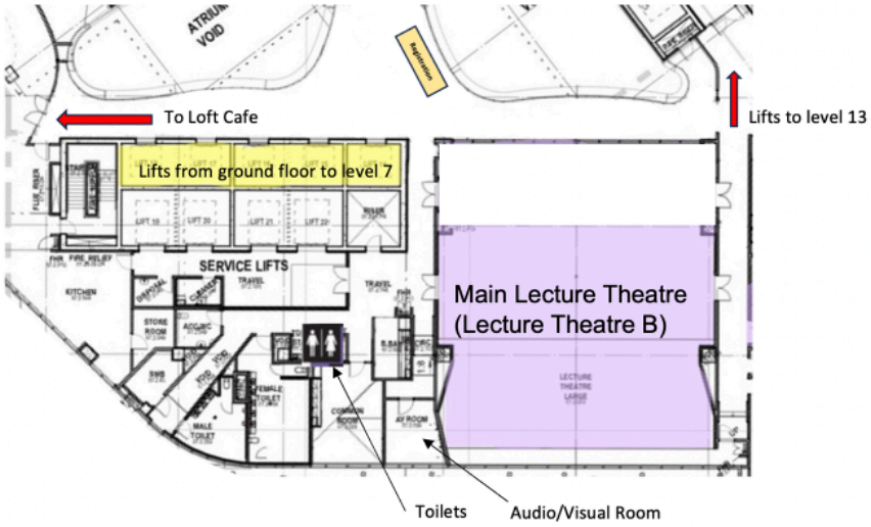
AUSTRALIAN BIOINFORMATICS AND COMPUTATIONAL BIOLOGY SOCIETY

General Information

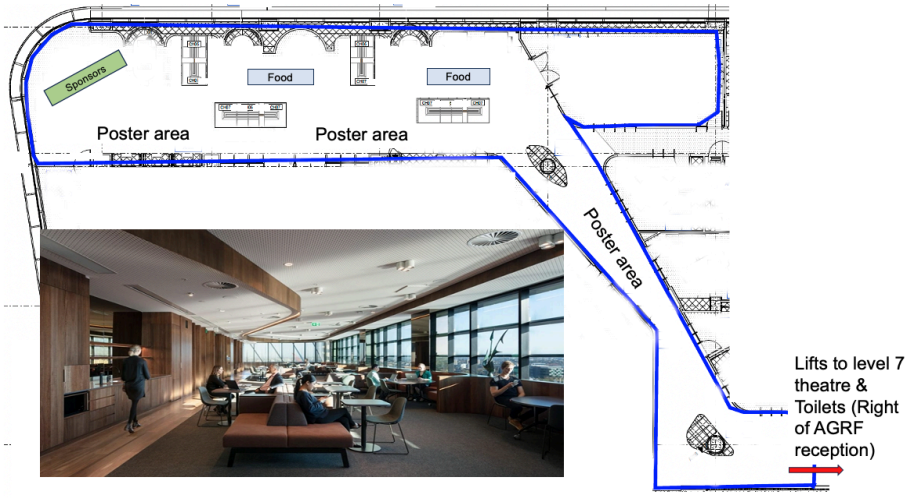
Event Details
8.45am-5.30pm
Peter MacCallum Cancer Centre, 305 Grattan St, Melbourne VIC 3052

Arrival & Registration
Upon arrival at Peter Mac, use the general lifts on the ground floor to proceed to the registration desk on Level 7. Registration is open from 8am.

The main program takes place in Lecture Theatre B.



The poster session, sponsor exhibits, as well as food and drink breaks will all take place on Level 13, accessible via research (purple) lifts.



Code of Conduct

All registrants to CBA have declared that they will adhere to the code of conduct; https://cancerbioinformatics.au/code-of-conduct/

Anyone who witnesses or is subjected to unacceptable behaviour should notify one of the designated safety officers. Current safety officers are made up of the CBA committee (green name badges, annotated “Organising Committee”). Reports can be made in person to a safety officer, emailed directly to a safety officer, or emailed to cancerbioinformaticsau@gmail.com.

Toilets

Toilets are located on Level 7 adjacent to Lecture Theatre B, and on Level 13 behind the lifts.

Wifi

Networks available:

- Eduroam
- Petermac_Free_Wifi (accept terms and conditions).

Online Streaming

Scientific talks will be streamed via a Zoom webinar. Completing this form will prior to the event to ensure timely processing, as the Zoom details will be sent via email once webinar registration is complete.

To accommodate the webinar participants:

- Quickly introduce yourself when asking a question
- Wait for the microphone so online attendees can hear you

Getting to Peter Mac

Due to Metro Tunnel works, accessibility by car and parking in the area around Peter Mac is limited. Three tram lines travel near Peter Mac. These are tram line 19 (stop number 10) and the 58 and 59 tram lines (stop 14). See Peter Mac resources for more information: https://www.petermac.org/about-us/locations/melbourne-campus

Prizes

Speakers and poster presenters have self-nominated for prize categories for students and early career researchers. Prizes will be awarded subject to sufficient eligible participants in long and short talks, and in posters.

Two ‘People’s Choice’ Poster Prizes will be awarded. Voting for the People’s Choice will close after the conclusion of the afternoon tea break.

Use this link to vote for the People’s Choice awards. Votes are anonymous but limited to one per person (via Google account login).

BYO Lanyards

In an effort to reduce waste, we request attendees to please bring their own lanyards on the day. For example these could be from a previous conference or event. There will be a limited number of generic lanyards available at the registration desk for those who cannot source their own.

Masks & Illness

We would like to remind everyone to be vigilant in preventing the spread of disease. If you are feeling unwell or test positive to COVID-19 or another virus then we request you please stay home.

In addition to protecting colleagues, Peter Mac is a hospital with many immunocompromised patients. Current Peter Mac guidelines recommend practising excellent hand hygiene and correct usage of PPE.

Masks are not mandatory in the areas of the building where the symposium is being held, but will be available at the registration desk if desired.

Posters and Lightning Talks

Posters are numbered and indicated in the program. Dots for attaching posters can be collected from the registration desk.

Lightning talks for selected posters will be presented at the end of either Session1 or Session 2, as indicated in the running order in, located on Page 26. If you have a lightning talk please sit near the front during your session.

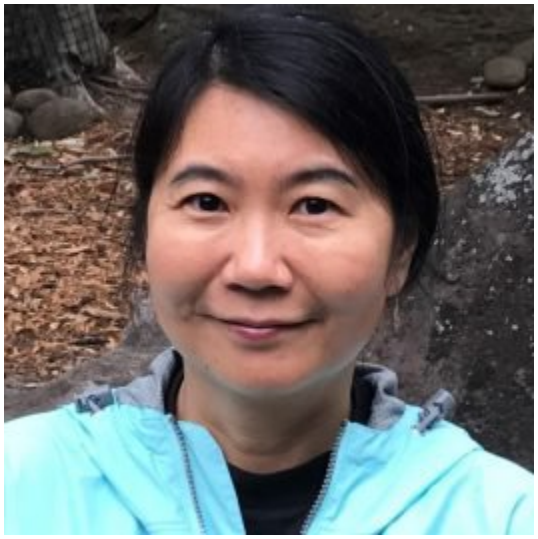
The poster session will run during lunchtime on Level 13. Posters can be taken down at the conclusion of afternoon tea, or after the conclusion of CBA.

2024 Cancer Bioinformatics Australia Schedule	
08:00	Registration Opens (Level 7, Peter MacCallum Cancer Centre)
8:45-9:00am	Welcome & Acknowledgement of Country (Tu Nguyen-Dumont)
Session One - Clinical Translation (Chair: Mark Cowley)	
9:00-9:45am	**International Speaker: Jinghui Zhang (St. Jude Children's Research Hospital, USA) - Sponsored by Monash University**
9:45-10:00am	Peter Priestley "WiGiTS: A universal NGS suite of tools for cancer research and diagnostics"
10:00-10:15am	Jie Mao "Data Analysis Pipeline for ex vivo High-Throughput Drug Screening: Enhancing Clinical Decision-Making and Drug Efficacy Analysis"
10:15-10:30am	Lightning Talks (Chair: Sehrish Kanwal)
10:30-11:00am	MORNING TEA - Level 13
Session Two - Single-cell & Spatial (Chair: Saskia Freytag)	
11:00-11:30am	**National Speaker: Alistair Forrest (Harry Perkins Institute of Medical Research) - Sponsored by Illumina**
11:30-11:35am	Joel Moffet "Multi-Omic Spatial Reconstruction of Glioma in 3 Dimensions"
11:35-11:50am	Sirui (Cathy) Weng "Archetype analysis to identify tumour cell states and transitions across related samples from the same patient from single-nuclei RNAseq data"
11:50-11:55am	Tongtong Wang "Unravelling pre-implantation transcriptomic signatures in cancer, friend or foe?"
11:55-12:10pm	Michael Roach "Advancements in segmentation-free typing with Xenium and CosMX"
12:10-12:15pm	Feng Wang "Conformational and Spectroscopic Analysis of 4-Anilinoquinazoline TKIs"
12:15-12:30pm	Lightning Talks (Chair: Andrew Lonsdale)
12:30-1:45pm	LUNCH & POSTERS - Level 13

Session Three - Multi-omics (Chair: David Goode)	
1:45-2:30pm	**International Speaker: Anders Skanderup (Genome Institute of Singapore, Singapore) - Sponsored by Peter MacCallum Cancer Centre**
2:30-2:45pm	Peinan Zhao "Molecular characterization of melanocyte subpopulations in human epidermis based on single-cell RNA sequencing"
2:45-3:00pm	Sanjay Krishna "CANELIB: a comprehensive platform for LC-MS-based cancer neoantigen discovery using whole exome sequencing data"
3:00-3:15pm	Lachlan Cain "ATAC of the Clones: Genetic subclone identification and copy number estimation in cancer from single-cell ATAC-seq"
3:15-3:45pm	AFTERNOON TEA - Level 13
Session Four - Methods (Chair: Claire Sun)	
3:45-4:15pm	**National Speaker: Katherine Pillman (The University of South Australia) - Sponsored by University of Melbourne**
4:15-4:30pm	Sung-Young Shin "Leveraging Machine Learning and Patient-Derived Explants to Discover Predictive Biomarkers of CDK4/6 Inhibitor Efficacy in Prostate Cancer Patients"
4:30-4:35pm	Charla Lu Chai "Understanding Mechanisms That Drive Disease Relapse in Early Stage, Low-risk Colorectal Cancer"
4:35-4:50pm	Liyang Fei "BARBIE: An R package for analysing barcode count data from clonal tracking experiments"
4:50-4:55pm	Pablo Acera Mateos "Cryptico: A Computational Approach for Exploring Cryptic Exons in the Search for Novel Therapeutics"
4:55-5:10pm	Ashley Weir "IdentifiHR: predicting homologous recombination deficiency in high-grade serous ovarian carcinoma through gene expression."
5:10-5:15pm	Nathan Hall "Predicting NSCLC patient response to immunotherapy using IIF profiling"
5:15-5:30pm	Closing Remarks & Prizes (Alicia Oshlack)
6:00pm-Late	DRINKS & NETWORKING @ The Castle

Invited Talks

Jinghui Zhang - St. Jude Children's Research Hospital



Professor Zhang is a computational biologist focused on the integrative analysis of large-scale, multi-dimensional genomic data to understand the initiation and progression of diseases. Her research interest has been in the development of highly accurate and sensitive computational methods for analyzing large-scale genomic data, especially in the area of detecting and analyzing genetic variations and somatic mutations. She participated in the development of BLAST and led the genetic variation analysis of the first assembled human genome. Her group has developed computational tools for analyzing genetic variations and somatic mutations.Understanding the genetic origin of pediatric cancers is currently the main focus of her research group which is primarily responsible for analyzing the next-generation sequencing data generated from the St. Jude Children's Research Hospital - Washington University Pediatric Cancer Genome Project (PCGP). Recently her team developed a suite of algorithms, CREST, CONSERTING and CICERO, for discovery of structural variations and copy number alterations in cancer genome at high accuracy. She has developed mathematical models for mapping clonal evolution trajectory from diagnosis to relapse for pediatric solid tumor and leukemia.

Mathematical modeling of therapy-related clonal evolution and clonal hematopoiesis in pediatric cancer patients and long-term survivors

Understanding the short-term and long-term therapy-related effect on the genomes of pediatric cancer and survivors is essential for reducing the mortality associated with cancer relapse and the accelerated physiological aging of long-term survivors. We present mathematical modeling of therapy-related mutagenesis processes revealing spatial pattern of tumor cell infiltrating in high grade gliomas and timing of emerging clones that confer therapy-resistance under the selective pressure of exposure to cytotoxic agents. We also show the dynamics of age- versus therapy-related clonal hematopoiesis (CH) in long-term survivors of pediatric cancer with a median follow-up time of 23.5 years. CH in survivors is associated with exposures to alkylating agents, radiation, and bleomycin. Therapy-related CH shows significant enrichment in STAT3, characterized as a CH-gene specific to Hodgkin lymphoma survivors, and TP53. Single-cell profiling of peripheral blood samples revealed STAT3 mutations predominantly present in T-cells and contributed by SBS25, a mutational signature associated with procarbazine exposure. Serial-sample tracking reveals that larger clone size is a predictor for future expansion of age-related CH clones, while therapy-related CH remains stable decades post-treatment. These data depict the distinct dynamics of these CH subtypes and support the need for longitudinal monitoring to determine the potential contribution to late effects.

MONASH University

Anders Skanderup - Genome Institute of Singapore

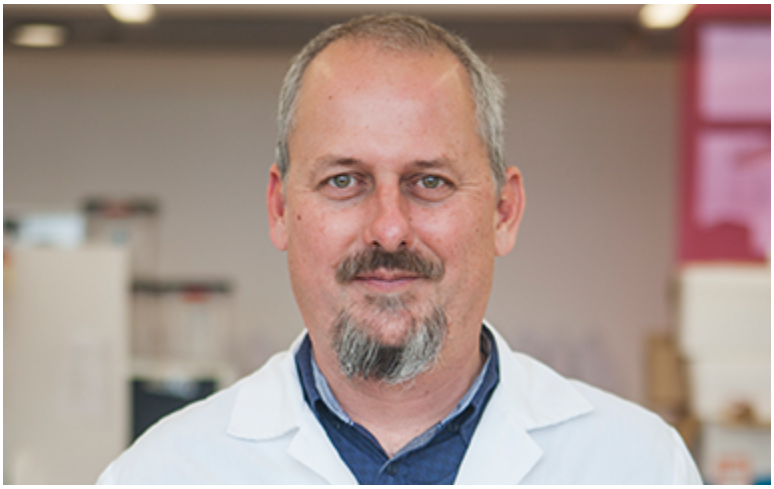
Anders Skanderup is a Group Leader at the Genome Institute of Singapore. His research focuses on computational and data-intensive approaches to deciphering the molecular basis of cancer and improving treatments. Recent contributions include computational methods for identifying cancer mutations and non-coding driver mutations, as well as deconvolution approaches for tumor transcriptomes and cell-free DNA. He is also a principal investigator in large national programs for lung cancer and cancer liquid biopsies. Dr. Skanderup has published 90 research papers (h-index 53). He holds adjunct positions in the Department of Computer Science at the National University of Singapore, the Cancer Science Institute of Singapore, and the National Cancer Center Singapore. He obtained a B.S. in Computer Science and a Ph.D. in Bioinformatics from the University of Copenhagen, Denmark. Subsequently, he completed postdoctoral training at Memorial Sloan-Kettering Cancer Center in New York, where he developed and applied computational methods to study cancer using high-dimensional cancer genomic datasets.

Turning noise into insights: computational approaches and applications of tumor transcriptome deconvolution

Tumors comprise a multifaceted cellular environment that both constrains and supports cancer evolution and malignant progression. Molecular profiling of tumors is commonly done in ‘bulk’, mixing signals from individual cell types. This creates noise that obscures the transcriptome profiles of individual cell types in the tumor. Unfortunately, experimental methods to characterize individual cells in the tumor microenvironment, such as imaging and single-cell–based approaches, cannot be applied retrospectively to existing large-scale bulk tumor datasets. Here, we discuss how a simple computational framework can take advantage of the noise in bulk tumor transcriptome profiles to learn the underlying gene expression profiles of individual cell types. We present recent methods developed by our group in this domain and highlight their applications in revealing novel insights into tumor metabolism, immune checkpoint blockade drug efficacy, and therapeutic resistance.

Peter Mac
Peter MacCallum Cancer Centre
Victoria Australia

Alistair Forrest - Harry Perkins Institute of Medical Research



Professor Forrest BSc (Hons), MIT, PhD was born in Western Australia and obtained his BSc (Hons) in Biotechnology at Murdoch University in 1993. He then moved to Brisbane and while working as a research assistant at the QIMR completed a Masters in Information Technology at the Queensland University of Technology. Shortly afterward, he completed his PhD in Bioinformatics at the Institute for Molecular Bioscience at the University of Queensland. During his time in Brisbane he was involved in both 'wet' and 'dry' science, generating and analysing some of the first microarrays used in Australia and invented a strand specific RNA-seq protocol that heralded the start of the RNA-seq revolution. In 2007 he moved to RIKEN Yokohama Japan on a CJ Martin Fellowship. Over the past seven years he has progressively been promoted at RIKEN taking on more senior roles, and is currently scientific coordinator of the FANTOM5 (Functional Annotation of the mammalian genome) project consisting of a consortium of over 250 scientists in 20 countries. This has used single-molecule sequencing to generate a map of promoters and enhancers across a large collection of human and mouse primary cells, cancer cell lines and tissues. The work has recently been published in the prestigious journal Nature along with a collection of 18 additional satellite papers in specialised journals such as Nature Biotechnology, Genome Research and Blood.

Spatial transcriptomics reveals discrete tumour microenvironments and autocrine loops within ovarian cancer subclones

High-grade serous ovarian carcinoma (HGSOC) is a polyclonal disease characterised by the presence of subclones with distinct cancer genotypes. This intratumoural heterogeneity is linked to recurrence, chemotherapy resistance, and overall poor prognosis. Here, we used spatial transcriptomics platforms (10x Genomics Visium and NanoString CosMx Spatial Molecular Imaging (SMI)) to examine genetic heterogeneity of HGSOC cells and their association with infiltrating populations in samples from patients treated with neoadjuvant chemotherapy. We found evidence of multiple tumour subclones with different copy number alterations co-existing within individual tumour sections. Examining gene expression differences between subclones we found evidence that their cell-to-cell communication networks may be rewired by differences in ligand and receptor expression levels. We hypothesise that this may modulate their interactions with stromal and immune cells and likely also leads to the creation of subclone specific autocrine loops.

Katherine Pillman - The University of South Australia

Dr Katherine Pillman is a bioinformatician and RAH Florey Fellow at the Centre for Cancer Biology in South Australia. She uses a variety of transcriptomic methods and next-gen sequencing data types to dissect gene regulation through analysis of microRNA biology and targeting, circular RNAs, alternative splicing, epigenetic modifications, gene regulatory networks and expression. She obtained her PhD in Molecular Biology from the University of Adelaide in 2009 working on transcriptional regulation in barley plants. Her postdoctoral research at Oregon State University involved RNA-seq analysis of stress-responsive gene regulatory networks in potato plants. In 2012, she returned to Australia to take up her current role as Lead Bioinformatician in the Gene Regulatory Section at the Centre for Cancer Biology. Career highlights include a Cell paper which identified the first protein known to control the formation of circular RNAs, and the subsequent discovery that this protein, Quaking, is a key emerging player in splicing in cancer. Other work has identified genome-wide microRNA regulatory networks controlling cell invasion and uncovered the cooperative role of microRNAs in cancer.

