Program Booklet

26 Oct: IHEC Annual Meeting
27 – 28 Oct: IHEC Science Days

HKUST, Hong Kong

Table of Contents

Important information

Pre-arrival

Currency

The legal tender in Hong Kong is the Hong Kong dollar (HKD), which is pegged to the US dollar at a rate of about **7.80 HKD to 1 USD**, although exchange rates may fluctuate slightly.
Currency converter by WORLDFIRST (For reference only)

Time Differences

Hong Kong Time is defined as **UTC+8**

Electricity & Voltage

The standard electrical voltage in Hong Kong is **220 volts AC, 50Hz**. Most hotel bathrooms also have outlets for 100 volts, but if not, you will need a transformer for any appliance or electrical equipment. The majority of electrical outlets in Hong Kong take **a three-pronged UK-style plug** (Details). You can buy an inexpensive adaptor for your electrical equipment at most convenience stores.

Internet Access

In HKUST campus, **eduroam** is available. Please visit the Wi-fi Services page by HKUST for more information.

Free Wi-Fi services are widespread in Hong Kong. In particular, there are many **Wi-Fi.HK** hotspots widely available across the city in locations such as major tourist attractions, shopping malls, parks and gardens, visitor centers, sports venues, museums, ferry terminals, cooked food centers and government buildings, which offer wireless internet connection completely free or free for a period of time. Download the **Wi-Fi.HK app** to search for a list of hotspot locations, where you can connect to the internet with the Wi-Fi.HK SSID.

Upon arrival

Travelling by Taxi in HK

Taxis are easily available from the airport. You may expect to pay approximately HKD 300-380 for the journey. **Note that only HK dollars (cash) will be accepted by your driver.** The taxi drop-off point on campus is near the red sundial at the North Entrance. For heading to IAS, you may use the map of driving route to IAS (download here) to give the taxi driver directions. For on-campus housing destinations, please ask the security guard at the main gate for directions.

Program Schedule

26 Oct: IHEC Annual Meeting

			Venue:
08:00 – 09:00	*Registration*		
09:00 – 10:15	**Working group sessions**		
		Chair	
	Assay Standards and Quality Control	*Martin Hirst*	IAS LT
	Bioethics	*Yann Joly*	IAS 2042
10:15 – 10:45	*Coffee break*		G/F Lobby
10:45 – 12:00	**Working group sessions**		
		Chair	
	Data Ecosystem / Integrative Analysis	*Guillaume Bourque*	IAS LT
	Communication	*Stephanie Weber*	IAS 2042
12:00 – 13:00	*Lunch*		G/F Lobby
		Chair	
13:00 – 13:15	**Welcome & General Remarks**	*Eric Marcotte*	IAS LT
13:15 – 15:15	**Consortium updates**	***Presenter***	
13:15 – 13:35	Canada	*Eric Marcotte / Martin Hirst*	
13:35 – 14:00	EU	*Tomasz Dylag / Philip Rosenstiel / Maja Jagodic*	
14:00 – 14:10	Germany	*Jörn Walter*	
14:10 – 14:25	Hong Kong	*Karl Herrup / Danny Leung*	
14:25 – 14:35	Japan	*Toshikazu Ushijima*	
14:35 – 14:45	Singapore	*Shyam Prabhakar*	
14:45 – 14:55	South Korea	*Young Jin Kim / Suman Lee*	
14:55 – 15:05	USA 4D Nucleome	*Ananda Roy*	
15:05 – 15:15	USA ENCODE	*Jason Hilton*	
15:15 – 15:25	New Zealand	*Aniruddha Chatterjee*	
15:25 – 16:00	*Coffee break*		G/F Lobby
		Chair	
16:00 – 17:00	**Working group updates**	*Martin Hirst*	IAS LT
		Presenter	
	Assay Standards and Quality Control	*Martin Hirst*	
	Bioethics	*Yann Joly*	
	Data Ecosystem	*TBD*	
	Integrative Analysis	*Guillaume Bourque*	
	Communication	*Stephanie Weber*	
		Chair	
17:00 – 18:00	**Future of IHEC**	*Martin Hirst / Eric Marcotte*	
18:05 – 19:00	**IHEC ISSC**	*Closed meeting*	
18:30 / 19:00	***Transport to IHEC dinner*** *(in Sai Kung)*		

27 Oct: IHEC Science Days – Day 1

Chiang Chen Lecture Theater (LT-J), HKUST

09:00 – 09:15 **Opening remarks**

09:15 – 10:45 **Session 1: Single Cell Analyses**
Exploring the physical genome
William Greenleaf

Decoding the gene regulation network in human germline cells by single-cell functional genomics
Fuchou Tang

Transcriptional and epigenetic states of oligodendroglia in development and disease: insight from single cell omics
Gonçalo Castelo-Branco

(Selected talk) *DNA hypermethylation encroachment at CpG island borders in cancer*
Susan Clark

10:45 – 11:15 Coffee break

11:15 – 12:45 **Session 2: Developmental Epigenetics**
Conservation and divergence of chromatin reprogramming in early animal development
Wei Xie

Variable silencing of the repeat genome - implications for epigenetic inheritance
Anne Ferguson-Smith

Crosstalk of N-α-acetyltransferase and DNA methylation in development and disease
Li-Jung Juan

(Selected talk) *Enhancing cellular reprogramming by directed factor evolution*
Ralf Jauch