Deciphering Cancer Complexity: Insights from Transcriptional and Post-transcriptional Regulatory Networks

Gene regulatory networks orchestrate fundamental cellular processes such as proliferation, differentiation, and apoptosis. In cancer, aberrations within these regulatory networks often drive tumourigenesis and disease progression. Over the past decade, our research has focused on the roles of diverse regulators such as transcription factors, splicing factors, circRNAs and microRNAs in driving oncogenesis. Tumour heterogeneity, both within individual tumours and across different tumours, poses a significant challenge in cancer research. Leveraging single-cell RNA-seq and ATAC-seq technologies, we have used the next-generation Gene Regulatory Network methods like Cell Oracle to modelling transcriptional regulatory networks in neuroblastoma at a single-cell resolution. Through in silico perturbations, we have explored regulatory microcontexts to identify transcription factors with strong potential as oncogenic drivers, currently undergoing wet-lab validation. Discovering and characterising these factors and their regulons marks the initial phase of dissecting and explaining transcriptome heterogeneity, thereby enabling prediction of which tumours will respond to targeted treatments and which are likely to relapse. Systems biology has repeatedly shown us that, though technically challenging to study, post-transcriptional regulatory factors such as microRNAs, circRNAs and splicing factors play crucial roles in cancer biology by acting as central nodes in interconnected regulatory networks. As a hybrid wet/dry lab group, we aimed to work at the cutting edge of these fields over the past decade, deep-diving into what sequencing data can teach us about these layers of regulation and developing analyses as required. We have discovered key splicing regulators and export factors that control the formation and sub-cellular localisation of circular RNAs in cancer and likewise, microRNAs and splicing factors that play vital cooperative roles in influencing cell invasion and cancer progression. Our findings underscore the importance of modelling networks when seeking to understand cancer biology and offer promising avenues for future systems approaches to personalised medicine.

THE UNIVERSITY OF MELBOURNE

Session One - Clinical Translation (Chair: Mark Cowley)

Peter Priestley - Hartwig Medical Foundation

WiGiTS: A universal NGS suite of tools for cancer research and diagnostics

Peter Priestley [1], Charles Shale [1], Hong Wing Lee [1], Stephen Watts [2], Luan Nguyen [1], Thomas Garrety [1], Junran Cao [1], Matthew Cooper [1], Oliver Hoffman [2], Edwin Cuppen [3]
[1] Hartwig Medical Foundation Australia
[2] University of Melbourne Centre for Cancer Research, Victoria Australia
[3] Hartwig Medical Foundation, Netherlands

Many tools and analyses are available to identify and interpret cancer relevant variants and biomarkers based on next generation sequencing. However, creating a comprehensive bioinformatics pipeline for cancer diagnostics and/or large scale research projects remains a complex and time consuming task. Key challenges include adapting variant callers to the specifics of the experimental setup and integration of multiple tools and analyses which may require extensive optimisation, annotation and filtering. Furthermore, upgrades to a pipeline and the accompanying sample validation is often equally challenging.

To address this problem, we have developed WiGiTS, an open source, comprehensive suite of tools for analysis of both whole genome and transcriptome sequencing (WGTS) and large capture panel data. The functionality of WiGiTS includes basic alignment and variant calling, but also advanced features such as tissue of origin analysis, homologous recombination deficiency classification, telomere analysis, as well as immune related analyses including HLA typing, CDR3 identification and neoepitope predictions. Clinically relevant findings are summarised in an open source molecular report. WiGiTS is compatible with multiple sequencing technologies including Illumina and Ultima and is suitable for analysis of both FFPE and fresh frozen material.

The WiGiTS suite has been implemented as a Nextflow pipeline called ‘oncoanalyser’ which is available in nf-core and can be run on any cloud or HPC cluster. The pipeline leverages resource files based on a database of more than 6,000 cancer genomes. WiGiTS can also be adapted to run on any exome or targeted NGS data by generating panel specific resources via a simple training procedure based off a representative set of samples. Pre-trained resources are available for the TSO500 panel. WiGiTS is inexpensive to run and provides a highly practical option for either a diagnostic pipeline or for uniform reanalysis of large cohorts of both WGTS and large panel data.

Jie Mao - Children’s Cancer Institute

Data Analysis Pipeline for ex vivo High-Throughput Drug Screening: Enhancing Clinical Decision-Making and Drug Efficacy Analysis

Jie Mao[1], Gabor Tax[1,2], Roxanne Cadiz[1], Loretta MS Lau[1,2,3], Paul G Ekert[1,2,4], M Emmy M Dolman[1,2], Chelsea Mayoh[1,2]
[1]Children's Cancer Institute, Lowy Cancer Centre, UNSW Sydney, Kensington, NSW, Australia
[2]School of Clinical Medicine, UNSW Medicine & Health, UNSW Sydney, Kensington, NSW, Australia
[3]Kids Cancer Centre, Sydney Children’s Hospital, Randwick, NSW, Australia

[4]Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia

To identify additional treatments beyond genomic-driven therapy recommendations in high-risk paediatric cancer patients, the ZERO Childhood Cancer Precision Medicine Program integrates ex vivo high-throughput drug screening (HTS) on patient-derived samples. An automated pipeline is essential for promptly analysing and reporting therapy recommendations while patients are still under clinical care.

Here we developed a pipeline for single agent and combination HTS data analysis. The automated pipeline involves evaluating drug screen quality, fitting dose-response curves, calculating drug efficacy parameters and synergy scores, and identifying potential efficacious drugs. All drug response data are presented through an in-house web interface that is integrated with matched patient molecular profiling data, providing a platform for curation of drug recommendations.
This pipeline was implemented on ZERO patients enrolled from February 2022, and has currently processed 72 tumor samples, sixty of which had a potentially efficacious drug recommendable to treating clinician. The automated data processing has eliminated errors that would arise from manual handling, and outputs results in under 30 minutes per sample. To enhance clinical utility, our reporting extends beyond traditional metrics such as area under the dose response curve (AUC) and half inhibitory concentration (IC50). We incorporated innovative drug efficacy parameters including half lethal concentration (LC50) and response at peak serum concentration, providing insights into cell death and treatment efficacy within a clinically actionable window. Importantly, our pipeline has unearthed unexpected findings, such as the synergistic effects of panRAF and MEK inhibitors in tumors with NF1 loss, leading to the initiation of new clinical trials.

Our automated HTS data analysis pipeline is a powerful tool for expanding precision medicine to support drug treatment recommendations and broadens the therapeutic options available to high-risk pediatric cancers.

Session Two - Single-cell & Spatial (Chair: Saskia Freytag)

Joel Moffet - The Walter and Eliza Hall Institute of Medical Research

Multi-Omic Spatial Reconstruction of Glioma in 3 Dimensions

Joel J.D. Moffet [1,2] , Jurgen Kriel [1,2], Oluwaseun E. Fatunla [1,2], Tianyao Lu [1,2], Jordan J. Jones [3], Samuel J. Roberts-Thomson [4], Andrew P. Morokoff [3], Jim R. Whittle [1,2,5], Saskia Freytag [1,2], Sarah A. Best [1,2]
[1] Personalised Oncology Division, The Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia
[2] Department of Medical Biology, University of Melbourne, Melbourne, Australia
[3] Department of Surgery, Royal Melbourne Hospital, Melbourne, Australia
[4] Department of Anatomical Pathology, Royal Melbourne Hospital, Melbourne, Australia
[5] Department of Medical Oncology, Peter MacCallum Cancer Centre, Melbourne, Australia

Gliomas are cancers of the brain that present a five-year survival rate for patients of 22%, with the current standard of care unable to prevent disease progression. The leading edge of the tumour is thought to be a key architecture involved in both recurrence and progression, but spatially resolved characterisation of

this region is lacking. Increasing the dimensionality at which we interrogate glioma into space and across omics platforms will improve our understanding of the underlying mechanisms causing biological dysfunction, and elucidate the interactions occurring between immune cells, tumour states, and the neuronal network at the diffuse leading edge. We are developing pipelines to process data from spatial transcriptomics and metabolomics to model the leading edge across serial sections of lower grade glioma. Interpolation between sections via Kriging and Machine Learning permits the integration of multi-omic technologies into a 3D model of disease architecture. Validated with mouse brain MERFISH data and performed on human gliomas with Xenium and CosMx technologies, Kriging displays promising accuracy (60-70%) in predicting the spatial changes to cellular neighbourhoods across serial sections spaced over 200 microns apart, offering a cost-effective approach to reconstructing tissues in 3D. There is significant need to improve our understanding of the leading edge in glioma, which will be enhanced by innovative spatial multi-omic strategies. For example, spatial transcriptomics has identified transient increases of expression for Myelin-associated genes such as MBP across the tumour leading edge, which vary in intensity across serial sections, highlighting the importance of a 3D model. Furthermore, abundance of the metabolite spermidine in the tumour core can be aligned to transcriptomic data to reveal association with and potential regulation of infiltrating T cells. Investigating glioma in its natural 3D and multi-modal context will improve our ability to identify spatial drivers of tumorigenesis and develop effective treatments for patients.

Sirui (Cathy) Weng - The Walter and Eliza Hall Institute of Medical Research

Archetype analysis to identify tumour cell states and transitions across related samples from the same patient from single-nuclei RNAseq data

Sirui (Cathy) Weng [1,2], Anna Trigos[1, 2, 3]
[1] Multi-omic Evolution and Ecosystems Laboratory, Laboratory Research, Peter MacCallum Cancer Centre, Melbourne
[2] Sir Peter MacCallum Department of Oncology, The University of Melbourne, Melbourne
[3] Monash University, Melbourne

Single-cell transcriptomics allows the discovery of novel cell populations that reveal intra-tumour heterogeneity. Having multiple tumour samples from a single patient provides a unique opportunity to validate these populations and investigate their recurrence. However, there are several challenges in comparing cell populations across related patient samples. Data integration techniques can theoretically help identify common or unique cell states between tumour samples, but currently available integration methods overcorrect, resulting in removal of real biology. Inter- and intra-lesion genetic subclones confound clustering and integration algorithms. Finally, traditional clustering results in a discrete classification of cells, which is not suitable to identify cells in transition or intermediate states between multiple cell fate trajectories.

To address these issues, we applied archetype analysis in a cohort of single-nuclei RNAseq data from 9 metastatic prostate cancer patients, with 2-6 metastatic lesions per patient. Archetype analysis dissects tumours into distinct cell subpopulations performing various survival critical tasks by fitting cells into convex polytopes. Archetype analysis dissects tumour cells into distinct subpopulations that express specialized gene modules by finding cells near the vertex of convex polytopes fitted using an unsupervised approach in a multidimensional space. This resulted in the identification of 4 archetypes per sample, which are the most distinct populations of cells in each sample. Next, consensus signatures for each archetype were derived from overlapping differentially expressed genes between archetype populations across samples. We identified 6 recurrent tumour cell states across our cohort, including

those related to adenocarcinoma and neuroendocrine pathologies, stress, hypoxia, immune signalling and cell cycle. We found that recurrent cell states are found across tumours of different patients, but that each patient has its own signature of transition states that can be found across metastatic lesions, which are independent of genetic clonal composition and patterns of metastatic spread.

Our archetype analysis enabled the identification of cell state signatures that dissect tumour cells into distinct populations that reflect cell state dynamics and regulatory mechanisms underlying cell plasticity. Our results provide previously underappreciated information about tumour plasticity and evolution in metastatic disease.

Tongtong Wang - Peter MacCallum Department of Oncology

Unravelling pre-implantation transcriptomic signatures in cancer, friend or foe?

Tongtong Wang [1,2], Janith A Seneviratne [1,2], David L Goode [1,2], Alicia Oshlack [2,3], Melanie A Eckersley-Maslin [1,2,4]
[1] Peter MacCallum Cancer Centre
[2] Sir Peter MacCallum Department of Oncology, The University of Melbourne
[3] School of Mathematics & Statistics
[4] Department of Anatomy and Physiology, University of Melbourne, Melbourne, Victoria, Australia

Background: Cellular plasticity, the capacity of cells to transition between states, is pivotal in cancer evolution. This fundamental process is observed not only in cancer but also during the development, repair, and maintenance of equilibrium. This study draws inspiration from the exceptional cellular plasticity exhibited during pre-implantation (PIE) embryo development, marking the pinnacle of plasticity during totipotency establishment. In early developmental stages, transcriptional control depends on maternal factors, but at the four-to-eight-cell stage, the zygotic genome starts actively transcribing. To detect genome reactivation akin to preimplantation states, we present PIECanceR, a machine-learning model to identify plasticity reactivation using transcriptomic data.

Aim: To explore the significance of transcriptional signatures associated with early embryonic plasticity in cancer.

Methods: A single-cell RNA sequencing atlas of early human embryos was assembled from pre-existing datasets. We selected genes with the highest variability in pan-cancer datasets and delineated their expression pattern during early embryonic stages as training features. A k-Top Scoring Pairs (kTSP) model was trained to classify cellular states along the developmental trajectory. Validation of the model involved simulating pseudobulk samples infused with embryonic signals against a background of 30 tissue types, with an increasing proportion of embryonic signals. Multiple datasets, including Genomic Data Commons (GDC), Genotype-Tissue Expression (GTEx), and the Cancer Cell Line Encyclopedia (CCLE), were then classified using PIECanceR to identify relevant patient samples, tissues, and cell lines enriched for the plastic signature.

Results: We have built a machine-learning model, PIECanceR, that successfully classifies early embryonic stages and demonstrated its capacity to score bulk RNA-seq samples in simulated samples. When applied to cancer samples, PIECanceR predicted elevated plasticity in a subset of patients, with significant differences in overall survival benefits.

Conclusions: PIECanceR translates developmental plasticity into cancer plasticity. This tool enables the discovery of reactivation of preimplantation plasticity from transcriptomic data in oncogenic settings. Investigation into the mechanisms and significance of ZGA in the context of cancer cells is ongoing.

Michael Roach - University of Adelaide

Advancements in segmentation-free typing with Xenium and CosMX

Michael Roach [1,2,3], David Cook [4], Kellie Wise [1,2], Monika Mohenska [1,2], Flipe Segato Dezem [5], Luke Zhang [5], Maycon Marção [5], Yutian Liu [5], Jiwoon Park [6], Christopher Mason [6], Jasmine Plummer [5], Luciano Martelotto [1,2]
[1] Adelaide Centre for Epigenetics, University of Adelaide, South Australia, Australia
[2] South Australian immunoGENomics Cancer Institute, University of Adelaide, South Australia, Australia
[3] Flinders University, South Australia, Australia
[4] Cancer Therapeutics Program, Ottawa Hospital Research Institute, Ottawa, Canada
[5] Center for Spatial OMICs, St Jude Children's Research Hospital, Memphis, TN, USA
[6] Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA

Single cell transcriptomics has revolutionized our understanding of the pathways driving differentiation and evolution of cell types. The ability to capture transcripts and associate them with individual cells provides the resolution required to identify minor cell types, cell states, and transitions. The maturation of spatial transcriptomics has further revolutionized the field by capturing the spatial information of transcripts and cell features. Two platforms are proving to be popular options in this space: 10X Genomicsâ€™ Xenium, and NanoString's CosMX.

Both Xenium and CosMX rely on cell staining and imaging to identify cell features which are then used to draw cell boundaries (cell segmentation). However, some tissue types are more difficult to segment or may exhibit significant cell heterogeneity, leading to incorrect cell boundaries or missing cells.

Segmentation-free typing can alleviate these limitations by grouping neighboring transcripts according to a grid and can annotate a tissue region based on the transcripts present in that location. We demonstrate the utility of segmentation-free typing for filling annotation gaps in both prostate and breast cancer clinical samples on the Xenium and CosMX platforms. We evaluate both segmentation-free and cell segmentation-based annotations with gold-standard single cell RNA-seq. While algorithms for spatial transcriptomics analyses are nascent, the developments so far are promising.

Feng Wang - Swinburne University of Technology

Conformational and Spectroscopic Analysis of 4-Anilinoquinazoline TKIs

Feng Wang [1]
[1]Swinburne

In the field of drug discovery and research, acquiring quantitative structure-activity relationship (QSAR) data, crucial for medical and chemical purposes, remains a daunting challenge, especially within automated computational processes. A significant obstacle in this regard is the need to achieve a comprehensive representation of three-dimensional (3D) conformations for flexible molecules. This study addresses this challenge by employing a newly developed conformational sampling method to analyze a class of potent 4-anilinoquinazoline drugs targeting the epidermal growth factor receptor (EGFR).

Specifically, we used a robust quantum mechanical conformational sampling method to study a class of potent tyrosine kinase inhibitor (TKIs), i.e., 4-anilinoquinazoline derivatives to study their low-energy conformations through intelligent computing coupled with quantum mechanical DFT calculations. The properties of these conformers were assessed using time-dependent density functional theory (TD-DFT) methods for their UV-Vis spectra for optical reporting. Subsequently, we compared these properties of the TKI derivatives to examine substituent impacts on the conformation and drug potency of the inhibitors for development of new such class of TKIs.

Session Three - Multi-omics (Chair: David Goode)

Peinan Zhao - Alfred Hospital, Monash University

Molecular characterization of melanocyte subpopulations in human epidermis based on single-cell RNA sequencing

Peinan Zhao* [1], Fumihito Noguchi* [1], Christopher Chew [1], Gamze Kuser Abali [1], Pacman Szeto [1], Youfang Zhang [1], William Berry [1], Magnus Zethoven [2], Malaka Ameratunga [1], Isobel Leece [1], Jen G. Cheung [1], Miles Andrews [1], Geza Paukovics [3], Nicholas C. Wong [4], Anthony T. Papenfuss [5,6], Mark Shackleton [1]
[1] Cancer Development and Treatment Group, Department of Medicine Research Laboratories, Alfred Hospital, Monash University, Victoria, Australia
[2] Bioinformatics Core Facility, Peter MacCallum Cancer Centre, Victoria, Australia
[3] Flow Cytometry Core Facility, Alfred Research Alliance, Victoria, Australia
[4] Monash Bioinformatics Platform, School of Translational Medicine, Monash University, Victoria, Australia
[5] The Walter and Eliza Hall Institute, Melbourne, Victoria, Australia
[6] Department of Medical Biology, University of Melbourne, Victoria, Australia

The biological and molecular mechanisms that underpin the malignant transformation of normal melanocytes to melanomas are largely unknown. In part, this is due to the limited understanding of normal human melanocyte homeostasis and how melanocytes respond to oncogenic insults such as ultraviolet radiation (UVR). This is particularly true for interfollicular epidermal melanocytes, which have the highest levels of UVR exposure and from whence most melanomas are thought to arise. These knowledge gaps impede the development of strategies for active, targeted prevention of melanoma formation.

We thus evaluated epidermal melanocytes transcriptionally, phenotypically and functionally after isolating them from human skin. Using single-cell RNA sequencing (scRNA-seq), we identified multiple transcriptionally distinct subpopulations within human epidermal melanocytes. RNA velocity analysis revealed subpopulations in different states of melanocytic differentiation, and immunohistochemistry staining demonstrated their distinct anatomical distribution throughout follicular and interfollicular epidermal compartments. Notably, one melanocyte subgroup, marked by increased expression of neurotrophic receptor tyrosine kinase 2 (NTRK2) and genes associated with ribosome biogenesis, exhibited molecular characteristics of progenitor cells. This subpopulation displayed human embryonic stem cell (hESC)-derived melanoblast markers, and their anatomical localization corresponded to that of intermediate melanocyte progenitors.
NTRK2+ melanocytes demonstrated enhanced clonogenicity after UVR exposure in primary cell cultures and in ex vivo whole skin explants. In contrast, NTRK2- melanocytes were suppressed by UVR.

Furthermore, scRNA-seq data of ex vivo melanocytes after UVR exposure revealed the upregulation of genes associated with cell proliferation within NTRK2-expressing melanocytes. In mouse back skin, the ratio of Ntrk2+ melanocytes increased within 24 hours of UVB irradiation, suggesting a proliferative response in these cells in vivo following UVR.

We thus report the discovery in human epidermis of a putative melanocytic cell hierarchy, and of a candidate melanocyte progenitor subpopulation that responds proliferatively to UVR and is thus a candidate cell of origin of melanoma.

Sanjay Krishna - Monash University

CANELIB: a comprehensive platform for LC-MS-based cancer neoantigen discovery using whole exome sequencing data

Sanjay Krishna [1], Mohammad Shahbazy [1], Anthony W. Purcell [1], Chen Li [1]
[1] Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Australia

Cancer-specific neoantigens refer to self-peptides with mutations that are presented by the human leukocyte antigen (HLA) complex on the cancer cell surface for T cell recognition. These neoantigens can serve as promising targets for the development of cancer vaccines and antigen-based immunotherapy. Liquid Chromatography-Mass Spectrometry (LC-MS) has become the mainstream technique for detecting the spectra of HLA-presented peptides and identifying them by searching a reference proteome library against the raw spectra data. However, the reference proteome library often only contains canonical wild-type protein/peptide sequences and lacks cancer-specific mutations, hindering the identification of cancer-specific neoantigens. To tackle this challenge, we developed a comprehensive platform, termed CANELIB, for LC-MS-based identification of cancer-specific neoantigens. To construct cancer-specific mutant peptide libraries, we collected numerous missense mutations identified by whole exome sequencing (WES) from various sources, including the Catalogue of Somatic Mutation in Cancer (COSMIC), the Cancer Cell Line Encyclopedia (CCLE), and published literature. To date, CANELIB documents 854,813 mutant peptides of 33 cancer types from 1428 cancer cell lines and 52 tumour tissues. CANELIB contains a large-scale user-friendly knowledgebase for users to browse, search, and download mutant peptide libraries for their cancer of interest. In addition, CANELIB provides customisable source codes for users to generate their mutant peptide libraries using WES data. To assess the usability of CANELIB, we searched our raw spectra data of MDA-MB-231 (a triple-negative breast cancer cell line), HCT116 (a colorectal cancer cell line), and various melanoma cell lines using both PEAKS Online and MSFragger. The results demonstrate that our libraries successfully identified mutated peptides that were not previously found using the canonical proteome library, thereby providing promising mutant peptide candidates for further immunogenicity assessment. We anticipate that CANELIB can accelerate the identification of cancer neoantigens, thereby better informing the development of targeted immunotherapies.

Lachlan Cain - Peter MacCallum Cancer Centre

ATAC of the Clones: Genetic subclone identification and copy number estimation in cancer from single-cell ATAC-seq

Lachlan Cain [1], Anna Trigos [1,2,3]
[1] Peter MacCallum Cancer Centre

[2] The University of Melbourne
[3] Monash University

Heritable gene expression changes in cancer may have genetic causes, such as copy number aberrations (CNAs), or non-genetic causes such as epigenetic modifications. Unfortunately, CNAs and epigenetic modifications are not easily distinguished in ordinary single-cell RNA-seq and ATAC-seq analyses of cancer. To address this problem, we present ATAClone, a tool for identifying genetic subclones in single-cell ATAC-seq and estimating their DNA copy number independently of any epigenetic modifications.

ATAClone implements both pre- and post-processing steps to minimise the impacts of confounding technical bias and epigenetic signals while preserving DNA copy number signal. Additionally, ATAClone performs copy number estimation jointly among cells in a "clustering-first" approach, borrowing concepts from single-cell RNA-seq analysis such as variance-stabilising transformations, principal component analysis, and graph-based clustering. These innovations enable ATAClone to resolve more genetic subclones with more precise copy number estimates than competing methods.

Using a diversity of 10X multiome single-cell RNA-seq/ATAC-seq cancer samples, we find that genetic subclones identified by ATAClone are largely consistent with existing RNA-seq and ATAC-seq-based copy number estimation methods though more comprehensive and readily identifiable. Moreover, using matched whole-genome DNA sequencing as a ground truth, we show that copy number estimates from ATAClone are more accurate than those derived with existing methods.

Finally, we demonstrate the utility of ATAClone in the multiome setting with a case study of a prostate cancer tumour consisting of multiple genetic subclones which independently differentiate from an adenocarcinoma to a neuroendocrine phenotype an inference which would not be possible with ordinary single-cell ATAC-seq or RNA-seq analysis.

ATAClone represents an important tool for disentangling the heritable genetic and non-genetic changes which underlie cancer from single-cell ATAC-seq data. When combined with orthogonal analyses, ATAClone enables separate measurement of CNAs and epigenetic modifications, providing deeper insight into the evolutionary history and adaptive forces driving a tumour.

Session Four - Methods (Chair: Claire Sun)

Sung-Young Shin - Monash University

Leveraging Machine Learning and Patient-Derived Explants to Discover Predictive Biomarkers of CDK4/6 Inhibitor Efficacy in Prostate Cancer Patients

Sung-Young Shin[1],[2], Margaret M. Centenera[3],[4], Joshua T. Hodgson[3],[4], Lisa M. Butler[3],[4], and Lan K. Nguyen[1],[2]
[1]Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC, Australia
[2]Cancer Program, Biomedicine Discovery Institute, Monash University, Clayton, VIC, Australia
[3]South Australian Immunogenomics Cancer Institute and Freemasons Foundation Centre for Men's Health, University of Adelaide, Adelaide, SA, Australia
[4]South Australian Health and Medical Research Institute, Adelaide, SA, Australia

CDK4/6 inhibitors, used in conjunction with hormone therapy, has significantly improved the prognosis of hormone receptor-positive cancers. Despite their initial success, a substantial fraction of tumours exhibits a limited responsiveness to CDK4/6 inhibitors from the outset, posing a major clinical challenge.
The unique genetic composition and molecular alterations of an individual patient can impact their responses and likelihood of experiencing resistance to CDK4/6 inhibitors. Thus, it is essential to identify molecular biomarkers that predict patients' sensitivity and/or resistance to CDK4/6 inhibitors. So far, several individual genes have been suggested as potential biomarkers for predicting sensitivity to CDK4/6 inhibition, but none of these alone has provided clinical utility.

In this study, we employed advanced machine learning approaches and leveraged prostate cancer patient-derived explants (PDEs) to identify predictive biomarkers for Ribociclib, a clinically-approved CDK4/6 inhibitor. Specifically, we utilise an ensemble feature selection approach and integrate domain knowledge into the feature selection process by considering the network distance between drug targets and genes.

Among the seven machine learning models tested, we found that support vector regression (SVR) and artificial neural networks (ANN) exhibited the best performance in predicting drug sensitivity. Pathway enrichment analysis of the selected biomarkers reveals a high enrichment of pathways associated with reactive oxygen species, DNA repair, and cell cycle checkpoints. Critically, the optimal selection of 10 features predicted the drug response of PDEs with an 80% accuracy rate, significantly surpassing the 60% accuracy achieved using putative CDK4/6 inhibitor biomarkers reported in the literature.
In summary, our method employs AI, PDEs and domain knowledge to identify biomarkers for Ribociclib drug responses. These predictive biomarkers undergo validation using independent preclinical models and hold the potential to become clinically effective biomarkers for prostate cancer patients.

Charla Lu Chai - Monash University

Understanding Mechanisms That Drive Disease Relapse in Early Stage, Low-risk Colorectal Cancer

Lu Chai [1,2], David Nickless [4], Paul J. McMurrick3, Vana Madelli [3], Jiangning Song [5,6,#], Helen E. Abud [1,2,#], Rebekah M. Engel [1,2,3,#]
[1] Department of Anatomy and Developmental Biology, Monash University, Wellington Road, Clayton, Australia.
[2] Stem Cells and Development Program, Monash Biomedicine Discovery Institute, Monash University, Wellington Road, Clayton, Australia.
[3] Cabrini Monash University Department of Surgery, Cabrini Hospital, 183 Wattletree Road, Malvern, Australia.
[4] Anatomical Pathology Department, Cabrini Pathology, Cabrini Hospital, Malvern, Australia.
[5]Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Australia.
[6] Monash Data Futures Institute, Monash University, Melbourne, Australia.
[#] These authors contributed equally to this study.

Despite the relatively high 5-year survival rates for patients diagnosed with early-stage colorectal cancer (CRC), up to 20% of those who have undergone curative surgery will experience disease relapse. Both local and distant disease recurrences are linked to a poor prognosis. Adjuvant chemotherapy is rarely used for patients with early-stage, low-risk CRC (T1-3, N0, M0) due to the lack of survival benefit, while existing methods for assessing the likelihood of disease recurrence have low predictive validity. As such, there is an urgent need to accurately predict the likelihood of disease recurrence and to understand the mechanisms that drive disease relapse. In this study, we focus on predicting recurrence and identifying biomarkers that are correlated with disease recurrence in early-stage CRC through the development of a

multi-omics framework. First, we developed a deep-learning-based computational framework using publicly available whole-slide images (WSI, n = 607, available in The Cancer Genome Atlas, TCGA) of histopathology slides for predicting overall survival rates in CRC patients. Next, we transferred the learned information to a Graph Convolutional Networks (GCN) network to predict disease recurrence by adding WSIs of CRC patients from our private dataset, with the inclusion of patients across multiple tertiary institutions, each linked to extensive clinical data prospectively entered in the Cabrini Monash Colorectal Neoplasia Database (CMCND). Meanwhile, the features of multiplex immunofluorescence images on tissue microarrays (TMAs) from the same patient dataset will be integrated with those of WSIs and the clinical records of correlated patients to connect WSIs with clinical and proteomic profiles of interest. Altogether, the developed framework enables the characterisation of the interactions among the diagnostic regions of interest and generates the inference of disease recurrence. By analysing the saliencies of prediction outcomes, our approach is capable of detecting biomarkers associated with the underlying mechanisms contributing to disease recurrence.

Liyang Fei - Peter MacCallum Cancer Centre

BARBIE: An R package for analysing barcode count data from clonal tracking experiments

Liyang Fei [1,2], Jovana Maksimovic [1,2], Alicia Oshlack [1,2,3]
[1] Peter MacCallum Cancer Centre, Melbourne, Victoria 3000, Australia
[2] Sir Peter MacCallum Department of Oncology, The University of Melbourne, Victoria 3010, Australia
[3] School of Mathematics and Statistics, The University of Melbourne, Victoria 3010, Australia

In cancer research, comparing the clonal composition of pools of cells before and after perturbations reveals important features such as treatment resistant clones. Recent advances in clonal tracking protocols have revolutionised our ability to track progeny cells back to their original progenitor cell. One popular technique relies on each original cell incorporating a unique DNA barcode and passing it down to its offspring, whose barcodes are subsequently sequenced. The count of each barcode detected in the sample indicates the number of progeny cells derived from the original cell. However, there is currently a lack of bioinformatic tools for robustly analysing barcode count data from bulk sequencing of DNA barcodes.

Here, we introduce BARBIE, an R package designed to analyse barcode count data across sample groups. In addition to providing useful data structures and novel functions for cleaning up, summarising and visualising barcode count data, BARBIE implements two statistical tests for comparing sample groups:
1) Differential barcode occurrence - identifies initial cells more likely to survive or divide under specific conditions compared to others;
2) Differential barcode proportion - identifies initial cells significantly expanding in one condition compared to others.

We have applied BARBIE to real datasets of barcoded Hematopoietic Stem Cells undergoing various perturbations. BARBIE can handle complex experimental design by using the limma linear modelling framework to highlight the effects of prioritised factors while taking into account other variables. Results demonstrate that BARBIE can identify biologically interesting barcodes showing statistically significant occurrence or proportion changes. Furthermore, BARBIE is interoperable with existing barcode sequence processing packages like BARtab.

Overall, BARBIE is the first R package to implement generalised statistical methods for robustly testing barcode significance in occurrence or proportion between sample groups. This is a major advance over the more commonly used descriptive methods that do not account for variability between samples.

Pablo Acera Mateos - Children's Cancer Institute

Cryptico: A Computational Approach for Exploring Cryptic Exons in the Search for Novel Therapeutics

Pablo Acera Mateos[1], Keith Sia[2], Nivedita Iyer[1], Shi Yan Lee[1], Greg Arndt[2], Ian Street[2], Antoine de Weck[1]
[1] Computational Drug Discovery Biology. Therapeutic Innovation Group. Therapeutic Discovery. Children's Cancer Institute
[2] Therapeutics Innovations for Kids (THINK). Therapeutic Innovation Group. Therapeutic Discovery. Children's Cancer Institute.

The development of new cancer therapies is constrained by the proteome's limited druggability, estimated at only 10-15%. Additionally, the increasing recognition of non-coding RNAs' roles in disease underscores the urgency for innovative therapeutics that move beyond conventional protein-targeting strategies. The emergence of RNA-targeting small molecules presents a promising avenue for novel therapeutic approaches, potentially expanding the range of actionable targets. Recent advancements include the use of splicing modulator compounds (SMCs) to suppress gene expression by introducing deleterious cryptic exons in a sequence specific manner. In this study, we developed a computational method to systematically identify potential targets for SMCs, specifically cryptic exons within the human genome. We evaluated the therapeutic potential of these exons based on their ability to disrupt gene expression through PTCs, frameshift mutations, or the alterations of protein folding due to the introduction of novel peptide chains. Our analysis revealed 41,664 previously unannotated cryptic exons, some of which are predicted to interfere with several oncogenes traditionally considered undruggable. We further validated the existence of these cryptic exons by modifying one of the spliceosome components to match a specific exon within the RAS pathway. Our findings provide a valuable resource for exploring potential therapeutic targets of SMCs and lay the groundwork for future drug screening efforts.

Ashley Weir - The Walter and Eliza Hall Institute

IdentifiHR: predicting homologous recombination deficiency in high-grade serous ovarian carcinoma through gene expression.

Ashley L Weir [1], Samuel C Lee [1], Mengbo Li [1], Chin Wee Tan [1], Susan J Ramus [2], Nadia M Davidson [1]
[1] The Walter and Eliza Hall Institute, Parkville, VIC, Australia
[2] School of Clinical Medicine, UNSW Medicine and Health, University of NSW Sydney, Sydney, NSW, Australia

Approximately half of all high-grade serous ovarian carcinomas (HGSC) have therapeutically targetable defects in the homologous recombination (HR) DNA repair mechanism. While there are genomic methods to identify HR deficient (HRD) patients, there are no gene expression-based tools to predict HR repair status in HGSC specifically, despite the research and clinical utility. We aimed to build the first HGSC-specific model to predict HR repair status using gene expression. We separated The Cancer Genome Atlas (TCGA) cohort of HGSCs (n = 361) into training (n = 288) and testing (n = 73) sets and labelled each case as being HRD or HR proficient (HRP) based on the clinical gold standard for classification, being a score of HRD genomic damage in addition to HR-gene mutations. Using the training set, we performed differential gene expression analysis using limma-voom between HRD and HRP samples. The normalised and scaled expression of the 2604 significantly differentially expressed

genes was then used to tune and train a penalised logistic regression model. Our model, IdentifiHR, is an elastic net penalised logistic regression model that uses the expression of 209 genes to predict HR status in HGSC. These genes capture known regions of HR-specific copy number alteration, which impact gene expression levels, and preserve the genomic damage signal. IdentifiHR has an accuracy of 85% in the TCGA test set and of 91% in an independent cohort of 99 samples of the Australian Ovarian Cancer Study (AOCS), collected from primary tumours before (n=74/99) and after autopsy (n=6/99), in addition to ascites (n=12/99) and normal fallopian tube samples (n=7/99). Further, IdentifiHR outperforms existing gene expression-based models for HR status, being BRCAness and MutliscaleHRD. IdentifiHR is an accurate model to predict HR status in HGSC using gene expression, that is currently being built as an R package.

Nathan Hall - University of Melbourne

Predicting NSCLC patient response to immunotherapy using IIF profiling

Nathan E Hall[1], Jared Mamrot[1] & Robyn A Lindley[1,2]
[1] GMDx Genomics Ltd, Melbourne, Victoria, Australia
[2] Faculty of Medicine, Dentistry & Health Sciences, Department of Clinical Pathology, University of Melbourne, Melbourne, Victoria, Australia

Background: Aberrant activity of AID/APOBEC deaminase enzymes can cause 'off-target' somatic mutations in cancerous cells. Mutations associated with deaminases and other mechanisms can be quantified using metrics relating to motif usage, strand bias, transitions/transversions, codon context, and amino acid changes. Collectively, these metrics form an Innate Immune Fitness (IIF) profile. The aim of this project was to conduct IIF profiling on a cohort of Non-Small Cell Lung Cancer (NSCLC) patients treated with immune checkpoint inhibitors (ICI), and use the IIF profiles to build and evaluate a predictive model. Methods: Whole exome and progression-free survival (PFS) data was obtained from Rizvi 2015, Hellmann 2018, Miao 2018, Fang 2019, Frigola 2021, and Ravi 2023 (n=516). IIF profiles were generated using CRIS (v5.0.1; GMDx Genomics Ltd). Patients were classified as a â€˜Responderâ€™ (PFS>12 months or Complete Response; n=149) or â€˜Non-Responderâ€™ (PFSâ‰¤12 months; n=367). Machine learning models were generated using XGBoost (H2O.ai platform) and evaluated using multiple rounds of cross-validation. Patient response predictions were collated for each patient and a consensus â€˜IIF Scoreâ€™ was calculated. Multivariable analysis of IIF Score, TMB (10mut/Mb) and PD-L1 status was conducted using a Cox proportional-hazards model. Results: Predictive accuracy of IIF Scores was 75% (Sensitivity=56%; Specificity=83%) with a Hazard Ratio (HR) of 0.43 (0.32-0.57; p<0.001; corrected for TMB and PD-L1). In comparison, the predictive accuracy of TMB was 69% (Sensitivity=58%; Specificity=73%), corrected HR=0.74 (0.57-0.95; p=0.02). â€˜PD-L1 >50% accuracy was 63% and corrected HR=0.68 (0.47-0.98; p=0.04). IIF Score AUC was significantly higher than TMB (0.77 vs 0.70; p<0.001). Conclusions: IIF Score was the strongest predictor of patient response to ICI. Despite the inherent limitations of combining data from multiple cohorts, IIF Score outperformed TMB and PD-L1 in predictive accuracy, HR and AUC. These results support the use of IIF profiling as a novel ICI biomarker in NSCLC patients.