12:45 – 13:15 **Rapid Fire Talks**
Chromatin Integration Labeling Technology (ChILT): an immunoprecipitation-free method for low-input epigenomic profiling
Hiroshi Kimura

Integrative chromatin profiling of primary liver cancers reveals epigenetic vulnerability for effective combination immunotherapy
Alfred Cheng

Histone H3K9 methyltransferase G9a in oocyte is essential for preimplantation development but dispensable for CG methylation protection
Wan Kin Au Yeung

Global DNA methylation levels regulates PD-L1 expression in melanoma
Aniruddha Chatterjee

MyoD induced enhancer RNA interacts with hnRNPL protein via CAAA motif to activate target

gene transcription during myogenic differentiation
Huating Wang

Allelic binding of human transcription factors to genetic variants that are associated with Type-2-Diabetes predisposition
Jian Yan

13:15 – 15:15 *Lunch & Poster session*
IHEC EXEC meeting (Closed meeting)

15:15 – 16:30 **Session 3: Epigenetic Regulation of Repetitive Sequences**
Single-cell chromatin accessibility landscapes in cell fate reprogramming
Jonathan Loh

Transposable elements in normal and cancer epigenome
Ting Wang

Expanded potential stem cells of mammals
Pentao Liu

16:30 – 17:00 *Coffee break*

17:00 – 18:15 **Session 4: Cancer Epigenetics**
Transcriptional Matching Identifies the Cell Lineage and Timing for the Origin for Childhood Posterior Fossa Tumors
Michael Taylor

Role of transcription factors and epigenetic enzymes in hematopoiesis and leukemogenesis
Marjorie Brand

A vicious combination of Tet repression and increased Dnmt activity underlies aberrant DNA methylation in human disorders
Toshikazu Ushijima

19:00 ***Conference dinner***
China Garden Restaurant, G/F, HKUST

28 Oct: IHEC Science Days – Day 2

Chiang Chen Lecture Theater (LT-J), HKUST

9:00 – 10:15 **Session 5: Transcriptional Regulation**

Functional elucidation of non coding regulatory elements
Jay Shin

Transgenerational epigenetic inheritance: From the epigenome to behavior
Isabelle Mansuy

Engaged RNA polymerases clarify regulatory architectures of the epigenome
John Lis

10:15 – 10:35 Coffee break

10:35 – 11:55 **Session 6: 3D Chromatin and RNA Structure**

A 3D code in the human genome
Erez Liberman-Aiden

Genome organization of dengue and Zika viruses
Yue Wan

(Selected talk) *Structural cell biology of eukaryotic nuclei*
Lu Gan

(Selected talk) *Inference of genomic spatial organization from a whole genome bisulfite sequencing sample*
Simon Heath

11:55 – 12:25 **Rapid Fire Talks**

Epigenomic signatures in sperm associated with Body Mass Index (BMI) and male infertility
Sarah Kimmins

A supervised method for enhancer identification and linkage to target genes
Marcel Schulz

Characterization of the Molecular Consequences of CIC-knockout and Neomorphic IDH1 R132H Mutation on Transcriptomic and Epigenomic Landscapes
Dongsoo (Stephen) Lee

Genome-wide DNA methylation profiling identifies convergent molecular signatures associated with idiopathic and syndromic forms of autism in post-mortem human brain tissue
Chloe Wong

Regulation of peripheral heterochromatin domain organization via histone and non-histone protein methylation
Shravanti Rampalli-Deshpande

Systematic 3D Genome Architecture Analysis in Xenopus tropicalis
Chunhui Hou

12:25 – 13:25 *Lunch*
NIH 4D Nucleome Workshop - Venue: Mr and Mrs Lee Siu Lun Lecture Theater (LT-K)

13:25 – 14:25 *Poster session*

14:25 – 15:40 **Session 7: Computational Epigenetics**
Decoding epigenetic programs in differentiation and disease
Christina Leslie

3D Genome organization in cancer cells
Feng Yue

Understanding transcriptional regulation by mining epigenomics data
Kevin Yip

15:40 – 16:00 *Coffee break*

16:00 – 16:30 **Selected short talks**
(Selected talk) *The ethics of epigenetics research: An overview of the activities of the IHEC Bioethics Workgroup*
Yann Joly

(Selected talk) *Epigenetic Discrimination: Emerging Applications of Epigenetics Calling for Ethical Scrutiny*
Charles Dupras

16:30 – 18:00 **Session 8: Disease Epigenetics**
Coordinated signatures of DNA methylation and gene expression in human aging
Karen Conneely

Epigenomic trajectories to neuropsychiatric and neurodegenerative disease
Jonathan Mill

Applied epigenomics: insights into the pathogenesis of Multiple Sclerosis
Maja Jagodic

(Selected talk) *Single nuclei chromatin accessibility analysis reveals epigenetic heterogeneity of mouse brain regions*
Sebastian Preissl

18:00 – 18:15 **Closing Remarks**

18:30 ***Transport to social event:***
The HK Wine and Dine Festival *(in Central)*

Science Day 1 – Invited Talks

Session 1: Single Cell Analyses

William Greenleaf: *Exploring the physical genome*

Fuchou Tang: *Decoding the gene regulation network in human germline cells by single-cell functional genomics*

Gonçalo Castelo-Branco: *Transcriptional and epigenetic states of oligodendroglia in development and disease: insight from single cell omics*