Lightning Talks and Posters

Session 1 - 10:15-10:30am

P1: Priyadarshana Ajithkumar - University of Ottago

Epigenomic profiling of paired colorectal cancer samples to identify epigenetic drivers of early metastasis

Priyadarshana Ajithkumar [1], Sharon Pattison [2], John McCall [3], Euan J. Rodger [1] and Aniruddha Chatterjee [1]
[1] Department of Pathology, Dunedin School of Medicine, University of Otago, Dunedin, New Zealand
[2] Health New Zealand Te Whatu Ora – Capital, Coast and Hutt Valley, New Zealand
[3] Department of Surgical Sciences, Dunedin School of Medicine, University of Otago, Dunedin, New Zealand

Colorectal cancer (CRC) is the third leading cause of death from cancer worldwide and metastasis is the leading cause of CRC-associated mortality. Lymph node metastases are an important factor in determining both stage (stage III) and treatment options for CRC patients. Patients diagnosed with advanced-stage colorectal cancer (stage III or higher) have a poorer prognosis compared to those with early stage disease (stage I or stage II). Epigenetic modifications, such as aberrant DNA methylation patterns, are emerging as a key factor in cancer metastasis. There is a lack of comprehensive understanding regarding the specific DNA methylation alterations and mechanisms that lead to CRC lymph node metastasis. Identifying genes and pathways associated with DNA methylation changes and lymph node metastasis could aid in the discovery of prognostic biomarkers. This study presents comprehensive whole-genome scale DNA methylation and transcriptome profiling from primary tumours and lymph node metastasis samples from the same patients providing valuable insights into the molecular landscape of CRC progression and metastasis. We identified 462 significantly differentially methylated CpGs (DMCs) that exhibited a methylation difference > 20% (false discovery rate adjusted P value < 0.001) between primary and lymph node samples. Among these DMCs, 403 were hypermethylated and 59 were hypomethylated in lymph node samples compared to primary CRC samples. Pathway enrichment analyses revealed associations between the DMCs and key biological processes, including lymphocyte migration into lymphoid organs, cell adhesion and germinal center B cell differentiation. The identification of these DNA methylation alterations and their correlation with critical biological pathways highlights the pivotal role of epigenetic changes in CRC lymph node metastasis. Understanding the molecular mechanisms underlying these epigenetic changes could help to identify prognostic markers.

P2: Islam Alagawani - Swinburne University of Technology

DFT-based conformational and NMR analysis of a SARS-CoV-2 Mpro inhibitor ML188

Islam Alagawani [1], Feng Wang [1], Vladislav Vasyliev [2]
[1] Department of Chemistry and Biotechnology, School of Science, Computing and Engineering Technologies, Swinburne University of Technology, Melbourne, Victoria 3122, Australia
[2] National Computational Infrastructure, Australian National University, Canberra, ACT 0200, Australia

In the pursuit of effective therapeutics against SARS-CoV-2, the main protease (Mpro) has emerged as a promising target due to its crucial role in viral replication. ML188 (Fig 1), initially designed as a non-covalent inhibitor to target SARS1-Mpro, has been reported to bind SARS2-Mpro with an IC50 potency of 1.5 ± 0.3 μM in cellular assays [1]. Understanding the structural dynamics and conformational flexibility of small molecule inhibitors targeting SARS-CoV-2 is crucial for developing COVID-19 inhibitors. In this study, we use a recently developed robust quantum mechanical conformational sampling method [2] to determine low-lying conformer clusters of ML188 in solution (CDCl3) and in complex with Mpro,

along with their properties such as NMR spectra. The DFT-based B3LYP/6-31G method is employed, yielding the electronic energies of up to 20 ML188-A conformers above the global minimum structure (Fig 1). Our preliminary results contribute to the understanding of ligand-protein interactions for favorable conformations both with and without interaction with Mpro, providing valuable insights for further Mpro inhibitor development.

P4: Krutika Ambani - Peter MacCallum Cancer Centre

How normalisation can affect everything in Cut and Run differential binding analysis

Krutika Ambani [1], Antonio Ahn [1,2], April Watt [1,2], Shom Goel [1,2], Alicia Oshlack [1]
[1] Peter MacCallum Cancer Centre, Melbourne, VIC, Australia
[2] Sir Peter MAcCallum Department of Oncology, University of Melbourne, Melbourne, VIC, Australia

Cut&Run (Cleavage Under Targets and Release Using Nuclease) is an increasingly popular method of profiling protein binding (transcription factors, histone modifications, etc) across the whole genome. In differential binding analysis, normalisation is an essential step to remove technical variation, so that the remaining variation between experimental conditions are biological. Choosing the most reliable normalisation method is critical for accurate and meaningful biological interpretations in downstream analysis. Previous analysis of Cut&Run data has shown that spike-in normalisation has been important for assessing binding positions however this has not been explored in the context of differential binding. Here, we examine various normalisation methods (background, spike-in, library-size) and show that the choice of method can result in huge discrepancies in downstream differential analysis. We further evaluate and make recommendations on the most reliable normalisation method.

P6: James Comben - Peter MacCallum Cancer Centre

Multi-omics approach to understanding extramedullary multiple myeloma

James Comben [1, 2], Amit Khot [1,2,5], Elaine Sanij [1,3,4] and Anna Trigos [1,2,3]
[1] Peter MacCallum Cancer Centre
[2] The University of Melbourne
[3] Monash University
[4] St. Vincent's Institute of Medical Research
[5] Royal Melbourne Hospital

Multiple Myeloma (MM) is a cancer of plasma B cells which is usually restricted to the bone marrow. However, in a subset of patients representing up to 8% upon primary diagnosis and up to 20% of patients with relapsed disease, MM cells escape the bone marrow and develop MM lesions in other organs which is known as extramedullary multiple myeloma (EMM). This form of the disease is highly aggressive and patient outcomes remain poor, despite the wealth of treatment options for MM patients. At present there are no therapeutic approaches specifically developed for EMM. With a rising incidence of EMM in the age of novel treatment options such as CAR-T and bi-specific antibodies, it is critical to understand the biology that underpins the ability of MM cells to survive and grow outside of the bone marrow whilst resisting current treatments. This will enable the development of novel treatments specifically designed for these patients.

In this project we will seek to understand clonal evolution in the development of extramedullary disease, identify non-genetic changes influencing the development and persistence of EMM and understand the role of the local microenvironment in modulating EMM. We will employ a multi-omics approach to profile MM patient samples from bone marrow and EMM lesions from a retrospective cohort. Whole genome

sequencing, single nuclei RNA sequencing, spatial transcriptomics and bulk methylation data will be generated and analysed and integrated computationally using a combination of approaches, including bioinformatics, network biology, machine learning, simulations, and computational modelling.
This project will generate and analyse the largest integrated multi-omics data set for the study of EMM, finally shedding light on the biological processes at play in EMM. This will inform the creation of novel personalised treatments directed towards EMM patients leading to improved patient outcomes.

P11: Andrea Gillespie - Peter MacCallum Cancer Centre

ADAR editing aware alignment reveals novel class of mega-intergenic RNAs

Andrea Gillespie[1,2], Brian J. Liddicoat[1,2], Wendy Jia[1,2], Andrew B. Das[1,2], Laura Scolamiero[1,2], Naiara G. Bediaga[1,2], Yih-Chih Chan[1,2], Christa van der Maarel[1,2], Andrew A. Guirguis[1,2], Kelsy Prest[1,2], Laure Talarmain[1,2], Cassandra Litchfield[1,2], Enid Y. N. Lam[1,2], Sebastian Hollizeck[1,2], John R. Horton[3], Benjamin Kroeger[4], Tao Wang[5], Francesca Zappacosta[5], Roland S. Annan[5], Rab K. Prinjha[6], Melissa B. Pappalardi[7], Michael T. McCabe[7], Ching-Seng Ang[8], Matthew S. Graus[9], Mathias Francois[9], Kieran F. Harvey[1,2,4], Sarah-Jane Dawson[1,2,10], Xiaodong Cheng[3], Mark A. Dawson[1,2,10]
[1] Peter MacCallum Cancer Centre, Victoria, Australia
[2] Sir Peter MacCallum Department of Oncology, The University of Melbourne, Victoria, Australia
[3] Department of Epigenetics and Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center, Texas, USA
[4] Department of Anatomy and Developmental Biology, and Biomedicine Discovery Institute, Monash University, Victoria, Australia
[5] Medicinal Science and Technology, GlaxoSmithKline, Pennsylvania, USA
[6] GlaxoSmithKline R&D Stevenage, UK
[7] Cancer Epigenetics Research Unit, Oncology, GlaxoSmithKline, Pennsylvania, USA
[8] Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Victoria, Australia
[9] The David Richmond Laboratory for Cardiovascular Development: Gene Regulation and Editing, The Centenary Institute, The University of Sydney, New South Wales, Australia
[10] The University of Melbourne Centre for Cancer Research, The University of Melbourne, Victoria, Australia

As DNA hyper-methylation is a common hallmark of cancer, several hypo-methylating agents have been approved for cancer treatment. We have developed a novel first-in-class catalytic inhibitor that only targets DNA methyltransferase 1 (DNMT1i). This unique drug has allowed us for the first time to specifically interrogate the effects of catalytic inhibition of DNMT1 and in doing so gain a better understanding of transcriptomic consequences of maintenance in DNA methylation. We found that DNMT1i and consequent relocation of chromatin bound DNMT1 does not induce large changes in the expression of coding genes, rather it reactivates endogenous retroviral elements which contain dsRNA structures that stimulate the cell-intrinsic viral mimicry response. These dsRNAs are specifically edited by ADAR which converts adenosine to inosine by deamination making it structurally similar to guanine and therefore binds with cytosine. In terms of sequencing edited transcripts, ADAR-induced edits result in either A>G or T>C (opposite strand) conversions with respect to the reference genome. These mismatches present difficulty in alignment in regions where hyper-editing occurs as many reads in these regions are discarded in a typical alignment. By utilizing an alignment algorithm designed to ignore these specific base conversions we discovered heavily edited, extraordinarily long (>200Kb) transcripts which are induced by the presence of DNMTi. These comprise an as yet undescribed class of RNAs which we have termed mega-intergenic RNAs (mintRNA) stemming from long tandem repeat (LTR) promoters functioning as treatment induced non-annotated transcription start sites (TSS, TINATs) as seen with other epigenetic therapies. The mintRNAs are all driven by TINATs, encompass multiple repetitive elements and contain several predicted dsRNAs, indicating possible induction of viral mimicry. This study elucidates how DNMT1 regulates intergenic transcription highlighting the importance of DNA methylation in supressing transcription of these intergenic regions.

P13: Calandra Grima - Peter MacCallum Cancer Centre

Everything old is new again: exploring a probe-based approach for single-cell RNA-sequencing

Calandra Grima [1], Melanie Neeland [2], Shivanthan Shanthikumar [2], Alicia Oshlack [1,3,4], Jovana Maksimovic [1,2,3]
[1] Peter MacCallum Cancer Centre, Computational Biology Program, Parkville, Australia
[2] Murdoch Children's Research Institute, Respiratory Group, Parkville, Australia
[3] The Sir Peter MacCallum Department of Oncology, The University of Melbourne, Parkville, Australia
[4] School of Mathematics and Statistics, The University of Melbourne, Parkville, Australia

Single-cell RNA-sequencing (scRNA-seq) provides a rich view of cellular activities but fixed tissue archives are incompatible with scRNA-seq. New scRNA-seq protocol Flex fixes fresh samples with paraformaldehyde upon collection, allowing analysis of archived tissues and facilitating better batch design. Understanding the unique quirks of any new scRNA-seq technology is vital for accurate bioinformatic analysis, so we aim to identify and correct technical biases in Flex count data.
Flex is a Chromium-compatible probe-based protocol developed by 10X Genomics. Unlike existing protocols that utilise poly-A capture of transcripts, Flex instead captures transcripts using a microarray-like set of probes that hybridise to specific positions along target transcripts. The human probe set covers all isoforms of 18,532 coding genes and transcripts can be targeted by multiple different probes. The hybridised probes, rather than transcripts, are sequenced and counts of probes targeting the same gene are summed. Downstream analysis treats Flex data alike 3' scRNA-seq data.

Generations of technologies from microarrays to scRNA-seq have highlighted the importance of understanding how technical biases influence gene expression data – something which is currently unknown for Flex. Our analyses have identified technical variation in the performance of different probes targeting the same transcripts. We are investigating how this variation relates to characteristics like probe nucleotide composition and hybridisation position along its target transcript. For example, we found that extreme probe GC content tends to result in lower probe counts. We are also investigating the impact of differences in the number of probes targeting each gene and transcript on gene counts, as this may over- or under-represent certain genes.

Once the technical biases in Flex data are understood, we will develop an alternative approach for summarising Flex probe counts into gene counts that accounts for these technical factors, ensuring that the counts produced by Flex accurately reflect gene activity within cells.

P16: Alexander Hill - Swinburne University of Technology

Structural investigations of first-generation EGFR targeting tyrosine kinase inhibitors using molecular dynamics

Alexander Hill [1]
[1] Swinburne University of Technology

Tyrosine kinase inhibitors (TKI's) that target the epidermal growth factor receptor (EGFR) have been successful anti-cancer drugs, with first-generation designed ligands Erlotinib and Gefitinib receiving FDA approval. Although resistance to these drugs arise in the way of EGFR protein mutations, newer generations of these TKI's have been developed to overcome these issues. Despite this, Erlotinib and Gefitinib can still be used as a first-line treatment as well as being supplemented with a third generation TKI drug, Osimertinib. Understanding of the structure/properties relationships of these inhibitors is still lacking, especially since small substitutions to the molecule can greatly alter the IC50 and other important

properties. This research has focused on the atoms on the tail end of the TKI's, whereas the head of the molecule docks into the rigid ATP binding site, the tail chains are solvent-exposed and experience more movement. Using molecular dynamics and post-processing methods, we have removed one tail chain on three different TKI's to observe the effects on stability and binging affinity. Gefitinib was of particular interest, as Erlotinib and most other first generation drugs have symmetrical tail chains, but Gefitinib doesn't so it could provide more insight into how important symmetry of these chains is. Results showed that even though these tail chains aren't in the binding pocket, they can still form significant intramolecular bonds with nearby protein residues, and can impact the binding energy of the molecule and thereby its stability, both positively and negatively. Providing a greater understanding of these inhibitors and aiding in the design and development of the next generation of these anti-cancer drugs.

P17: Nhan Hoang - Monash University

Targeting the PI3K-Hippo-ERK signalling crosstalk network to devise effective combinatorial treatment for breast cancer

Nhan Hoang [1,2], Milad Ghomlaghi [1,2], Tao Zhang [1,2], Sung-Young Shin [1,2], Lan K. Nguyen [1,2]
[1] Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia
[2] Biomedicine Discovery Institute, Monash University, Clayton, Victoria 3800, Australia

The PI3K-mTOR signalling pathway controls critical survival processes and is aberrantly activated in 40% of breast cancer. Thus, targeting aberrant PI3K signalling for breast cancer treatment has been a major research priority. Mounting evidence suggests significant and complex crosstalk between the PI3K pathway and other oncogenic pathways, and that these enable tumour cells to bypass the effect of PI3K-based therapy. However, our understanding of PI3K-related crosstalk and how the crosstalk network mediates the effect of PI3K-based therapy remains limited. In this study, we analysed the dynamic interplay between the PI3K and the Hippo/YAP and ERK MAPK pathways using a systems biology approach that integrates biological experiments and predictive modelling. We systematically perturbed the constituent pathways using small-molecule drugs and examined the dynamic response of the whole network. We found that PI3K inhibition using a clinically-relevant drug (alpelisib) significantly increased YAP nuclear localisation and expression of its target genes in a time- and dose-dependent manner in breast cancer cells. This suggests YAP signalling as a novel escape route for tumour cells to bypass PI3K inhibition. We confirmed that dual inhibition of PI3K and YAP indeed synergistically blocks cancer cell growth and induces cell death. In addition, the data obtained from biological experiments were utilised to construct a new computational model of the integrated PI3K-Hippo-ERK network, which helped predict additional synergistic drug combinations. We are currently validating these experimentally. Together, our interdisciplinary study provides new quantitative frameworks for the analysis of signalling crosstalk and discovery of effective combination therapies for breast cancer.

P18: Yunhui Jeong - Monash University

Timing is Key: Rationalizing Drug Treatment to Enhance FGFR-Directed Combination Therapy in Triple Negative Breast Cancer

Yunhui.Jeong[1], Tao Zhang[1], Roger Daly[1][2], Sung-Young Shin[1], Lan K. Nguyen[1][2]
[1] Biomedicine Discovery Institute, Cancer Program, Monash University, Clayton, Victoria 3800, Australia
[2] Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia

Triple-negative breast cancer (TNBC) is a highly aggressive form of breast cancer, which presents limited targeted treatment options. Targeting FGFR signalling is emerging as a potential therapeutic strategy for TNBC, with several FGFR inhibitors currently undergoing clinical trials. However, resistance to FGFR-targeted therapy poses a significant challenge, and more effective combination therapy strategies are needed. Unfortunately, most drug combination studies do not consider the impact of drug treatment timing and order on anti-tumour efficacy, which hampers the optimization of maximum therapeutic benefits.

In this study, we aimed to address this challenge by focusing on the FGFR system using a quantitative systems biology approach that combined mechanistic mathematical modelling and experimental studies. We constructed a mechanistic model of the integrated FGFR signalling network in TNBC MDA-MB-453 cells, which we calibrated with time-course signalling drug response data. We then used the model to perform systematic simulations of pair-wise concurrent and sequential drug combinations targeting the network nodes, comparing their effects under various conditions. Our model predictions suggested that sequential inhibition of FGFR4 followed by AKT may be more effective than concurrent inhibition, and we used model simulations to inform the optimal time gap between drugs. Our experimental results supported the model prediction, showing that FGFR4 inhibition with either BLU9931 or H3B-6527 triggered potent reactivation of AKT (and ERK signalling) in MDA-MB-453 cells. In short-term proliferation assays, sequential treatment with BLU9931 and then AKT inhibitor MK-2206 was significantly better than single-drug treatments in blocking cancer cell viability, and better than concurrent BLU9931/MK-2206 treatment in blocking the tumour colony formation capacity of these cells.

Overall, our study provides deeper insights into the role of timing in combination treatment efficacy, and offers a novel mechanistic framework that enables the quantitative assessment and prediction of effective sequential combination regimens. Our findings suggest that sequential targeting of FGFR4 and AKT may be a useful therapeutic strategy for TNBC.

P22: Yunduo (Dawson) Lan - Monash University

Engineering network-based dynamic features to predict drug sensitivity

Yunduo (Dawson) Lan [1] Sungyoung Shin [2] Lan K Nguyen [3]
[1] Cancer Program, Biomedicine Discovery Institute, Monash University, Victoria, Australia
[2] Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Victoria, Australia.

Computational discovery of cancer drug response biomarkers typically employs machine learning pipelines to analyse large pharmacogenomic datasets. However, these -omics datasets, while extensive, lack critical information about intracellular protein signalling dynamics, which can determine drug response outcomes. The static nature of these datasets thus limits their ability to predict drug response mechanistically.

Here, we present a method to engineer 'dynamic' features from established Ordinary Differential Equation (ODE) models of cancer signalling networks. These ODE models incorporate mechanistic biological knowledge as kinetic equations based on biochemical principles, such as Michaelis–Menten kinetics. Using transcriptomic data, the ODE model is individualised to specific cell lines, simulating protein dynamics based on the cell line's transcriptomic profile. Dynamic features are then computed from the simulation data using mechanistic rules to capture the characteristics of the generated protein dynamics time-series.

As a demonstration, we applied this approach to an FGFR4 signalling network model to predict drug response to FGFR4 inhibitors, as FGFR4 overexpression is linked to various cancers, including hepatocellular carcinoma and breast cancer. Clustering analysis showed correspondence between dynamic features and distinct protein dynamics, while feature selection demonstrated a weak but significant association between FGFR4 network features and response to a specific FGFR4 inhibitor, FGFR_0939. Ongoing studies involve integrating ODE model-based dynamic features with -omics data to derive better biomarkers of drug response in cancer. Overall, our work offers a novel strategy for predicting cancer drug sensitivity, bridging the gap between -omics data and mechanistic understanding.

P23: Zhu Jun Law - Monash University

Detection of germline copy number variations in a small targeted-sequencing panel: applications for population genomic screening

Zhu Jun Law [1], Jason A Steen [1], Paul Lacaze [2], Tu Nguyen-Dumont [1,3] on behalf of the DNA Screen investigators
[1] Clinical Genomics Laboratory, School of Translational Medicine, Monash University, Melbourne, Australia
[2] Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia
[3] Department of Clinical Pathology, The University of Melbourne, Melbourne, Australia.

The DNA Screen study is a pilot study of population genomic screening for Hereditary Breast and Ovarian Cancer (HBOC), Lynch Syndrome (LS) and Familial Hypercholesterolemia (FH) in Australia. DNA Screen recruited >10,000 participants, aged 18-40 years.

DNA Screen uses a custom panel targeting the coding regions of 9 genes associated with high risk of HBOC (BRCA1, BRCA2, PALB2), LS (MLH1, MSH2, MSH6) and FH (APOB, PCSK9, LDLR). To date, medically actionable variants called from the DNA Screen dataset have consisted of single nucleotide variants and indels (up to 20bp). Yet, it is estimated that up to 10%, 29% and 10% of pathogenic variations are attributable to copy number variations (CNVs) in HBOC, LS and FH, respectively.

The initial reporting phase of the study didn't include CNVs because most tools developed for CNV calling have been designed for whole genome or large targeted sequencing (exome or large panels) data. The analytical validity of small panels such as ours (88 kbp total) for the detection of CNVs is still unknown.

In this study, we tested ExomeDepth and VS-CNV (Golden Helix). Firstly, simulated BAMs representing copy number (CN) losses (CN=0,1) and gains (CN=3,4) were generated, and the recall rates were calculated for each tool to allow robust statistical thresholds to be determined. Secondly, we applied both tools to a set of 1000 DNAscreen samples, further refining our filtering process (ie, removing frequently recurring CNVs unlikely to be of clinical significance).

Our preliminary analyses demonstrate that both tools have the potential to call CNVs from our very small panel. By expanding DNA Screen's scope to include CNV calling, this study will realise the potential for the DNA Screen test to identify more individuals in the population at risk of life-threatening conditions, so they can access life-saving early interventions, rather than late-stage treatment.

P24: David Le - Peter MacCallum Cancer Centre

3D Spatial Analysis of Cancer Tissues

David Le [1], David Kaplan [2], Anna Trigos [3]
[1] Peter MacCallum Cancer Centre, Victoria, The University of Melbourne
[2] Peter MacCallum Cancer Centre, The University of Melbourne
[3] Peter MacCallum Cancer Centre, The University of Melbourne, Monash University

Spatial technologies have revolutionised our study of cancer tissues. A major limitation of current spatial approaches is the study of cancer at a 2D level, by profiling only a single slice of the tumour. This results in a loss of critical 3D information, potentially leading to the misrepresentation of tissues structures, an over or underestimation of cell interactions, and overlooking key spatial features. Fortunately, 3D spatial imaging technologies are beginning to emerge. This provides an opportunity to get a timely understanding of what information is lost in 2D compared to 3D and develop tools that are capable of capturing spatial tumour organisation features that can only be appreciated, quantified, and extracted from 3D spatial data. In this project we aimed to develop a simulation tool and spatial analysis toolkit for 3D spatial data. We developed spaSim-3D, an R package which allows the generation of 3D tissue structure data, including spherical, ellipsoidal, cylindrical, and network-like structures of cells of multiple cell types, allowing capturing the main patterns of cancer tissue organisation in 3D. We also developed SPIAT-3D, an R package with over 30 functions for the analysis of 3D spatial data.
Testing our 3D spatial analysis toolkit on a diverse set of 3D simulated and real tissue data, we have been able to determine the ability of commonly used spatial metrics in 2D to capture key 3D patterns, identified key features of 3D data that are not captured in 2D, as well as developed novel metrics that are capable of capturing 3D-specific features.
Our 3D simulation and spatial analysis toolkit, spaSim-3D and SPIAT-3D, will empower researchers to study tumour organisation and cell interactions in 3D. By enabling the capture and quantification of novel 3D-specific features, our tools will help the development 3D spatial analysis pipelines in cancer research.

Session 2 - 12:15-12:30pm

P27: Anna Mealy - Hudson Institute of Medical Research

Examining the role of interferon epsilon in pancreatic cancer

Anna Mealy [1], Nicole Campbell [1], Paul Hertzog [1]
[1] Hudson Institute of Medical Research

Pancreatic cancer has the highest mortality rate among all cancer types. It is the fourth leading cause of cancer death in Australia, with a one-year survival rate of 20% and a five-year survival rate of 8%. Pancreatic ductal adenocarcinoma (PDAC) constitutes the majority (80–90%) of cases of pancreatic cancer.

Interferons are a family of cytokines that perform vital functions in the innate and adaptive immune responses. The role of interferons in tumorigenesis is complex, exerting both cytotoxic affects as well as the promotion of cell survival depending on the levels and duration of cytokine stimulation.
A subset of type I IFN, IFN epsilon, has been found to play an important anti-tumour role in ovarian cancer and it has been suggested it may be sufficient to suppress immune evasion and metastasis in pancreatic cancer.

A region on chromosome 9 contains the type I IFN gene cluster located at 9p13-22. Multiple prominent tumour suppressor genes are located on 9p21.3 such as cell cycle inhibitors CDKN2A and CDKN2B.

A recent study has shown that co-deletion of type I IFN with CDKN2A/B occurs in up to 60% of PDAC samples. The loss of CKDN2A provides an environment with enhanced proliferative capacity with co-deletion of type I IFNS providing an additional advantage of enhanced immune evasion.
The objective of this study was to examine IFN epsilon expression in pancreatic cancer and to determine if expression alters throughout cancer progression and metastasis.

Using clinical and RNA sequencing data of pancreatic adenocarcinoma from The Cancer Genome Atlas Project (TCGA) database, differential expression of type I IFN was examined in different pathological stages of pancreatic cancer. Gene set enrichment analysis was performed to determine what pathways were involved in the expression of type I IFNs.

P28: Michael Nakai - Peter MacCallum Cancer Centre

Isopod: Detecting differential transcript usage between cell types from long-read single cell data

Michael Nakai[1], Belinda Phipson[2], Nadia Davidson[2], Alicia Oshlack[1]
[1] Peter MacCallum Cancer Centre, Victoria, Australia, [2] Walter and Eliza Hall Institute, Victoria, Australia

Transcriptomics conventionally relies on short read sequencing to generate gene expression level data. While short reads are both accurate and cheap to generate, they are unable to unambiguously reconstruct full length transcripts due to their sequence overlap. Long reads have the potential to overcome this limitation as the extended read length allows for accurate reconstruction of the original transcript, preserving structural variation. In singe cell data this allows for the investigation of transcript-level expression and usage differences, where the proportion of each isoform within a gene is compared between cell types. Pinpointing transcripts exhibiting differential transcript usage (DTU) can also highlight cancer-driving isoform switching events within tumor cells. Currently there are few established methods for this analysis, and a relative lack of publicly available data to test these methods on.

We first aim to simulate transcript-level single cell data to test DTU methods. Average gene counts are first generated using a gamma distribution, and counts are split between isoforms using a log-transformed linear model based on a real dataset. Individual cell counts are then simulated by sampling from Poisson distributions based on each isoform's mean counts. Isoform switching events are simulated between groups of cells by swapping the gene count proportions allocated to the two most abundant isoforms. The simulation tool is also able to simulate differential gene expression, along with gene and cell count outliers.

We aim to utilise single cell data structures by presenting a permutation-based method to estimate DTU between cell types or clusters. Our method, Isopod, takes a transcript-level counts table which is first passed through a filtering function, discarding low count and sparsely represented genes, reducing noise generated by aligners in the form of ambiguous, low count reads. DTU is then calculated at the transcript and gene level by performing a summed pseudobulk for each transcript between cell types, calculating and storing a chi-squared p-value, and repeating the process after randomly shuffling cells between cell types. A permutation p-value is calculated after 10,000 permutations. Isopod is able to find isoform switching events in public datasets, and has a consistently high sensitivity when tested on simulated data creating using the aforementioned tool.

P29: Stefanie Navaratnam - Peter MacCallum Cancer Centre

Comparing the Mutation Profile of Epigenetic Modifier Genes by Evolutionary Age in Cancer

Stefanie Navaratnam [1][2], David Goode [1]
[1] Peter MacCallum Cancer Centre
[2] The University of Melbourne

The hallmarks of cancer represent phenotypes acquired during tumour development and can be likened to the loss of multicellular features. In cancer, gene regulatory networks (GRNs) that evolved to enforce multicellularity through regulation of genes inherited from unicellular ancestors are extensively rewired. Given their global effects on gene regulation, we posit the impacts of alterations to epigenetic modifier genes on GRN rewiring in cancer is related to at what point they emerged during the evolution of multicellularity. To investigate, we focused on the KMT2 family of methyltransferases, whose members emerged at different points during evolutionary time. KMT2F/G originated in unicellular species, KMT2A/B at the earliest stages of multicellularity, while KMT2C/D appeared later in the evolution of multicellularity.

We compared rates of somatic mutations between KMT2 genes of unicellular and multicellular origin in data from Pan-Cancer Analysis of Whole Genomes and the PanCancer Atlas for 33 tumour types. Across all cancers, a significant association exists between gene age and proportion of KMT2-mutated samples. Globally, rates of nonsense mutations and frameshift deletions increase significantly from older to younger KMT2 genes. This persists within 21 tumour cohorts, and notably in all eight tumour types with more than 400 samples and mutations in at least 10% of samples for any KMT2 gene. Additionally, at the protein level multicellular KMT2s have a higher proportion of predicted deleterious mutations in functional domains than the unicellular KMT2F/G.

More frequent and functionally important mutations to multicellular KMT2s likely inhibit their normal regulatory behaviour. Mutations to KMT2 genes likely alter network interactions such that multicellular regions are rewired, leading to a simultaneous loss in multicellular features and unchecked regulation of unicellular features (i.e., cancer hallmarks). Future analyses will investigate the transcription and binding profiles of KMT2 methyltransferases to examine functional differences by age.

P30: Aayushi Notra - Center for Cancer Biology

Single Cell Analysis of an in-vitro Timecourse Model to study the Development of Neuroblastoma

Aayushi Notra [1], Katherine Pillman [1], Nick Warnock [1], Gregory Goodall [1], Yeesim kew-Goodall[1]
[1] Center for Cancer Biology, SA, Australia

Introduction: Neuroblastoma is one the most common extracranial cancers originating from an atypical differentiation pattern of the developing sympathetic nervous system. It is a highly heterogenous cancer and remains one of the leading causes of oncology-related paediatric deaths.

Objective: To gain an in-depth molecular-level understanding of the factors driving abnormal differentiation, consequently leading to disease formation.

Method: Single-cell RNA-seq data from our in-vitro iPSC differentiation model was analysed using the Seurat R package. Clustering analysis was done to identify cell-type populations at each developmental stage (day 10,17 and 31). Model suitability was evaluated through integration with publicly available single-cell data from the developing fetal adrenal medulla. Subsequently, CellOracle analysis was performed using day 17 single cell and publicly available fetal adrenal medulla data. In silico perturbations targeted filtered transcription factors from the base gene regulatory network. Transcription factors

showing consistent perturbation effects across both datasets were identified and chosen for further analysis.

Results: Analysis of individual timepoints revealed highly heterogeneous cell populations. Integrated analysis uncovered shared cell populations across all timepoints and datasets, suggesting close replication of natural differentiation in our in-vitro iPSC model. CellOracle analysis identified 14 transcription factors potentially significant in sympathetic neuronal differentiation.

Future Direction: The candidate transcription factors identified from cell oracle analysis will undergo in vitro perturbation to validate their functional significance.

P33: Antonietta Salerno - University of New South Wales

Unveiling the effects of copper-chelation therapy in Neuroblastoma immune microenvironment with a multi-omics approach

Antonietta Salerno [1], Jourdin R.C. Rouaen [1] [2], Orazio Vittorio [1][2]
[1] School of Biomedical Sciences, Faculty of Medicine and Health, University of New South Wales, Sydney, NSW, Australia
[2] Children's Cancer Institute Australia, Lowy Cancer Research Centre, University of New South Wales, Sydney, NSW, Australia

Recent advancements in anti-GD2 immunotherapy have displayed promise in improving the survival rates of high-risk neuroblastoma patients. However, challenges persist owing to the immunosuppressive nature of the tumor microenvironment (TME). This study employs a cutting-edge multi-omics approach, combining Nanostring GeoMx Digital Spatial Profiling (DSP) and single-cell BD Rhapsody RNA profiling, to comprehensively investigate the effects of copper-chelators on neuroblastoma dynamics, as co-adjuvant agents for anti-GD2 therapy.
DSP analysis reveals that copper chelation induces an immune-permissive TME, marked by higher immune infiltration, upregulated pro-inflammatory pathways, and destabilized oncogenic signaling. Single-cell sequencing further unveils that copper chelation downregulates tumoral Mycn expression and its targets. Moreover, we observed activation of various immune cell types crucial for antibody-dependent cell-mediated cytotoxicity (ADCC). Notably, this therapy resulted in a five-fold increase of infiltrated neutrophils exhibiting an anti-tumor N1 phenotype, thus emerging as key players in copper-chelation-induced immune modulation.
Interestingly, by uncovering the ligand-receptor interactions among cell types in treatment and control, we observed a shift from the tumor to the neutrophils as the primary hub in the network of cell signaling interactions, favoring anti-tumor response. This study proposes a novel mechanism of action for copper chelation agents, able to hamper tumor progression by revitalizing neutrophils-directed anti-tumor immunity.

P36: Angus Shoppee - Monash University

Feature-directed Analysis of Splice Events: A novel tool to identify soluble receptor isoforms

Angus Shoppee [1], Meredith O'Keeffe [1]
[1] Monash University & Biomedicine Discovery Institute, Melbourne, Australia

Alternative splicing of transcripts allows individual genes to give rise to a variety of structurally and functionally distinct protein isoforms. There is an emerging understanding that many genes canonically encoding membrane-embedded receptors can also produce soluble receptor isoforms through alternative splicing. These soluble receptor isoforms can play regulatory roles in signalling networks, thereby

contributing to homeostasis or dysregulation, and some isoforms have been identified as potential prognostic biomarkers in cancer settings. However, current understanding of the expression and function of these molecules is limited, underpinned by a lack of tooling to facilitate the discovery and characterisation of soluble receptor isoforms at scale.

We present a novel algorithm developed to perform genome-wide, unbiased screening for splice events affecting receptor transmembrane regions in RNA sequencing data. These screen results can subsequently facilitate further bioinformatic or lab-based analysis. In a proof of concept, we apply this algorithm to publicly available data to demonstrate the variety of soluble receptor isoforms that exist at the transcript level. We show that in a cohort of melanoma patients, several receptors identified by our algorithm exhibit differential splicing rates between tumours that were subsequently responsive to immune checkpoint blockade versus those that were non-responsive.

We hope that the application of this tool to the vast existing body of RNA sequencing data will contribute to an improved understanding of receptor alternative splicing, assisting in the identification of new biomarkers and therapeutic targets.

P37: Montana Spiteri - The Walter and Eliza Hall Institute

Epigenetic and Transcriptional Insights from a Brain Cancer Peri-Operative Clinical Trial for Gliomas

Montana Spiteri [1,2]
[1] Personalised Oncology Division, WEHI, Australia, 3052
[2] Department of Medical Biology, University of Melbourne, Australia, 3052

Patients with glioma, a heterogeneous group of primary brain cancers, have an extremely poor 5-year survival of only 22%. The intricate heterogeneity, restrictive blood-brain barrier and complex tumour microenvironment associated with this cancer are obstacles that have stymied advancements of treatments that effectively target these brain tumours. Perioperative, neoadjuvant, phase 0 studies provide a unique opportunity to study drug pharmacokinetics, pharmacodynamics, and help to identify predictive biomarkers. Moreover, the longitudinal trial design is particularly beneficial for examining the direct effects of a new drug on critical biological outcomes. Here, we present epigenetic and transcriptional insights from matched pre-treatment and post-treatment patient samples from the first peri-operative clinical trial for glioma. Leveraging a multi-omics approach, we conducted methylation array (N = 5), single nuclei RNA sequencing (snRNA-seq; N = 5) and spatial transcriptomics (N = 4) of matched tissue samples collected at biopsy (time of diagnosis) and resection (on-treatment). We developed a comprehensive analysis pipeline for the snRNA-seq data, which includes an ambient RNA removal step using CellBender to enhance data quality. We also employed a variational autoencoder from the scVI toolkit to integrate data across samples, amplifying the robustness of our findings. Results show heterogeneous responses to treatment but with significant consistent changes in tumour composition and peripheral immune infiltration. Interestingly, clonality assessment using Numbat CNV haplotype caller revealed persistent loss of clonality elicited by the treatment across all patients. Overall, these insights demonstrate the feasibility of the peri-operative clinical trial approach to uncover novel mechanisms of response in patients with glioma enhancing understanding of glioma biology and treatment efficacy.

P38: Dingyin Sun - Hudson Institute of Medical Research

Identify potential oncogenic gene fusion events in pediatric CNS tumors and bone & soft tissue sarcomas

Dingyin Sun [1], Claire Sun [2], Ron Firestein [1,2]
[1] Hudson Institute of Medical Research, Victoria, Australia,
[2] Monash University, Victoria, Australia

Pediatric cancers differ from adult tumours, especially by their low mutational rate, meaning that the many targeted therapied developed for adult cancers cannot be directly uses in pediatric settings, which highlight the need for identifying more paediatric specific targets.
The initiation and progression of pediatric cancer involve diverse oncogenic mechanisms, with gene fusions being particularly prevalent in children compared to adults. These fusion genes arise from chromosomal rearrangements, resulting in the abnormal juxtaposition/fusion of two genes. This aberration can lead to the overexpression of oncogenes or the suppression of tumor suppressor genes (TSGs). Consequently, identifying gene fusions in pediatric cell lines holds promise for elucidating tumorigenesis mechanisms and uncovering potential therapeutic targets, as well as diagnostic or prognostic biomarkers. In my study, the focus will be on pediatric central nervous system (CNS) tumors and bone & soft tissue sarcomas. By concentrating on these specific areas, the aim is to identify oncogenic gene fusions within subcategories of pediatric CNS tumors and sarcomas, potentially revealing novel therapeutic avenues.