Susan Clark: (Selected talk) *DNA hypermethylation encroachment at CpG island borders in cancer*

Session 2: Developmental Epigenetics

Wei Xie: *Conservation and divergence of chromatin reprogramming in early animal development*

Anne Ferguson-Smith*: Variable silencing of the repeat genome - implications for epigenetic inheritance*

Li-Jung Juan: *Crosstalk of N-α-acetyltransferase and DNA methylation in development and disease*

Ralf Jauch: (Selected talk) *Enhancing cellular reprogramming by directed factor evolution*

Rapid Fire Talks

Hiroshi Kimura: *Chromatin Integration Labeling Technology (ChILT): an immunoprecipitation-free method for low-input epigenomic profiling*

Alfred Cheng: *Integrative chromatin profiling of primary liver cancers reveals epigenetic vulnerability for effective combination immunotherapy*

Wan Kin Au Yeung: *Histone H3K9 methyltransferase G9a in oocyte is essential for preimplantation development but dispensable for CG methylation protection*

Aniruddha Chatterjee: *Global DNA methylation levels regulates PD-L1 expression in melanoma*

Huating Wang: *MyoD induced enhancer RNA interacts with hnRNPL protein via CAAA motif to activate target gene transcription during myogenic differentiation*

Jian Yan: *Allelic Binding of Human Transcription Factors to Genetic Variants that Are Associated with Type-2-Diabetes Predisposition*

Session 3: Epigenetic Regulation of Repetitive Sequences

Jonathan Loh: Single-cell chromatin accessibility landscapes in cell fate reprogramming

Ting Wang: *Transposable elements in normal and cancer epigenome*

Pentao Liu: *Expanded potential stem cells of mammals*

Session 4: Cancer Epigenetics

Michael Taylor: *Transcriptional Matching Identifies the Cell Lineage and Timing for the Origin for Childhood Posterior Fossa Tumors*

Marjorie Brand: *Role of transcription factors and epigenetic enzymes in hematopoiesis and leukemogenesis*

Toshikazu Ushijima: *A vicious combination of Tet repression and increased Dnmt activity underlies aberrant DNA methylation in human disorders*

William Greenleaf
Department of Genetics, Stanford University

Single Cell Analysis

Exploring the physical genome

Abstract:
Chromatin accessibility is a powerful lens to explore mechanisms of gene expression regulation, as regions of increased chromatin accessibility represent genetic elements that have the potential to regulate gene expression. To define the open chromatin landscape in primary human tissue, we collected single-cell chromatin accessibility profiles across 10 populations of immunophenotypically defined human hematopoietic cell types and constructed a chromatin accessibility landscape of human hematopoiesis to characterize differentiation trajectories. We find variation consistent with lineage bias toward different developmental branches in multipotent cell types. We observe heterogeneity within common myeloid progenitors (CMPs) and granulocyte-macrophage progenitors (GMPs) and develop a strategy to partition GMPs along their differentiation trajectory. Furthermore, we integrated single-cell RNA sequencing (scRNA-seq) data to associate transcription factors to chromatin accessibility changes and regulatory elements to target genes through correlations of expression and regulatory element accessibility. Overall, this work provides a framework for integrative exploration of complex regulatory dynamics in a primary human tissue at single-cell resolution. We have also recently completed a survey of the chromatin accessibility landscape in primary human tumor tissue, providing a catalog of regulatory elements across human cancers.

Fuchou Tang
School of Life Sciences, Peking University

Single Cell Analysis

Decoding the gene regulation network in human germline cells by single-cell functional genomics approaches

Abstract:
Human germline cells are crucial for maintenance of the species. However, the developmental trajectories and heterogeneity of human germline cells remain largely unknown. We performed single-cell RNA-seq and DNA methylome sequencing analyses of human germline cells in female and male human embryos spanning several critical developmental stages. We found that female fetal germ cells (FGCs) undergo four distinct sequential phases characterized by mitosis, retinoic acid signaling, meiotic prophase, and oogenesis. Male FGCs develop through stages of migration, mitosis, and cell-cycle arrest. Individual embryos of both sexes simultaneously contain several subpopulations, highlighting the asynchronous and heterogeneous nature of FGC development. Moreover, we observed reciprocal signaling interactions between FGCs and their gonadal niche cells, including activation of the bone morphogenic protein (BMP) and Notch signaling pathways. Our work provides key insights into the crucial features of human germline cells during their highly ordered mitotic, meiotic, and gametogenetic processes in vivo.

Gonçalo Castelo-Branco
Department of Medical Biochemistry and Biophysics, Karolinska Institutet

Single Cell Analysis

Transcriptional and epigenetic states of oligodendroglia in development and disease: insight from single cell omics

Abstract:
Oligodendrocytes are glial cells that mediate myelination of neurons, a process that allows efficient electrical impulse transmission in the central nervous system (CNS). An autoimmune response against myelin triggers demyelination in multiple sclerosis (MS). Oligodendrocyte precursor cells (OPCs) can initially differentiate and promote remyelination in MS, but this process eventually fails in progressive MS. In order to clearly define transcriptional and epigenetic states of OPCs and other oligodendrocyte lineage cells during development and disease, we have performed single-cell RNA-Seq and ATAC-Seq of cells of the oligodendrocyte lineage from healthy mouse CNS and also from a mouse model of MS. We identified several cell states/populations, representing unique stages during the process of differentiation, myelination and final stages of maturation. Our results also indicate that diverse embryonic progenitor cells of the oligodendrocyte lineage from different regions of the CNS converge into cell states compatible with differentiation at postnatal stages, while subsequent divergence of the mature terminal differentiated oligodendrocytes occurs in the juvenile/adult CNS, as the neuronal circuitry matures. We also identified unique OLs and OPCs in the experimental autoimmune encephalomyelitis (EAE) mouse model of MS, expressing genes involved in antigen processing and presentation and immunoprotection. Thus, our results indicate a previously unanticipated heterogeneity of the oligodendrocyte lineage during development and disease.