P40: Ashlee Thomson - University of Adelaide

Improving Accuracy and Reducing Bias in Genomic Alignment for B-cell Acute Lymphoblastic Leukemia Patients: A Pan-Genome Graph Approach

Ashlee Thomson [1,2], Jacqueline Rehn [1,2], Susan Heatley [1,2,3], Laura Eadie [1,2], Elyse Page [1,2], Caitlin Schutz [2], Barbara McClure [1,2], David Yeung [1,2,4,5], James Breen [6,7], Deborah White [1,2,3,5,8]
[1] Faculty of Health and Medical Sciences, University of Adelaide, Adelaide SA
[2] Blood Cancer Program, Precision Cancer Medicine Theme, South Australian Health & Medical Research Institute (SAHMRI), Adelaide SA
[3] Australian and New Zealand Children's Oncology Group (ANZCHOG), Clayton Victoria
[4] Haematology Department, Royal Adelaide Hospital and SA Pathology, Adelaide SA
[5] Australasian Leukaemia and Lymphoma Group (ALLG), Richmond Victoria
[6] Black Ochre Data Labs, Indigenous Genomics, Telethon Kids Institute, Adelaide SA
[7] James Curtin School of Medical Research, Australian National University, Canberra ACT
[8] Australian Genomics Health Alliance (AGHA), The Murdoch Children's Research Institute, Parkville Victoria

Introduction
A current limitation of traditional genome analysis workflows is the use of a linear reference genome for sequence alignment which is made mostly from European ancestry individuals and does not capture the unique genomic variation contained within global populations. Aligning sequencing reads to this linear reference genome requires bioinformatic tools to adjust for discrepancies between sequencing reads and the reference, potentially leading to "reference bias". This bias is intensified when handling data with substantial genomic variability, as seen in B-cell acute lymphoblastic leukemia (B-ALL) patients. We aimed to build a B-ALL-specific pan-genome graph that incorporates clinically significant genomic alterations, to enhance alignment precision and eliminate reference bias.
Methods
Using the GRCh38 reference genome as a base, we constructed a pan-genome graph using the Variation Graph (VG) toolkit that contains 7278 genomic alterations identified in B-ALL patients as listed in the Catalogue of Somatic Mutations In Cancer. We integrated genome annotations into the graph to facilitate exon and splice junction identification. To analyse mapping performance, we aligned Illumina paired-end

RNA sequencing data from 10 B-ALL patients to both the spliced pan-genome graph using VG's mapping function and to the GRCh38 linear reference genome using STAR aligner as a comparison.
Results
Preliminary mapping results demonstrate that on average, 98.35% of paired-end sequencing reads map to the pan-genome graph compared to an average of 95.47% which map to the linear reference. This difference represents ~4 million paired-end sequencing reads. We also see an increase in the reporting of genomic insertions and deletions larger than 10 base pairs.
Conclusion
This spliced pan-genome graph is effective and comparable to the linear reference genome when aligning B-ALL patient sequencing reads. Further analysis will involve expanding mapping metric analysis and investigating the specific variations detected.

P41: Lauren Tjoeka - Peter MacCallum Cancer Centre

Genotyping single-nuclei transcriptomes using custom probe-based technology (10x Genomics Single Cell Gene Expression Flex)

Lauren Tjoeka [1,2], Laura Twomey [1], Amanda Lee [1], David Yoannidis [1], Hugo Saunders [1], Sara Alaei [1], Tim Semple [1], Luciano Martelotto [3], AOCS [1], Alicia Oshlack [1,2,4], Elizabeth Christie [1,2]
[1] Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia
[2] Sir Peter MacCallum Department of Oncology, The University of Melbourne, Parkville, Victoria, Australia
[3] The South Australian Immunogenomics Cancer Institute, Faculty of Health and Medical Sciences, The University of Adelaide, Adelaide, South Australia, Australia
[4] School of Mathematics and Statistics, University of Melbourne, Parkville, VIC, Australia

Background:
Most women with BRCA1/2 mutations in their high grade serous ovarian cancer (HGSC) respond well to primary treatment. However, many will relapse within 2-3 years, acquire treatment resistance, and succumb to their disease. A common mechanism conferring resistance to platinum-based chemotherapy and PARP inhibitors are reversion mutations. Reversions are secondary somatic mutations in BRCA1/2 observed in up to 50% of patients with recurrent HGSC. BRCA1/2 reversions frequently appear to be subclonal and it is unknown how tumour cells without reversions persist through multiple lines of treatment despite having DNA repair deficiency conferred by germline or somatic mutations in BRCA1/2. This study aims to understand those resistance mechanisms, evident transcriptionally, within reversion-negative cells that co-exist with reversion-positive cells in patient tumour samples.

Method:
In order to identify the cells with and without reversions, genotyping of single nuclei or cells is required, to enable differential gene expression and gene set enrichment analysis of reversion-positive and reversion-negative nuclei. Genotyping of transcripts from 3' or 5' single-cell RNA sequencing is constrained to detection of variants positioned proximally to the 3' and 5' capture sites. Given that BRCA1/2 mutations: germline, somatic and reversions, are typically positioned >1 kb from transcript ends, 3' or 5' single-cell RNA sequencing technology are not feasible methods for genotyping of nuclei in this context. To circumvent this technical limitation, we are testing a probe-based single-nuclei RNA sequencing approach with custom-designed probes targeting known BRCA1/2 germline, somatic and reversion mutations coupled with a whole transcriptome read-out to delineate transcriptional differences between genotyped reversion-positive and reversion-negative nuclei (10x Genomics Single Cell Gene Expression Flex). Preliminary data anticipated May 2024.

Future Directions:

I anticipate identifying transcriptional changes occurring in cells without reversions that contribute to their ability to survive treatment and will investigate whether reversion-positive cells have additional resistance mechanisms with the findings subject to in vitro or orthogonal validation studies.

P43: Shunhan Yao - Medical College of Guangxi University

A Radiograph Dataset for Training Deep Learning Models for the Classification, Localization, and Segmentation of Primary Bone Tumors

Shunhan Yao[1,2], Qingfeng Chen[1], Jiangning Song[2]
[1]Medical College of Guangxi University, Nanning, China
[2] Biomedicine Discovery Institute and Monash Data Futures Institute, Monash University, Melbourne, Australia.

Primary malignant bone tumors are the third highest cause of cancer-related mortality among patients under the age of 20. X-ray scans is the primary tool for detecting bone tumors. However, due to the varied morphologies of bone tumors, few radiologists possess sufficient expertise to make a definitive diagnosis based on radiographs. With the recent advancement of deep learning algorithms, there is a surge of interest in computer-aided diagnosis of primary bone tumors. However, the lack of publicly available X-ray datasets for bone tumors has obstructed development in this field. To address this challenge, we established the BTXRD dataset in collaboration with multiple medical institutes. This dataset comprises 3,746 bone images (1,879 normal and 1,867 tumor), and each image has related clinical information and global labels. Furthermore, each tumor instance has its distinct mask and bounding box. The BTXRD dataset is freely available and can be used to support the development and evaluation of algorithms for diagnosing primary bone tumors.

P44: Hujun Yu - Peter MacCallum Cancer Centre

Characterisation of lymphangiogenic enhancers in development and cancer

Hujun Yu[1,4], Virginia Panara[6,7], Maria Rondon Galeano[1,4], Elizabeth Mason[1,4], Jovana Maksimovic[2,4], Katarzyna Koltowska[6,7], Shom Goel[3,4], Benjamin M. Hogan[1,2,5]
[1] Organogenesis and Cancer Program, Peter MacCallum Cancer Centre, Melbourne, VIC 3000, Australia
[2] Computational Biology Program, Peter MacCallum Cancer Centre, Parkville, VIC, 3010 Australia
[3] Peter MacCallum Cancer Centre, Parkville, VIC, 3010 Australia
[4] Sir Peter MacCallum Department of Oncology, University of Melbourne, Melbourne, VIC 3000, Australia
[5] Department of Anatomy and Physiology, University of Melbourne, Melbourne, VIC 3000, Australia
[6] Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden
[7] Beijer Gene and Neuro Laboratory and Science for Life Laboratories, Uppsala University, Uppsala, Sweden.

Metastasis is one of the hallmarks of cancer, and it is the primary cause of cancer morbidity and mortality. During metastasis, tumour promote lymphangiogenesis, the formation of lymphatic vasculature, and then access these new lymphatic vessels to spread to distant sites. The process of lymphangiogenesis is thought to be similar in cancer and in embryonic development, however this has never been actively tested in ideal model systems with high resolution genomic and cell biological approaches. It remains unclear whether there are cancer specific hallmarks and regulatory networks in tumour lymphangiogenesis.

The project will investigate the regulatory network of lymphangiogenesis in breast cancer, using mouse model with HER2 overexpression. To identify genes and gene regulatory elements, such as enhancers, involved in tumour lymphangiogenesis, we performed single-nuclei multiome sequencing to profile the transcriptomes and the chromatin accessibility in the tumour vasculature. By comparing the data on tumour lymphangiogenesis with developmental lymphangiogenesis (previously generated data in the lab),

we are able to determine the similarities, differences and mechanisms driving developmental lymphangiogenesis versus tumour lymphangiogenesis. After we identify key genes and enhancers involved in lymphangiogenesis in cancer, we will use functional genomics (eg. CRISPR) to perturb the candidates to test its impact on the formation of lymphatic vasculature induced by tumours and in development. This may potentially lead to the discovery of a novel therapeutic target that suppress tumour lymphangiogenesis in a specific manner and reduce metastasis in cancer patients.

P47: Huu Minh Liem Le - The University of Melbourne

Using long read RNA sequencing to identify differentially phasd protein domains

Huu Minh Liem Le [1], Nadia Davidson [2]
[1] The University of Melbourne
[2] Blood Cells and Blood Cancer, WEHI

Differential expression analysis aims to identify the transcriptional program that defines one group of cells with a distinct phenotype from another. This method has been pivotal to understanding the mechanisms that drive cancer, metastasis and treatment resistance. Traditionally, differential analysis involves measuring the gene expression under two or more conditions, and performing statistical tests to find the most significant changes between the expression profiles. Differential analysis can also examine differences at the level of transcript isoforms: Long read sequencing is growing in popularity for this application, as it overcomes the uncertainty in transcript quantification seen with short-read data. However, drawbacks exist in both gene-level (DGE) and transcript-level (DTE/DTU) approaches: DGE does not provide information on splicing events, and the large number of isoforms in a DTE/DTU reduces statistical power and is challenging to interpret. Inspired by this problem, our research group has developed an alternative feature for long read analysis — differential expression of phased protein domains. Protein domains are regions of proteins which are predicted to fold and function independently on one another. In this approach, the combination of protein domains within each transcript is identified and the counts of transcripts with the same domain combinations (DoCo) are aggregated. By performing differential analysis on DoCos using long read sequencing of cancer cell lines, we compared this approach with the conventional gene-level and transcript-level tests. Our approach conserved the distance relationship between samples, and resolved the data resolution conundrum while providing power close to DGE, but with an ability to resolve many biologically interpretable splicing changes. Additionally, the DoCo approach is applicable to single cell differential study. These findings suggest a potential use of DoCo in future differential analyses, notably in cancer contexts where cellular instability may give rise to various novel domain combinations.

Posters

P3: Sallam Alagawani - Swinburne University of Technology

Insights into Halogen-Induced Changes in 4-Anilinoquinazoline EGFR Inhibitors: A Computational Spectroscopic Study

Sallam Alagawani [1], Vladislav Vasilyev [2], Andrew H.A. Clayton [3], Feng Wang [1]
[1] Department of Chemistry and Biotechnology, School of Science, Computing and Engineering Technologies, Swinburne University of Technology, Melbourne, Victoria 3122, Australia
[2] National Computational Infrastructure, Australian National University, Canberra, ACT 0200, Australia

[3] Optical Sciences Centre, Department of Physics and Astronomy, School of Science, Computing and Engineering Technologies, Swinburne University of Technology, Melbourne, Victoria 3122, Australia

The epidermal growth factor receptor (EGFR) is a pivotal target in cancer therapy due to its significance within the tyrosine kinase family. EGFR inhibitors like AG-1478 and PD153035, featuring a 4-anilinoquinazoline moiety, have garnered global attention for their potent therapeutic activities. While pre-clinical studies have highlighted the significant impact of halogen substitution at the C3'-anilino position on drug potency, the underlying mechanism remains unclear. In this study, we investigated the influence of halogen substitution (X=H, F, Cl, Br, I) on the structure, properties, and spectroscopy of halogen-substituted 4-anilinoquinazoline tyrosine kinase inhibitors (TKIs) using time-dependent density-functional methods (TD-DFT) with the B3LYP functional. Our calculations revealed a planar conformation for the global minimum structure of each derivative, alongside a closely lying twisted conformation. Halogen substitution did not induce significant changes in the three-dimensional conformation of the TKIs, it did lead to noticeable alterations in electronic properties, such as dipole moment and spatial extent, which could impact interactions at the EGFR binding site. Spectral analysis unveiled a halogen-induced rightward of up to 7 nm redshift (range: 326-333 nm) in the absorbance maximum. The UV-visible spectra demonstrate that TKI-X compounds with higher potency generally exhibit absorption peaks at shorter wavelengths. This trend is exemplified by the peak wavelength of bromine at 326.71 nm, whereas hydrogen, possessing the lowest IC50 nM, displays a peak at 333.17 nm. This observation suggests a potential correlation between the potency of the compounds and their spectral properties. The introduction of different substituents (X) leads to varying changes in $\Delta\lambda max$ values, representing the discrepancy between the maximum absorption wavelengths of the calculated spectra for the calculated structures and those EGFR-DB. Certain TKIs-X exhibits positive delta values, including H, F, and I, while others such as Cl and Br, known for their higher potency, demonstrate negative delta values.

P5: Diego Chacon-Fajardo - Garvan Institute of Medical Research

Re-purposing non-oncology agent Itraconazole to target the dynamic cellular ecosystem of pancreatic cancer

Diego Chacon-Fajardo [1], Sean Porazinski [1], Jennifer Man [1], Howard Yim [2], Emad El-Omar [2], Anthony Joshua [1,3], Marina Pajic [1,3]
[1] Personalised Cancer Therapeutics Lab, Precision Cancer Medicine Program, Garvan Institute of Medical Research
[2] Microbiome Research Centre, St George and Sutherland Clinical School, UNSW, Sydney
[3] St Vincent's Clinical School, Faculty of Medicine, UNSW Sydney. Australia.

Pancreatic ductal adenocarcinoma (PDAC) has a 5-year survival of only 10% and persists as the third most common cause of cancer-related death in Western societies. Precision medicine is a likely future for all cancers treatment, specially those with high-mortality and molecularly heterogeneity such as Pancreatic Cancer (PC). Bioinformatic analyses have provided new insights of how novel molecular-guided therapies can modulate signalling within tumours and ecosystem. In this study, single-cell RNAseq is used to delineate the complex mechanisms by which a clinically-used, orally available anti-fungal agent itraconazole improves the overall anti-tumour response in an advanced in vivo PC model (KPC).
scRNAseq analyses have revealed a significant modulation across the tumour microenvironment post-itraconazole therapy. In cancer-associated fibroblasts (CAFs), downregulation of the pro-tumourigenic Cd105+ signature. Further, antigen-presenting CAFs (apCAFs) exhibited decreased expression of H2-Aa and Cd74, which act as decoys for corresponding receptors on Cd4+/Cd8+ T cells,

impairing their activation. Immunofluorescence further demonstrated decreased deposition of collagen and altered matrix remodelling within itraconazole-treated tumours.
In the immune environment, itraconazole treatment led to positive changes within macrophages, including enrichment of M1-like pro-inflammatory phenotype, confirmed via immunofluorescence, and further associated with positive signalling via Cxcl9-Cxcr3 and Tnf1rs-1a/1b ligand-receptors. In lymphoid cells, scRNAseq revealed dampening of regulatory T cells homing as well as functional potency signals associated with increased Cd8+ T-cell infiltration post-treatment, suggesting an improved response to immune checkpoint blockade.
Finally, analysis of epithelial-mesenchymal transition (EMT) in tumour subsets showed overall decrease in their mesenchymal phenotype, enhancing migration and modulation of the immune microenvironment. Furthermore, upregulation of Sepp1, Cd74 and Clu in specific tumour subtypes suggests favourable prognosis and presence of tumour infiltrating lymphocytes (TILs), whereas repression of Lcn2 and Ecm1 indicates downregulation of pro-metastatic mediators.
These integrated bioinformatic and molecular studies provide scientific rationale for the development of itraconazole and immunotherapy combination in PC.

P7: Ben Curran - Children's Cancer Institute

Pan-cancer Hierachical Classification of pediatric cancers using methylation data

Ben Curran [1]
[1] Children's Cancer Institute of Australia

Accurate and early diagnosis is essential for the management of pediatric cancers. In the ZERO childhood cancer study, whole genome sequencing, transcriptomics and methylation arrays are used to identify a cancer and the genetic variants.

In addition to tissue specific differences in methylation, changes in methylation can accompany tumor development. This can lead to differences in DNA methylation patterns between cells from different tissues and between normal and tumor cells from different tissues.

As part of the curation of pediatric cancer samples, data from methylation arrays is used to identify a cancer type and sub-type for many CNS and Sarcoma tumors, using a classifier trained on data from EPIC arrays. We have built a new multi-level classifier built using data from both v1 and v2 of the EPIC 850K array.

Patients enrolled in Zero have diagnoses that can be mapped to the ontology of cancer described by the World Health Organizations International Classification of Diseases for Oncology.

Individual classifiers are built for each node in the tree. A one-to-many differential methylation analysis between all children of a node in the tree is undertaken to identify a reduced set of features. A cross validated grid search was used to select the parameters for training a Random Forest model on this feature set. This set of parameters is then used to train a Random Forest model and validated on data from a hold-out set of samples.

The use of a new set of features and model parameters at each node in the tree decomposes a multi-class classification problem having hundreds of classes to a set of smaller, self-contained classification problems.

P8: Gunjan Dixit - Peter MacCallum Cancer Centre

With More Data Comes More Responsibility: Large scale single cell analysis of > 130 clinical samples

Gunjan Dixit [1,2], George Howitt [1,2], Victoria Streeton-Cook [1,2], Shivanthan Shanthikumar [3,4], Melanie Neeland [3], Jovana Maksimovic* [1,2] and Alicia Oshlack* [1,2]
[1] Computational Biology Program, Peter MacCallum Cancer Centre, Parkville, VIC, Australia
[2] Sir Peter MacCallum Department of Oncology, University of Melbourne, Parkville, VIC, Australia
[3] Respiratory Diseases, Murdoch Children's Research Institute, Parkville, VIC, Australia
[4] Respiratory and Sleep Medicine, Royal Children's Hospital, Parkville, VIC, Australia

This study presents a large-scale single-cell analysis using Flex Single-Cell RNA-Seq (scRNA-seq) technology to process 136 clinical samples from 56 healthy participants across three paediatric age groups: preschool (1-5 years), early childhood (6-11 years), and adolescence (12-16 years). Derived from five different upper airway tissues—nasal brushings, tonsils, adenoids, bronchial brushings, and bronchoalveolar lavage (BAL)—the samples were meticulously processed using the new 10X FLEX capture technology, facilitating sample storage and batching which is crucial given the opportunistic nature of clinical sample collection. This process yielded a comprehensive dataset of 796,759 cells (CellRanger), subsequently reduced to 440,047 high-quality cells after rigorous data cleaning to remove low-library sized cells, ambient RNA, doublets/multiplets and mitochondrial outliers.

The cell types were annotated by integrating harmonized data from two adult reference atlases (Lung Cell Atlas and Tonsil Atlas), enriched with insights gained through manual inspection of cell-type-specific markers. This analysis revealed distinct cellular profiles across different tissues: tonsils and adenoids predominantly contained B cells due to their lymphoid nature, while nasal and bronchial brushings were characterized by a higher presence of epithelial cells. BAL samples were mainly composed of macrophages. Advanced linear modelling further delineated how cell-type compositions shift with age.

The complexity of handling such extensive data requires strategic decision-making at various stages to achieve consensus and comprehend technical aspects arising not only from the new FLEX technology but also from different tissues. This atlas not only enhances our understanding of paediatric airway biology but also sets the stage for future research and clinical applications aimed at improving paediatric respiratory health.

P9: Timothee Froute - Peter MacCallum Cancer Centre

CaSP Project implementation: how to set up automatic data exchange between institutes for a clinical/research hybrid project?

Timothee Froute [1]
[1] Peter MacCallum Cancer Centre, Victoria, Australia

Omico's Cancer Screening Program (CaSP) provides free, large scale Comprehensive Genomic Profiling (CGP) to Australian cancer patients with incurable and advanced cancers, with the aim of identifying patients with potential matches to targeted therapy clinical trials. CaSP requires CGP to be carried out by NATA accredited pathology laboratories across Australia, with each lab receiving electronic requests and electronically returning pre-analytical status and progress data, as well as final CGP reports and sequence data. The CaSP protocol challenges normal laboratory specimen registration, processing and reporting procedures by mandating not one, but multiple timepoints for external data transfer to the Omico Application Programming Interface (API). PeterMac's Laboratory Information System (LIS) lacks a "ready

to use" feature capable of such automated discontinuous external data transfers. Our solution was to introduce an elaborate architecture leveraging traditional workflows. This involved automated emailing of electronic requests as PDF forms to our specimen registration team, and scheduled LIS data extractions to a custom MySQL database, allowing for more flexible data manipulation and exchange with the Omico API. As CaSP volumes increased, this architecture has allowed us to respond "on the fly" to resolve bugs and adapt to edge cases while maintain a high turnaround CGP service. This presentation will discuss how we set up such a system, how it may be upgraded in the future, and the different problems we encountered along the way.

P10: Milad Ghomlaghi - Monash University

Decoding Network Topologies Driving Adaptive Resistance to Targeted Cancer Therapies

Karina Islas Rios[1,2], Milad Ghomlaghi [1, 2], Sung-Young Shin[1, 2], and Lan K. Nguyen[1,2]
[1] Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia
[2] Biomedicine Discovery Institute, Monash University, Clayton, Victoria 3800, Australia

Targeted therapies represent a significant advancement in cancer treatment by disrupting function of cancer-specific molecules pivotal cell proliferation. However, the formidable challenge of drug resistance undermines their long-term efficacy. Despite their initial success, the emergence of drug resistance poses a significant challenge, limiting their long-term efficacy. Adaptive resistance, characterized by the rapid evolution of cancer cells to evade targeted therapies through intricate signalling networks, represents a challenging obstacle against effective treatment.

This study employs a multidisciplinary approach integrating advanced mathematical modelling techniques and introduces NetScan, an innovative web-based tool, to dissect the intricate network architectures recapitulating human cellular signalling pathways contributing adaptive resistance. Our methodology combines mathematical frameworks such as ordinary differential equations-based modelling, statistical analysis, and hierarchical clustering, all validated through biological experiments.

Through unbiased computational simulations and systematic analysis of over 16,000 network structures, we elucidate the underlying mechanisms fueling adaptive drug resistance against single-agent therapy. Contrary to conventional understanding, our findings unveil a diverse spectrum of network patterns contributing to adaptive resistance, extending beyond classical feedback loops. Furthermore, we identify network structures inducing hyper-activation of target proteins, potentially exacerbating cancer progression.
To pinpoint these critical network architectures within the human protein signallome, we introduce NetScan, a novel tool enabling real-time identification and visualization of defined network structures within cellular signalling networks. Validation using high-throughput experimental data confirms the emergence of resistance in proteins forming specific network structures predicted by mathematical models to mediate adaptive resistance.

In conclusion, our study significantly enhances our understanding of network-mediated drug resistance mechanisms in cancer and provide novel insights for the development of more effective therapeutic strategies.

P12: David Goode - Peter MacCallum Cancer Centre

GEPETO – A toolkit for identifying transcriptional programs driving clonal evolution from barcoded single-cell RNA-sequencing data

David Goode [1,2], Alex Casar Berazaluce[1,2], Brenda Sanchez Pichardo [3], Howard Bondell[2]
[1] Peter MacCallum Cancer Centre, Victoria, Australia
[2] The University of Melbourne, Victoria, Australia
[3] National Autonomous University of Mexico, Morelos, Mexico

Combining barcoded transcripts into single-cell RNA-sequencing (barcoded scRNA-seq) links the frequencies of clonal sub-populations to their gene expression profiles, revolutionising our understanding of how changes in gene regulation drive clonal evolution. Barcoded scRNA-seq can reveal how the activity of specific sets of co-expressed genes, known as Gene Expression Programs (GEPs), enhance or diminish the fitness of certain clones. This is especially important in cancer, where oncogenic and drug resistance GEPs confer survival advantages to clonal subpopulations of tumour cells, fuelling disease progression. However, existing single-cell analysis methods have many limitations when it comes to identifying and characterizing GEPs tied to clonal fitness from barcoded scRNA-seq data.

We introduce the Gene Expression Program Extraction Toolkit (GEPETO), to extract GEPs associated with user-defined fitness-associated Property Of Interest (POI), such as clone size or fraction of dividing cells, from barcoded scRNA-seq data. GEPETO employs Random Forest machine learning models, leveraging transcriptional profiles of barcoded clones to identify genes whose expression predicts clonal values of a user-defined POI. The most predictive genes are grouped into GEPs by clustering to find sets of genes commonly co-expressed together in the same cells. To evaluate validity of putative gene expression programs, GEPETO implements a permutation test to control false discovery rate, benchmarked on scRNA-seq data simulated by SPLATTER.

In applications to experimental data, GEPETO successfully uncovered known GEPs tied to differences in proliferation rates between clones. GEPETO identified multiple GEPs associated with increased clone size in both cell line and mouse models of Acute Myeloid Leukemia (AML). Activity of GEPs found by GEPETO could accurately predict AML clone size based on transcriptional profiles, with R2 values up to 0.8. One of the top GEPs included Slpi, a gene previously shown to enhance in vivo growth of AML cells.

Our findings shed light on how diverse programs related to translation, stemness, mRNA transport, and stress response contribute to clonal fitness and disease progression in AML. They demonstrate how GEPETO can reveal novel connections between gene co-expression and clonal fitness, with applicability to many contexts where clonal evolution and competition are observed.

P14: Anthony Hart - Monash University

Untangling the Triad: Network Topology, Drug Dose, and Drug Response Dynamics

Anthony Hart [1], Sungyoung Shin [1], Lan K. Nguyen [1]
[1] Cancer Program, Biomedicine Discovery Institute, Monash University, Victoria, Australia

Cancer therapeutics often target specific proteins that reside in complex, non-linear networks. The complexity and adaptability of these networks can undermine the ability of drugs to potently inhibit protein activity for sufficient durations. Moreover, the ability of protein activity to rebound following drug treatment has been shown to facilitate the development of drug resistance. Understanding protein responses over

time following drug treatment is crucial, yet the impact of drug dosage on drug response, while equally important, remains underexplored.

In this study, we systematically generated a series of network motifs and interrogated them using meta-dynamic network modelling, an analytical technique developed in our lab, to elucidate the relationship between network topology, drug dose, and drug response. Our findings support the ubiquitous paradigm of maximum tolerable dose for a large number of the network topologies we explored, as increasing drug dose generally led to a decrease in output proteins. However, our results also challenge this paradigm by identifying network topologies in which increasing drug dose beyond a threshold provides no appreciable increase in efficacy, suggesting that higher doses may not always be beneficial. Moreover, we demonstrate the existence of network topologies that can facilitate a protein-dynamic phenomenon we term "super-rebound," wherein increasing drug doses cause greater acute inhibition but stronger long-term target reactivation, exceeding pre-drug levels. This highlights the potential for drug desensitisation and underscores the importance of considering network topology and dose-dependent responses in drug development.

Overall, our findings suggest that optimal dosing strategies may vary depending on the specific network topology of the targeted proteins. Furthermore, our work emphasises the need of a systems-level perspective when understanding complex dynamics of drug responses, highlighting the need for more integrative approaches in cancer therapeutics research.

P15: Jingni He - University of Melbourne

Enhancing Disease Risk Gene Discovery by Integrating Transcription Factor-Linked Trans-located Variants into Transcriptome-Wide Association Analyses

Jingni He[1][2], Wanqing Wen[3], Jie Ping[3], Qing Li[2] , Zhishan Chen[3], Deshan Perera[2], Xiang Shu[4], Jirong Long[3], Qiuyin Cai[3], Xiao-Ou Shu[3], Wei Zheng[3], Quan Long[2], Xingyi Guo[3]
[1] Department of Clinical Pathology, University of Melbourne, VIC, AU
[2] Department of Biochemistry & Molecular Biology, University of Calgary, Calgary, Canada
[3] Division of Epidemiology, Vanderbilt University School of Medicine, Nashville, TN, USA
[4] Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Transcriptome-wide association studies (TWAS) have been successful in identifying disease susceptibility genes by integrating cis-variants predicted gene expression with genome-wide association studies (GWAS) data, while trans-located variants for predicting gene expression remain largely unexplored. Here, we introduce transTF-TWAS, which includes transcription factor (TF)-linked trans-located variants to enhance model building. Using data from the Genotype-Tissue Expression project, we predict alternative splicing and gene expression and applied these models to large GWAS datasets for breast, prostate, and lung cancers. Our analysis revealed 887 putative cancer susceptibility genes, including 465 in regions not yet reported by previous GWAS and 137 in known GWAS loci but not yet reported previously, at Bonferroni-corrected $P < 0.05$. we demonstrate that transTF-TWAS outperforms other existing TWAS approaches in both constructing gene prediction models and identifying disease-associated genes, as evidenced by simulations and real data analysis. These results have shed new light on several genetically driven key regulators and their associated regulatory networks underlying disease susceptibility.

P19: Chol-Hee Jung - University of Melbourne

Characterisation of mitiome genes, a complete set of nuclear encoded mitochondrial genes.

Chol-Hee Jung [1], Grace Hall [1], Bernard J. Pope [1,2], Christopher M. Hovens [2,3,4]
[1] Melbourne Bioinformatics, The University of Melbourne, Melbourne, VIC, Australia
[2] Department of Surgery, University of Melbourne, 5th Floor Clinical Sciences Building, Royal Melbourne Hospital, Grattan Street, Parkville, VIC, 3050, Australia
[3] Department of Urology, Royal Melbourne Hospital, Melbourne, VIC, 3050, Australia
[4] University of Melbourne Centre for Cancer Research, Victorian Comprehensive Cancer Centre, Parkville, VIC, 3050, Australia.

Mitochondria are a crucial component of eukaryotic cells, which serve as a cellular power house, generating adenosine triphosphate molecules by the oxidative phosphorylation system (OXPHOS). For humans, the mitochondrial genome is responsible for encoding 13 core subunits of OXPHOS pathway proteins, but the function of mitochondria requires approximately 1100 other protein coding genes which are encoded in the nuclear genome.
Despite the crucial function of those 1100 genes for cells, the variant status of this gene set in conditions such as cancer have been limited to only a subset of those genes in multiple studies, and they have been referred to as various names. Here, we propose the term ‘mitiome’ for the complete set of nucleus-encoded human mitochondrial protein coding genes and present an analysis of the characteristics of mitiome in terms of expression and mutational profiles across human cancers, including prostate cancer.
The human mitiome genes are evenly distributed across all chromosomes except for the Y chromosome. Their total lengths and coding region lengths are slightly shorter compared to other protein-coding genes. Mitiome genes are consistently highly expressed in both tumour and normal tissues. Their expression is higher than other protein coding genes by an order of magnitude but only slightly lower than other nuclear-encoded house-keeping genes. However, the tumour mutational burden of mitiome is the lowest when compared to house-keeping genes or the rest of the protein-coding genes, implying that they are less tolerant to somatic mutations.
This study is the first to profile mitiome as a functional entity in cells, which are not only of high importance for the function of mitochondria, but also exhibit distinct characteristics.

P20: David Kaplan - Peter MacCallum Cancer Centre

Cell spatial motifs in prostate cancer

David I Kaplan [1,2], Anna S Trigos [1,2,3]
[1] Trigos Laboratory, Mulit-omic Evolution and Ecosystems, Computational Biology Program, Laboratory Research, Peter MacCallum Cancer Centre, Victoria, Australia
[2] The Sir Peter MacCallum Department of Oncology, The University of Melbourne, Victoria, Australia
[3] Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash Medicine Nursing and Health Sciences, Monash University, Victoria, Australia

The current theory of tumour evolution states that following the emergence of a clonal population of tumour cells in an organ, further genetic – or epigenetic – mutations will develop, giving rise to sub-clonal populations, producing a heterogeneous population of tumour cells. Faced with selective pressures – such as immune response, nutrient restriction or therapeutic treatment – some of these subclones may be eliminated while resistant populations will survive and proliferate, driving the evolution of the tumour. This theory does not adequately capture the complex picture of tumour development, as it overlooks the diverse interactions between distinct populations of cells in a tumour that enable tumour survival. Here, we hypothesise that an evolutionary unit in cancer development is the tumour ecosystem. We anticipate

that this will be manifested as specific spatial configurations of tumour populations, clones and cells of the microenvironment.

To investigate this, we utilized spatial transcriptomics data from various stages of prostate cancer, from primary to metastatic disease, including metastatic lesions from different organs. We identified that transcriptional signatures of prostate cancer cell plasticity are distributed in discrete regions in the tissue, and that patterns of colocalization and spatial segregation were recurrent and predictable across samples, suggesting specific tumour cell motifs provide survival advantage. We also identified clonal populations across samples and narrowed down recurrent patterns of spread of clones within tissues, pointing towards the existence of tissue-level patterns of evolution.

Overall, our approach enables the study of cancer evolution at the ecosystem level. Our results point towards the existence of cell spatial motifs in cancer, which are unique to disease stage and patient, extending our current understanding of cancer evolution.

P21: Shivakumar Keerthikumar - Peter MacCallum Cancer Centre

Investigating The Effect of Bipolar Androgen Therapy (BAT) on Patient Derived Xenografts (PDX) of Castration Resistant Prostate Cancer using Single Cell Multiome Sequencing

Shivakumar Keerthikumar [1][2][3], Rosalia Quezada Urban [3], Georgia Cuffe [3], Melbourne Urological Research Alliance (MURAL) [3], Peter Lau [5], David L Goode [1][2], Renea Taylor [2][3][4], Gail Risbridger [1][2][3], Anthony Joshua [6], Mitchell Lawrence [1][2][3]
[1] Cancer Research Division, Peter MacCallum Cancer Centre, Melbourne, VIC, Australia
[2] Sir Peter MacCallum Department of Oncology, The University of Melbourne, Parkville, VIC, Australia
[3] Department of Anatomy and Developmental Biology, Monash Biomedicine Discovery Institute Cancer Program, Monash University, Clayton, VIC, Australia
[4] Department of Physiology, Monash Biomedicine Discovery Institute Cancer Program, Monash University, Clayton, VIC, Australia
[5] Australian Genome Research Facility, Melbourne, Australia
[6] The Kinghorn Cancer Centre, St Vincent's Hospital, Sydney, NSW, Australia

Bipolar androgen therapy (BAT) is a potential new treatment for patients whose prostate cancer has become resistant to standard hormone-blocking therapy. BAT involves cycling patients between low (castrate) and high (supraphysiological) levels of testosterone to control tumour growth. This is based on the concept that prostate cancer cells have a biphasic response to androgen levels. Clinical trials show ~30% of tumours respond to BAT, so there is growing interest in determining the features of responsive tumours. This study aimed to use patient-derived xenografts (PDXs) to compare the efficacy of BAT in different cases of metastatic advanced prostate cancer and to profile these tumours using the 10x single cell multiome- Assay for Transposase Accessible Chromatin (ATAC) and Gene Expression (GEX) technique.
We used PDXs from the Melbourne Urological Research Alliance spanning diverse forms of prostate cancer from patients who progressed on current treatments, including potent AR signalling inhibitors, chemotherapy and radioligand therapy. PDXs were treated with vehicle or BAT (fortnightly intramuscular injections, 1mg testosterone cypionate) for 6-weeks, or until tumours reached an ethical volume. Samples were further processed at AGRF for nuclei isolation using the 10x Chromium Next GEM Single Cell Multiome assay.
Three of seven of PDXs (42%) responded to BAT with a significant reduction in tumour volume compared to vehicle control. Using the single cell multiome approach, ~36000 cells were captured. The median high-quality fragments per cell was ~12000 for ATAC and the median genes per cell was ~2800 for GEX. Analyses of the dataset demonstrate widespread transcriptional changes in both responsive and non-responsive PDXs after acute BAT treatment. There were minimal acute changes in chromatin

between control and treatment groups, suggesting the importance of differential expression of existing AR-regulated genes. In summary, this study is revealing the inter- and intra-tumoural differences in the response of patient-derived models to BAT.

P25: Jiadong Mao - University of Melbourne

Φ-space: Integrative analysis of single-cell and spatial transcriptomics data in phenotype spaces

Jiadong Mao [1], Kim-Anh Lê Cao [1]
[1] Melbourne Integrative Genomics, School of Mathematics and Statistics, The University of Melbourne

Can you use an atlas of healthy lung samples to annotate cancerous lung cells? You could. But conventional cell typing methods only assign known cell type labels to (new) query cells. This leads to poor interpretability when the query cells contain out-of-reference cell states. Spatial transcriptomics (ST) technologies are now widely utilised in cancer research, but ST platforms have vastly different spatial resolutions and they are still fast evolving. We desperately need new computational tools to robustly transfer knowledge accumulated in bulk and scRNA-seq data to ST data from different platforms in a biologically meaningful way.

To address these challenges, we have developed Φ-space (Phi-space) for integrating bulk RNA-seq, single-cell multiomics and ST data in phenotype spaces. We define a phenotype space as a low dimensional space that uses bulk or single-cell references to phenotype query single cell on a continuous scale. Our phenotype space can include various type of information from the reference data, including cell type, culture type and patient disease condition.

We have already shown Φ-space's ability to successfully integrate bulk RNA-seq and single-cell multiomics data. Here we will illustrate the power of Φ-space in providing biological insights into cancerous ST samples by utilising scRNA-seq references containing only healthy samples. Our first case study revealed the subtle and heterogenous phenotypic signatures of Non-Small Cell Lung Cancer cells from multiple ST (NanoString CosMx) samples. In our second case study, we utilised three scRNA-seq references to jointly annotate a ST (BGI Stereo-seq) dataset from a mouse spleen sample with barcoded Acute Myeloid Leukaemia (AML) cells. Our results revealed subtle but significant grouping of AML clones in both the tissue and the phenotype spaces.

P26: Loza Martin - University of Tokyo

Human housekeeping cis-regulatory elements and their involvement in tumor suppression

Loza Martin [1], Alexis Vandenbon [2], and Kenta Nakai [1]
[1] The Institute of Medical Science, the University of Tokyo
[2] Institute for Life and Medical Sciences, Kyoto University

This study explores housekeeping cis-regulatory elements (HK-CREs) in the human genome. Through extensive multiomics analysis, we highlight the unique epigenetic features of these elements and explore their importance in vital biological processes beyond the regulation of housekeeping genes. Notably, we observe reduced activity of HK-CREs in cancer cells, particularly those near the telomere region of chromosome 19 and associated with zinc finger genes. Further analysis, including cancer samples, suggests the importance of these genes in housekeeping tumor suppressor processes. Overall, our findings highlight the importance of HK-CREs within the cells for preserving cellular integrity and stability.