Wei Xie
School of Life Sciences, Tsinghua University

Developmental Epigenetics

Conservation and divergence of chromatin reprogramming in early animal development

Abstract:
Upon fertilization, drastic chromatin reorganization occurs during early development. Deciphering the molecular events underlying this process is crucial for understanding both fundamental biology and in vitro fertilization. Previously, we have investigated chromatin reprogramming during mouse preimplantation development for chromatin accessibility, histone modifications, and 3D architecture. Recently, we also employed an improved ATAC-seq approach to investigate chromatin reprogramming in early human embryos. These studies collectively unveiled highly dynamic and non-canonical chromatin regulation during maternal-to-zygotic transition and genome activation in mammals. To explore to what extent chromatin reprogramming modes in early development are conserved across vertebrates, we further extended our analysis to zebrafish early embryos. In this talk, I will discuss our recent progress in understanding epigenetic inheritance and reprogramming during animal embryogenesis. These results will shed light on the conservation and divergence of the underlying molecular principles for chromatin reprogramming in early vertebrate development across different species.

Anne Ferguson-Smith
Department of Genetics, University of Cambridge

Developmental Epigenetics

Variable silencing of the repeat genome - implications for epigenetic inheritance

Abstract:
Genetic models of epigenetic inheritance can provide useful insights into mechanisms, stability and heritability of modified states. Endogenous retroviruses (ERVs) are a class of LTR retrotransposons representing around one quarter of the repetitive elements in the murine genome. Most are epigenetically silenced. However, in two classic mouse models, Agouti viable yellow (Avy) and Axin fused (Axin(Fu)), a member of the IAP class of ERV has inserted in the vicinity of the agouti and axin genes respectively, and been variably DNA methylated between individuals; this is associated with transcriptional variability of the associated genes, and non-genetically conferred phenotypic variation which can be transmitted across generations. Such alleles are known as metastable epialleles. Here we present a genome-wide systematic screen for novel endogenous metastable epialleles with or without an impact on transcription using mouse strain-specific datasets that we generated as part of the BLUEPRINT reference epigenome project (EUFP7 BLUEPRINT grant HEALTH-F5-2011-282510). The properties, impact and heritability of these alleles has been explored and provided useful insights into the impact of the epigenetic control of repetitive elements on the mammalian genome, and transgenerational epigenetic inheritance. In particular, we have generated a new model for quantifying the extent of transgenerational epigenetic inheritance and unravelling mechanisms involved.

Li-Jung Juan
Genomics Research Center, Academia Sinica

Developmental Epigenetics

Crosstalk of N-α-acetyltransferase and DNA methylation in development and disease

Abstract:
Mammalian N-α-acetyltransferase 10 protein (Naa10p) catalyzes N-α-acetylation of nascent proteins. We previously reported that human Naa10p overexpression is required for DNA methyltransferase 1 (DNMT1) to bind, to methylate and to down-regulate specific tumor suppressor genes such as E-cadherin, leading to lung tumorigenesis (J Clinc Invest, 2010). Importantly, mutation of Naa10p is linked to severe developmental defects including intellectual disability and infancy death in patients with Ogden syndrome. Consistently, our Naa10-null mice display partial embryonic lethality and growth defects and further link Naa10p mutation-associated human diseases to defective DNA methylation and genomic imprinting (Mol Cell, 2017). I will discuss how Naa10p crosstalks with DNA methylation in developmental control, especially in adipogenesis and neuronal function.

Jonathan Loh
A*STAR Institute of Molecular and Cell Biology

Epigenetic Regulation of Repetitive Sequences

Single-cell chromatin accessibility landscapes in cell fate reprogramming

Abstract:
Somatic cells are converted to pluripotent stem cells by inducing the expression of specific transcription factors. However, derivation of bona-fide induced pluripotent stem cells is limited by low efficiency of the process and often heterogeneously reprogrammed cell populations. Here, we developed a live-cell fluorescent probe which specifically stains and sort for cells undergoing early reprogramming. We applied scATAC-seq to analyze these intermediate reprogramming cells to decipher the underlying mechanism governing epigenetic changes during the process. Our study revealed that cells undergoing reprogramming can exhibit bi-lineage chromatin profiles. This is confirmed by the co-detection of myoblast and fibroblasts genes or myoblast and pluripotent gene within a single cell. Additionally we identified novel functional enhancer nodes that regulate reprogramming. When coupled with high-resolution 4C-seq, we were able to capture the distant target genes controlled by the enhancer nodes. Analyzing single-cell chromatin profile of cells undergoing reprogramming, demonstrated the dynamic switching of enhancers' activity resulting in rewiring of cell fates.