P31: Andrew Pattison - Monash University

Spatially coordinated immune responses in colorectal cancer

Andrew D Pattison [1, 2, 5], Rebekah M Engel [1, 2, 3], Liam Skinner [1, 4], David Nickless [6], Christine Georges [3], Paul McMurrick [3], Thierry Jarde [1, 2], Lochlan Fennell [1, 4], Helen E Abud [1, 2]
[1] Department of Anatomy and Developmental Biology, Monash University, Clayton, 3800, Australia
[2] Stem cells and Developmental Program, Monash Biomedicine Discovery Institute, Monash University, Clayton, 3800, Australia
[3] Cabrini Monash University Department of Surgery, Cabrini Hospital, Malvern, Australia
[4] Aging and Cancer Laboratory, Monash Biomedicine Discovery Institute, Monash University, Clayton, 3800, Australia
[5] Monash Genomics and Bioinformatics Platform, Monash University, Clayton, 3800, Australia
[6] Anatomical Pathology Department, Cabrini Pathology, Cabrini Hospital, Malvern, VIC 3144, Melbourne, Australia

Colorectal cancer (CRC) is the third most common cancer globally with the second-highest mortality rate. CRC can be grouped into DNA mismatch repair proficient (MMRp) and deficient (MMRd), with only ~50% of MMRd tumours responding to immunotherapy. It is not fully understood why only some MMRd tumours respond to immunotherapy, although the presence of CXCL13+ T cells and tertiary lymphoid structures (TLS) have been suggested as favourable prognostic markers. To better understand the tumour-immune microenvironment, we measured the expression of ~1000 genes in 846,469 cells from diverse tumour/normal tissues (23 samples, 15 donors) using the CosMx Spatial Molecular Imager. QC, cell type annotation and spatial niche identification were performed using a custom bioinformatic pipeline. Niches reflected known tissue structures including normal colonic crypts and tumour masses, as well as neutrophil rich (NR) and TLS niches. We characterised the cells within these niches by pseudobulk differential gene expression, revealing striking changes in chemokine signalling, T cell infiltration and myeloid cell polarisation. To obtain a larger sample set, we leveraged public single cell RNA-Seq data. We found cells from samples with more neutrophils had distinct innate inflammation related gene expression changes and the tumour cells themselves had an altered stemness profile. Samples with relatively more TLS (as defined by a spatial TLS signature) showed greater interferon signalling and had significantly increased numbers of proliferating B cells. This work advances the understanding of spatial organisation the CRC immune response and provides niche-specific targets for validation and further investigation.

P32: Bernard Pope - University of Melbourne

Ultrasensitive Detection of Circulating Tumour DNA enriches for Patients with a Greater Risk of Recurrence of Clinically Localised Prostate Cancer

Bernard Pope [1], Gahee Park [2], Edmund Lau [1], Jelena Belic [2], Radoslaw Lach [2], Anne George [2], Patrick McCoy [1], Anne Nguyen [1], Corrina Grima [1], Bethany Campbell [1], Chol-hee Jung [1], Emma-Jane Ditter [2], Hui Zhao [2], The Pan Prostate Cancer Group (PPCG), David C. Wedge [3], Daniel S. Brewer [4], Andy G. Lynch [5], Harveer Dev [2], Vincent J. Gnanpragasam [2], Nitzan Rosenfeld [2], Christopher M. Hovens [1], Niall M. Corcoran [1], Charles E. Massie [2]
[1] The University of Melbourne
[2] University of Cambridge
[3] University of Manchester,
4] University of East Anglia
[5] University of St. Andrews

Circulating tumour DNA (ctDNA) has demonstrated utility for diagnostic and prognostic applications in many cancer types. However, previous methods have proven less effective in localised prostate cancer. We assessed the limits of detection of ctDNA in this context using the high-sensitivity INVAR method, and tested the hypothesis that ctDNA detection is associated with high risk disease.

A total of 128 individuals with clinically localised prostate cancer were selected, and 27 healthy individuals were included as negative controls. Plasma cell-free DNA (cfDNA) samples from cases and controls were profiled using custom targeted sequencing panels, with saturating coverage of patient-specific mutations identified by WGS. We assessed ctDNA detection in cases using the highly sensitive INVAR pipeline, leveraging consensus sequencing alignments, background error modelling and integration of signals across thousands of patient-specific variants. Biochemical recurrence and metastasis-free survival curves were used to assess the relationship between ctDNA detection and disease progression.

We combined signals across the maximum number of genome-wide patient specific mutations and leveraged an established analysis pipeline that corrects for background error rates and calculates a global integrated mutant allele fraction. ctDNA was detected in 16% of cases. Furthermore, ctDNA detection was significantly associated with biochemical recurrence and shorter metastasis-free survival.

Our study provides clear insights into the required analytical sensitivity and potential utility of ctDNA analysis in localised prostate cancer. This raises the potential for including ctDNA detection as an additional tool for patient stratification in future neo/adjuvant treatment trials.

P34: Nicole Saw - Peter MacCallum Cancer Centre

A novel probe-based sequencing approach enables high-resolution single-cell analysis within archival patient FFPE samples

Nicole YL Saw [1], Sara Roth [1,2], Avraham Travers [1,2], Shienny Sampurno [1], Kelly M. Ramsbottom [1], Sara Alaei [1], Timothy Semple [1], Jeanne Tie [1,2], Alexander Heriot [1,2], David L. Goode [1,2], Ian A. Parish [1,2]
[1] Peter MacCallum Cancer Centre
[2] Sir Peter MacCallum Department of Oncology, The University of Melbourne

Single-cell sequencing-based methods have revolutionised biology by allowing high resolution transcriptional profiling of complex cell populations. FFPE (formalin-fixed, paraffin-embedded] tissue archives remain the largest bank of patient samples, however they have not previously been amenable to scRNAseq analysis due to RNA fragmentation and compromised cell integrity. Recently, a probe-based technique called single-nuclei pathology sequencing (snPATHOseq] was reported to circumvent these issues and enabled sensitive and efficient transcriptome profiling of single nuclei extracted from FFPE samples. We benchmarked snPATHOseq against matched fresh sample scRNAseq data using tissue from the same patient, and demonstrate higher cell yields by snPATHOseq alongside sensitive detection of key cell populations. In addition to recovery of both immune and tumour cell populations, we additionally were able to detect rare cell populations, such as lymph node stromal cells. Overall, these data demonstrate that snPATHOseq represents a powerful and sensitive approach for analysing banked archival patient tissues, with performance that is comparable to conventional scRNAseq approaches.

P35: Katherine Scull - Monash University

mirrorCheck: an R package facilitating informed use of the lfcShrink algorithm with DESeq2 for differential gene expression analysis of clinical samples

Katherine Elise Scull [1], Kiarash Behrouzfar [1], Paul Yeh [1,2]
[1] Blood Cancer Biomarkers Lab, Department of Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, VIC
[2] Monash Haematology, Clayton, VIC

The lfcShrink function in the DESeq2 package for differential gene expression (DGE) analysis is a sophisticated tool aimed at reducing noise from low abundance genes without resorting to arbitrary filters. When comparing multiple groups, the documentation suggests creating a new factor specifying one level per group, allowing a simple design formula with easily interpretable results. However, challenges can arise, especially when analysing clinical samples. Firstly, when using the “apeglm” algorithm for lfcShrink, one can only extract pairwise comparisons (termed contrasts) against the reference level. For other contrasts, the factor must be re-levelled so that one of the desired groups forms the reference. Thus, extracting all the desired contrasts can be laborious and error-prone. Secondly, when using RNAseq data from the bone marrow of patients with haematological malignancies, we noted that the list of differentially expressed genes from contrasts sometimes varied dramatically depending on which group is the reference level. Specifically, although we expect the two results sets to mirror each other, we found low consensus between the genes lists, with a bias towards up-regulated genes in both results. This called the reliability of the identified genes into question; however, filtering out low abundance genes prior to analysis substantially improved consensus for many affected contrasts. To facilitate the use of lfcShrink and help users assess whether they need to pre-filter data, we present mirrorCheck, an R package including two functions; each addresses one of the problems identified above. run.DESeq.all.contrasts automates extraction of all possible contrasts and outputs reports including Volcano plots and gene tables. compare.reciprocal.contrasts produces diagnostic plots, which concisely convey the reliability of each comparison, and outputs ‘consensus’ gene lists. We predict that mirrorCheck may be particularly useful for DGE analyses of cancer samples in translational research, which can be complicated by high biological variability and low sample numbers.

P39: Ryosuke Suzuki - Children’s Cancer Institute

Flexible, lightweight and customizable pipeline for high-throughput microscopy image analysis leveraging unbiased features extracted from deep learning models

Ryosuke Suzuki [1], Antoine de Weck [1] [2]
[1] Computational Drug Discovery Biology, Children’s Cancer Institute, Australia
[2] School of Clinical Medicine, UNSW Medicine

Cell morphology analysis based on microscopy imaging can provide profound insights into the effects of drug treatments, e.g. to quantify the viability effects of drugs, elucidate mechanisms of action (MoA) of unknown drugs, and map morphological changes to underlying genetic perturbations. Recent developments in high-throughput screening and computer vision technology are paving the way to a deeper understanding of biological processes from cell images. However, the established methods take the single-cell segmentation approach followed by hand-crafted feature extraction. Focusing on single cells could potentially mask the system-level effects such as inter-cellular relations or proliferation effect. Additionally, the segmentation process adds a resource-intensive computational step. Furthermore, using handcrafted features relies on known domain expertise, and might miss hidden representation beyond the current human knowledge. To fully leverage all the information that images contain, feature extraction in an unbiased manner using deep learning-based computer vision (CV) models is a promising alternative. In the last few years, there have been several papers published using CV models on cell image data. While some of them show significant performance by re-training the model on the cell image data, it may pose other problems such as the lack of generalizability (e.g. re-trained model is specific to a particular dataset) and resource intensiveness that hinders researchers from analyzing a variety of datasets. In this paper, we introduce and show a simpler pipeline based on a CV model without re-training but using commonly available natural image-based trained weights with the proposed novel correction method that

already achieves the state-of-the-art (SOTA) equivalent performance. We will also explore and show how the pipeline can discern different MoAs and gene groups.

P42: Feng Yan - The Walter and Eliza Hall Institute of Medical Research

Evaluation of de novo transcriptome assembly software for long-read RNAseq and its application to cancer

Feng Yan [1], Nadia M Davidson [1]
[1] Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, Victoria 3052, Australia

Long-read sequencing has significantly advanced transcriptomic studies by enabling the capture of comprehensive transcriptome with full-length splice variants and fusion genes. However, current analyses rely heavily on a high-quality reference genome and gene annotation. Recent development of de novo assembly software for long-read data allows the benefits of long-reads to be leveraged without a reference, for example, in non-model organisms and cancers with highly rearranged genomes.

However, despite this opportunity, benchmarking and protocols to guide the best approach for reference-free, long-read assembly and differential analysis are lacking. Here, we present a comprehensive evaluation of de novo long-read transcriptome assembly tools, RATTLE and RNAbloom2. We also compare against the state-of-the-art short-read de novo assembler, Trinity. Using genome-guided assembly from Bambu and spike-in sequin transcripts as ground truth, we evaluated assembly quality, transcriptome completeness, computational efficiency, and effect on downstream differential gene and transcript analysis on several long-read datasets from simulation and cancer cell lines (cDNA and dRNA).

We show that RATTLE assembled the longest transcripts, while RNAbloom2 assembled the most transcripts, albeit with redundancy issues. The computational efficiency became a bottleneck for RATTLE with larger datasets. As RNAbloom2 does not cluster transcripts into genes, we used the short-read clustering tool Corset for this purpose. Interestingly, the clustering was also improved when using Corset with RATTLE and Trinity, compared to their native gene-transcript mappings.

For downstream analysis, the differential gene expression results from all methods showed high consistency with the Bambu assembly and sequin truth, but for differential transcript-level analysis, RNAbloom2+Corset was the best de novo combination. We also highlighted de novo methods' capability to detect mutations, INDELs, and fusions missed by reference-based approaches in cancer cell line data.

Our study offers an opportunity for researchers to leverage the strengths of each method and suggests further improvement of current long-read assembly software.

P45: Hanyun Zhang - Garvan Institute of Medical Research

Deciphering fibroblast functional heterogeneity using non-negative matrix factorization

Hanyun Zhang[1], Kate Harvey[1], Daniel Lee Roden[1,2], Beata Kiedik[1,2], Eva Apostolov[1,2], Hue M. La[3], Luciano G. Martelotto[4], Fernando J. Rossello[3,5], Sean M. Grimmond[3], Alex Swarbrick[1,2]
[1] The Kinghorn Cancer Centre and Cancer Research Theme, Garvan Institute of Medical Research, Darlinghurst, New South Wales, Australia
[2] St Vincent's Clinical School, Faculty of Medicine, University of New South Wales, Sydney, New South Wales, Australia
[3] University of Melbourne Centre for Cancer Research, University of Melbourne, Melbourne, VIC 3000, Australia
[4] Adelaide Centre for Epigenetics, South Australian Immunogenomics Cancer Institute, The University of Adelaide, Adelaide, SA, Australia

[5] Murdoch Children's Research Institute, The Royal Children's Hospital, Melbourne, VIC 3052, Australia

Fibroblasts are crucial regulators of tumor progression, with their phenotypes heavily dependent on the surrounding environment. Diverse fibroblast populations have been found across different tumor types. However, the full spectrum of fibroblast function in tumors remains incompletely characterized. Here we employed non-negative matrix factorization (NMF) to capture the functional states of fibroblasts, represented as meta-programs (MPs). This analysis encompassed 30547 fibroblasts from 146 samples of prostate cancer, pancreatic ductal adenocarcinoma (PDAC), and breast cancer.
We identified 13 MPs, each consisting of 50 genes, and subsequently annotated them based on the predominant biological processes. Expression of MPs was quantified for individual cells using the AUCell algorithm. We observed an enrichment of extracellular matrix (ECM) organization MP in myofibroblast-like cancer-associated fibroblasts (myCAFs) of breast cancer and PDAC, consistent with the ECM remodeling role of myCAFs. Additionally, the MP associated with the regulation of leukocyte genes was enriched in inflammatory CAF (iCAF)-like fibroblasts in breast cancer, while an MP related to smooth muscle cell (SMC) proliferation was enriched in SMC-like fibroblasts in prostate cancer, highlighting the ability of this method to capture gene sets associated with the phenotypic functions of fibroblasts.
Beyond canonical fibroblast functions, we identified an MP composed of genes responsive to bone morphogenetic protein (BMP) stimulation, including CHRDL1, SFRP4, SFRP1, TGFBR3, GPC3, and NFIA. This MP exhibited enrichment in several fibroblast phenotypes in prostate cancer, as well as in the iCAF-like population in breast cancer. Notably, CAFs derived from BMP-treated mouse models of prostate and breast cancer have shown increased expression of inflammatory chemokines and pro-tumor activity. We hypothesized that this MP may characterize a CAF population differentiated into iCAF phenotype in response to the BMP stimulation. In summary, through NMF analysis, we identified MPs revealing heterogeneous function and regulation of fibroblasts within the tumor microenvironment.

P46: Mark Ziemann - Burnet Institute

Improved pathway enrichment analysis of Infinium methylation data

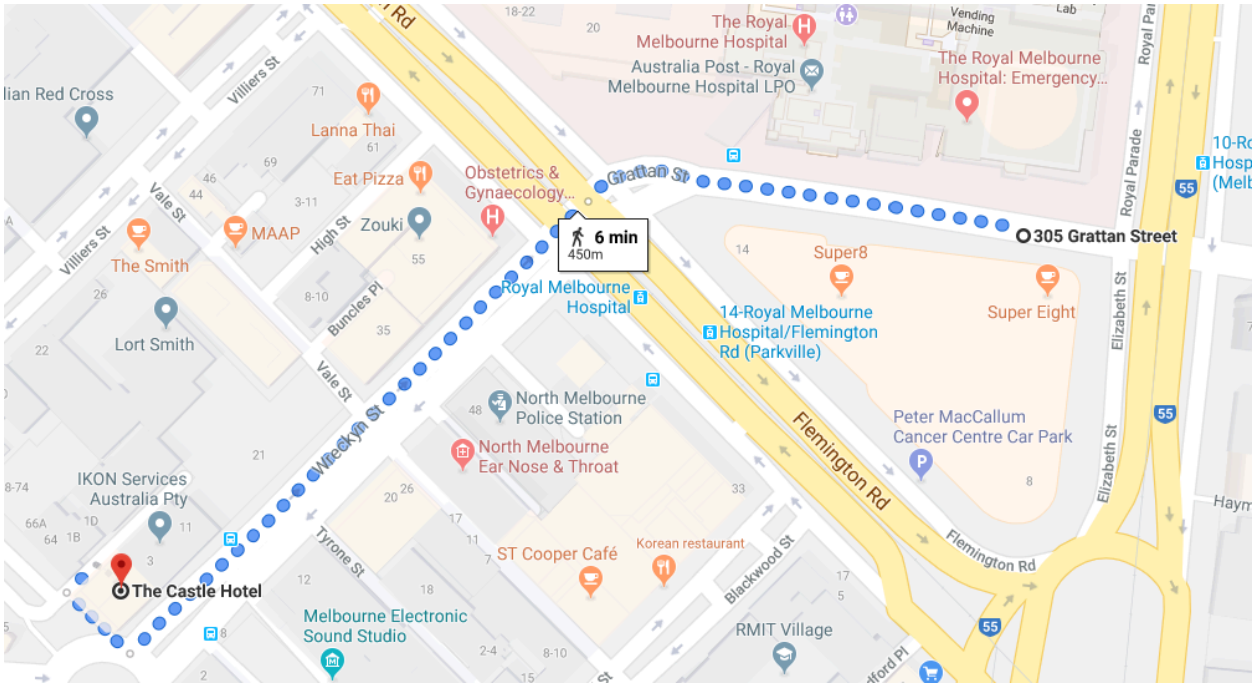
Mark Ziemann [1,2], Mandhri Abeysooriya [2,3], Anusuiya Bora [1,2], Séverine Lamon[3], Mary Sravya Kasu[2], Mitchell W. Norris[2], Yen Ting Wong[4,5], Jeffrey M. Craig[4,5]
[1] Burnet Institute, Melbourne, Australia
[2] Deakin University, School of Life and Environmental Sciences, Geelong, Australia
[3] Deakin University, School of Exercise and Nutrition Sciences, Institute for Physical Activity and Nutrition, Geelong, Australia
[4] Deakin University, School of Medicine, Geelong, Australia. Murdoch Children's Research Institute, Melbourne, Australia.

Infinium Methylation BeadChip arrays remain one of the most popular platforms for studying the epigenetics of cancer, but tools for downstream pathway analysis have their limitations. Functional class scoring (FCS) is a group of pathway enrichment techniques that involve the ranking of genes and evaluation of their collective regulation in biological systems, but the implementations described for Infinium methylation array data don't retain direction information, which is important for mechanistic understanding of genomic regulation. After evaluating several candidate FCS methods, we found that mean aggregation of probe limma t-statistics by gene followed by a rank-ANOVA enrichment test out performed existing over-representation analysis method in simulations. This method also showed improved sensitivity and robustness in an analysis of real lung tumour-normal paired datasets. Using matched RNA-seq data we examine the relationship of methylation differences at promoters and gene bodies with RNA expression at the level of pathways in lung cancer. We then analysed blood-based signatures for 14 prevalent and 19 incident disease states in a group of 18,413 participants. Results indicate epigenetic changes at the pathway level preceding diagnosis for cancer of the prostate, ovary, lung, breast and colon.

Social Event

We are planning an informal social event at the conclusion of CBA. It is walking distance from Peter Mac at The Castle Hotel (upstairs), from 6.00pm onwards.

The Castle Hotel
56 Courtney St
North Melbourne 3051
https://thecastlehotel.com.au/



Announcement: 2025

Please join us in Sydney for the next Cancer Bioinformatics Australia.
Information coming soon.

If you would like to be on the organising committee please contact:

Chelsea Mayoh
cmayoh@ccia.org.au