Ting Wang
Department of Genetics, Washington University

Epigenetic Regulation of Repetitive Sequences

Transposable elements in normal and cancer epigenome

Abstract:
Advances in next-generation sequencing platforms have reshaped the landscape of genomic and epigenomic research towards the understanding of roles of human transposable elements. It is now possible to map the epigenetic landscape of transposable element (TE) across many tissue and cell types as well as in diseases. It is also possible now to directly manipulate TEs genetically and epigenetically. In the presentation I will discuss tools developed for this purpose and present some results from investigating regulatory roles of transposable elements in normal and cancer cells, using public genomic and epigenomic resources.

Pentao Liu
School of Biomedical Sciences, The University of Hong Kong

Developmental Epigenetics

Expanded potential stem cells of mammals

Xuefei Gao, Monika Nowak-Imialek, Xi Chen, Heiner Niemann, Aarah Teichmann and Pentao Liu

Abstract:
Embryonic stem cells (ESCs) derived from the epiblast contribute to the somatic lineages and the germline upon reintroduction to the mouse blastocyst but are excluded from the extraembryonic tissues that are derived from the trophectoderm (TE) and the primitive endoderm (PrE). By inhibiting signal pathways implicated in the earliest embryo development, we recently established cultures of mouse expanded potential stem cells (EPSCs) from individual 8-cell blastomeres, by direct conversion of mouse embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs). A single EPSC can contribute to both the embryo proper and the TE lineages in chimera assay. *Bona fide* trophoblast stem cell (TSC) lines, extraembryonic endoderm stem (XEN) cells, and ESCs could be directly derived from EPSCs *in vitro*. Molecular analyses of the epigenome and single-cell transcriptome reveal enrichment for blastomere-specific signature and a dynamic DNA methylome in EPSCs. The knowledge of mouse EPSCs now enables establishing EPSCs of other mammalian species, where robust ESCs are not currently available. The EPSCs of these species have similar molecular features, and can differentiate to extraembyonic as well embryonic cell lineages *in vitro* and *in vivo*. The successful generation of EPSCs produces new tools for investigation of embryonic development, and opens a wealth of avenues for translational research in biotechnology, agriculture, and genomics and regenerative medicine.

Michael Taylor
The Arthur and Sonia Labatt Brain Tumour Research Centre

Cancer Epigenetics

Transcriptional matching identifies the cell lineage and timing for the origin for childhood posterior fossa tumors

Abstract:
Single cell sequencing (RNA-seq) of >80,000 cells from the developing fetal and post-natal cerebellum allows us to develop a 'scaffolding' of the cellular basis of cerebellar development. From there we compare the different cell types in the developing cerebellum to single cell sequencing of human medulloblastomas, ependymomas, and pilocytic astrocytomas. We can demonstrate from these comparisons that each of the different subgroups of medulloblastoma, ependymoma, and pilocytic astrocytoma come from specific populations in the developing cerebellum, including origins from cell populations that had not previously been identified. Beyond originating from a single cell type, childhood brain tumors appear to transcriptionally match, and likely arise from those specific cell lineages at very specific time points in development. These results identify the cell of origin for Group 3 medulloblastoma, Group 4 medulloblastoma, Shh medulloblastoma, PFA ependymoma, PFB ependymoma, and cerebellar pilocytic astrocytoma. Our results further support a model in which the transcriptome of childhood brain tumors is largely set by the specific time and cell lineage in which the transformation event occurs supporting the importance of epigenetics in these childhood cancers. Further, the restriction of the cell of origin for many brain tumors to fetal life, suggests that childhood brain tumors are imitated in utero, and offers a proximate explanation for their restriction to pediatric age groups.

Marjorie Brand
Sprott Center for Stem Cell Research, Ottawa Hospital Research Institute

Cancer Epigenetics

Role of transcription factors and epigenetic enzymes in hematopoiesis and leukemogenesis

Abstract:
Chromatin modifying enzymes and transcription factors play critical roles in regulating hematopoiesis. Furthermore, deregulation of the transcriptional and epigenetic regulatory network contributes to leukemic transformation. To better understand the mechanism through which TFs regulate cell fate decision in health and disease, we have used a multi-pronged approach that combines “omics”- based approaches together with in vivo models of hematopoiesis and leukemia. At the meeting, I will present our latest results on the role of the H3K27 demetylase UTX in the maintenance of T-cell acute lymphoblastic leukemia. Furthermore, I will show our new results on the importance of quantitative changes in transcription factor protein levels in the regulation of cell fate decisions.

Toshikazu Ushijima
Division of Epigenomics, National Cancer Center Research Institute

Cancer Epigenetics

A Vicious Combination of *Tet* Repression and Increased Dnmt Activity Underlies Aberrant DNA Methylation in Human Disorders

Abstract:

Aberrant DNA methylation is heavily involved in the development of human disorders, such as cancer, neurodegenerative disorders, and metabolic disorders. Aberrant methylation is known to be induced in relation to aging and to be accelerated by certain types of inflammation [Issa, J Clin Invest, 124:24, 2014; Ushijima, Clin Cancer Res, 18:923, 2012]. The accumulation level of aberrant DNA methylation in normal tissues can be correlated with cancer risk, as unequivocally demonstrated by a prospective clinical study [Maeda, Gut, 66:1721, 2017]. We recently compared CpG sites methylated by aging in gastric mucosae with those methylated by inflammation triggered by *H. pylori* infection, which is one of the most potent inducers of aberrant DNA methylation [Niwa, Cancer Res, 70:1430, 2010]. *H. pylori*-triggered inflammation indeed accelerated age-related methylation, but also induced methylation of CpG sites unaffected by aging, involving genes potentially responsible for carcinogenic processes [Nanjo, MS in preparation]. Mechanistically, *Il1b, Tnf,* and *Nos2* expression has been known to be associated with methylation induction in multiple animal models [Hur, Carcinogenesis, 32:35, 2011; Katsurano, Oncogene, 31:342, 2012]. We found that NF-κB activation, induced by IL-1β and TNF-α, increased expression of some miRNA genes, and that these miRNAs in turn repressed expression of three *Tet* genes, which lead to a decreased hmC content. At the same time, increased production of nitric oxide, due to increased *Nos2* expression, enhanced Dnmt enzymatic activity, as previously reported. Coexistence of *Tet* repression and increased Dnmt activity led to aberrant DNA methylation of a large number of CpG sites [Takeshima, MS submitted]. This suggests that a vicious combination of *Tet* repression and increased Dnmt activity, which can be present in specific biological conditions, such as *H. pylori*-triggered inflammation, neuroinflammation, and obesity, may underlie aberrant DNA methylation in a variety of human disorders.

Science Day 2 – Invited Talks

Session 5: Transcriptional Regulation

Jay Shin: *Functional elucidation of non coding regulatory elements*

Isabelle Mansuy: *Transgenerational epigenetic inheritance: From the epigenome to behavior*

John Lis: *Engaged RNA polymerases clarify regulatory architectures of the epigenome*

Session 6: 3D Chromatin and RNA Structure

Erez Liberman-Aiden: *A 3D code in the human genome*

Yue Wan: *Genome organization of dengue and Zika viruses*

Lu Gan: (Selected talk) *Structural cell biology of eukaryotic nuclei*

Simon Heath: (Selected talk) *Inference of genomic spatial organization from a whole genome bisulfite sequencing sample*

Rapid Fire Talks

Sarah Kimmins: *Epigenomic signatures in sperm associated with Body Mass Index (BMI) and male infertility*

Marcel Schulz: *A supervised method for enhancer identification and linkage to target genes*

Dongsoo (Stephen) Lee *: Characterization of the molecular consequences of CIC-knockout and Neomorphic IDH1 R132H mutation on transcriptomic and epigenomic landscapes*

Chloe Wong: *Genome-wide DNA methylation profiling identifies convergent molecular signatures associated with idiopathic and syndromic forms of autism in post-mortem human brain tissue*

Shravanti Rampalli-Deshpande*: Regulation of peripheral heterochromatin domain organization via histone and non-histone protein methylation*

Chunhui Hou: *Systematic 3D Genome Architecture Analysis in Xenopus tropicalis*

Session 7: Computational Epigenetics

Christina Leslie: *Decoding epigenetic programs in differentiation and disease*

Feng Yue: *3D Genome organization in cancer cells*

Kevin Yip: *Understanding transcriptional regulation by mining epigenomics data*

Selected short talks

Yann Joly: (Selected talk) *The ethics of epigenetics research: An overview of the activities of the IHEC Bioethics Workgroup*

Charles Dupras: (Selected talk) *Epigenetic discrimination: Emerging applications of epigenetics calling for ethical scrutiny*

Session 8: Disease Epigenetics

Karen Conneely: *Coordinated signatures of DNA methylation and gene expression in human aging*

Jonathan Mill: *Epigenomic trajectories to neuropsychiatric and neurodegenerative disease*

Maja Jagodic: *Applied epigenomics: insights into the pathogenesis of Multiple Sclerosis*

Sebastian Preissl*:* (Selected talk) *Single nuclei chromatin accessibility analysis reveals epigenetic heterogeneity of mouse brain regions*

Jay Shin
Laboratory for Advanced Genomics Circuit, RIKEN Center for Integrative Medical Sciences
Transcriptional Regulation

Functional elucidation of non coding regulatory elements

Abstract:
The recent FANTOM has unraveled a myriad of transcriptome diversity across 400 human cell types where many were considered non-coding RNAs (FANTOM5 consortium). Furthermore, single cell genomics technology is transforming how we conduct research by revealing greater complexity in both cellular and genetic levels. In the lab, we developed to capture the 5'-end of RNA molecules at the single cell resolution (i.e. single-cell CAGE). The method is revealing unique complexity of transcriptome at the single cell resolution when compared to the 3'-end based methods, allowing investigation towards the functional implications of both coding and non-coding RNAs such as eRNAs. We are further generating a functional reference set of lncRNAs across multiple human cell types using large-scale functional perturbation followed by molecular phenotyping with CAGE. Here, I will introduce the next FANTOM and reveal our current pipeline to fortify the encyclopedia of functional lncRNAs.

Isabelle Mansuy
Brain Research Institute, University of Zurich

Transcriptional Regulation

Transgenerational epigenetic inheritance: From the epigenome to behavior

Abstract:
Behavior in mammals is strongly influenced by environmental factors, particularly when experienced during early postnatal life. While positive factors can favor proper behavioral responses, negative factors such as traumatic events can alter behavior and induce diseases like borderline personality disorder, bipolar depression and antisocial behaviors. Such disorders are usually marked in individuals directly exposed but strikingly, they can also affect their offspring sometimes across several generations. The biological mechanisms underlying the transmission of trauma-induced symptoms from parent to offspring have recently started to be examined and are thought to involve non-genetic factors. This talk will present an experimental model of early traumatic stress in mice and show evidence that epigenetic factors are implicated in the expression and the inheritance of the effects of such trauma. This mouse model exhibits altered social behaviors, depressive-like symptoms, cognitive deficits, and impaired glucose regulation in adulthood. The symptoms are pronounced and persist throughout life, and are transmitted to the following offspring across several generations, through both females and males. They are associated with epigenetic alterations involving persistent changes in DNA methylation at the promoter-associated CpG island of several genes, in the brain of the offspring and the germline of their father. Further to DNA methylation, other epigenetic mechanisms involving regulation by non-coding RNAs are also involved. These findings suggest that epigenetic processes largely contribute to the impact of negative environmental exposure in early life on adult behavior and physiology, and its inheritance across generations.

John Lis
Department of Molecular Biology and Genetics, Cornell University

Transcriptional Regulation

Engaged RNA polymerases clarify regulatory architectures of the epigenome

Abstract:
The structural status of the chromatin across the genome is often assessed by examining the patterns of histone modifications and of chromatin's nuclease or transposase sensitivity. This information can be used to divide the genome into chromatin (epigenomic) states that predict functions of active gene transcription, repression, promoters, enhancers and other genomic units. The distributions of particular transcription factors also provide critical information for evaluating the functions of specific regions of the genome. Here, I describe insights into chromatin modification uncovered by mapping engaged RNA polymerases (RNAP). These transcriptional machines transcribe protein encoding genes, structural and regulatory RNAs, and are found at many active enhancers. Genome-wide Nuclear Run-on assays can localize RNAP levels across all genes, identifying the rate-limiting regulatory steps in transcription. Such assays, like GRO-seq, and its base-pair resolution version, PRO-seq, both have high sensitivity and more than a 100K-fold dynamic range. These methods identified promoter-proximal pausing as a major regulated step in transcription of most genes. Our latest nascent RNA assays preserve single-molecule information to map RNA 5'caps (or 5'uncapped) as well as their 3' ends simultaneously at both promoters and enhancers. These assays reveal two dynamic pause classes that are regulated during heat shock, and a structure of regulatory elements where transcription start sites and well-positioned nucleosomes form an interdigitated architecture. We are testing the functionality of this basic architecture in both high-throughput and locus specific assays that I will describe in this presentation.

Erez Lieberman-Aiden
The Center for Genome Architecture, Baylor College of Medicine

3D Chromatin and RNA Structure

A 3D code in the human genome

Abstract:
Stretched out from end-to-end, the human genome – a sequence of 3 billion chemical letters inscribed in a molecule called DNA – is over 2 meters long. Famously, short stretches of DNA fold into a double helix, which wind around histone proteins to form the 10nm fiber. But what about longer pieces? Does the genome's fold influence function? How does the information contained in such an ultra-dense packing even remain accessible? In this talk, I describe our work developing 'Hi-C' (Lieberman-Aiden et al., Science, 2009; Aiden, Science, 2011) and more recently 'in-situ Hi-C' (Rao & Huntley et al., Cell, 2014), which use proximity ligation to transform pairs of physically adjacent DNA loci into chimeric DNA sequences. Sequencing a library of such chimeras makes it possible to create genome-wide maps of physical contacts between pairs of loci, revealing features of genome folding in 3D. Next, I will describe recent work using in situ Hi-C to construct haploid and diploid maps of nine cell types. The densest, in human lymphoblastoid cells, contains 4.9 billion contacts, achieving 1 kb resolution. We find that genomes are partitioned into contact domains (median length, 185 kb), which are associated with distinct patterns of histone marks and segregate into six subcompartments. We identify ~10,000 loops. These loops frequently link promoters and enhancers, correlate with gene activation, and show conservation across cell types and species. Loop anchors typically occur at domain boundaries and bind the protein CTCF. The CTCF motifs at loop anchors occur predominantly (90%) in a convergent orientation, with the asymmetric motifs "facing" one another. Next, I will discuss the biophysical mechanism that underlies chromatin looping. Specifically, our data is consistent with the formation of loops by extrusion (Sanborn & Rao et al., PNAS, 2015). In fact, in many cases, the local structure of Hi-C maps may be predicted in silico based on patterns of CTCF binding and an extrusion-based model. Finally, I will show that by modifying CTCF motifs using CRISPR, we can reliably add, move, and delete loops and domains. Thus, it is possible not only to "read" the genome's 3D architecture, but also to write it.

Yue Wan
Genome Institute of Singapore

3D Chromatin and RNA Structure

Genome organization of dengue and Zika viruses

Abstract:
Dengue and Zika are closely related members of the Flaviviridae family of positive, single-stranded RNA viruses and are of global clinical importance. These viruses utilize an 11kb RNA genome for translation and replication, and much remains to be learnt about how the entire genome folds to enable virus function. Here, we performed high throughput RNA secondary structure and pair-wise interactome mapping on four dengue serotypes and four Zika strains within their virus particles and in infected cells, enabling us to visualize the first 3D model of their genomes. Structure probing of dengue and Zika genomes in virions identified structurally conserved regions with covariation and low synonymous mutation rates, as well as identified an RNA structure that correlates with strong ribosome pausing. Genome-wide interactome mapping reveals extensive long-range interactions, many of which exist in alternative conformations and/or are shared among the viruses. We observe that far-away interactions in virions tend to be disrupted in cell, suggesting that interactions may be disrupted due to genome translation and replication in vivo. This comprehensive structural resource of dengue and Zika viruses reveals that viral genome organization is much more complex than previously appreciated and deepens our understanding of the molecular basis for dengue and Zika pathogenesis.

Christina Leslie
Memorial Sloan Kettering Cancer Center

Computational Epigenetics

Decoding epigenetic programs in differentiation and disease

Abstract:
Epigenetic programs govern the diverse differentiation states of immune cells as well as their dysfunctional states in tumors. We recently described computational analysis of chromatin accessibility and gene expression data in a mouse cancer model to show that tumor-specific T cells differentiate to dysfunction through two discrete chromatin states: an initial plastic state that can be functionally rescued (i.e. through immunotherapy) and a later fixed state that is resistant to therapeutic reprogramming. Here we examine the epigenetic programs underlying another non-responsive T cell state, namely tolerance to self-antigen. Self-reactive CD8 T cells that escape negative selection in the thymus must acquire a tolerant phenotype to avoid autoimmune injury. Using experimental data in mouse genetic systems from the Schietinger lab at MSKCC, we decipher the chromatin states distinguishing tolerant T cells from functional (effector, memory) T cells, computationally identify transcription factors associated with the tolerance epigenetic signature, and determine whether it is possible to break tolerance. This study has potential therapeutic implications both for cancer immunology and autoimmunity.

Feng Yue
Department of Biochemistry and Molecular Biology, Pennsylvania State University

Computational Epigenetics

Genetic variants and its impact on 3D genome structure in cancer

Abstract:
Inherited non-coding genetic variants confer significant disease susceptibility in many cancers. However, the molecular processes of how germline variants contribute to somatic lesions are poorly understood. Here we performed targeted sequencing in 5,008 patients and identified a key regulatory germline variant in GATA3 strongly associated with Philadelphia chromosome-like acute lymphoblastic leukemia (Ph-like ALL). By creating an isogenic cellular model with CRISPR-Cas9 system, we showed that this variant activated a strong enhancer that significantly upregulated GATA3 transcription, which in turn reshaped the global chromatin accessibility and 3D genome organization. Remarkably, this genotype switch induced a chromatin loop between the CRLF2 oncogene and a distal enhancer, similar to the somatically acquired super-enhancer hijacking event in patients. Finally, we showed that GATA3 directly regulates CRLF2 and potentiates the oncogenic effects of JAK-STAT signaling in leukemogenesis. Our results demonstrated how a germline non-coding variant contributes to oncogene activation epigenetic regulation and 3-dimensional genome reprogramming.

Kevin Yip
Department of Computer Science and Engineering, The Chinese University of Hong Kong
Computational Epigenetics

Understanding transcriptional regulation by mining epigenomics data

Abstract:
Epigenomics has become an important tool for identifying transcriptional regulatory elements, deciphering their functional roles, and estimating consequences of their perturbations. Knowledge discovery from the raw data requires the identification of common patterns both across different genomic regions and across different cell and tissue types. The identified patterns can then be used to model the underlying biological mechanisms and make predictions in situations with no or noisy experimental measurements. In this talk, I will describe some of my work that follows this paradigm, with an aim of answering several basic questions: 1) Whether precise gene expression levels can be explained by epigenomics signals alone; 2) To what extent different regulatory elements and their various epigenomics signals jointly affect the same target gene; 3) Whether there are important types of epigenomic signals neglected in experimental designs based on current standard practices; and 4) How a more holistic view of transcriptional regulation can be obtained by integrating many different types of epigenomics data.

Karen Conneely
Department of Human Genetics, Emory University

Disease Epigenetics

Coordinated signatures of DNA methylation and gene expression in human aging

Abstract:
Epigenome-wide association studies in humans have reported thousands of age-differentially-methylated CpG sites, and recent studies show that age can be predicted from DNA methylation data with great accuracy across a wide range of cell and tissue types. However, the role of these DNA methylation changes remains unelucidated. In whole blood, the profile of associations between age and gene expression is often reported to align poorly with the age-methylation association profile. One possible explanation for this phenomenon is that many of the DNA methylation changes observed in whole blood are not directly functional but are marks left by another process. An alternative explanation is that statistical power in gene expression studies is limited by the size of contemporary studies and the inherent noisiness and transience of gene expression. In this work, we use integrated data on DNA methylation and gene expression in human blood cells in conjunction with the results of a large well-powered study of gene expression and age to further investigate whether age-related changes in DNA methylation associate with changes in gene expression. We find that CpGs that are 1) proximal to or 2) associated with age-associated genes display enrichment patterns consistent with coordinated signatures of aging in the transcriptome and methylome.

Jonathan Mill
University of Exeter

Disease Epigenetics

Epigenomic trajectories to neuropsychiatric and neurodegenerative disease

Abstract:
There is mounting evidence to support a role for developmentally regulated epigenetic variation in the molecular etiology of neuropsychiatric and neurodegenerative disorders. In this talk I will discuss on-going work from our group aimed at identifying epigenetic variation associated with a diverse range of neuropsychiatric phenotypes including schizophrenia, autism and dementia. I will present evidence for dynamic patterns of DNA modifications (5mC and 5hmC) across human brain development, highlighting how the prenatal period is a time of considerable epigenomic plasticity in the brain, and the importance of neurodevelopmentally-dynamic loci in neurodevelopmental disease phenotypes. I will also describe the impact of genetic variation on the epigenome during brain development, presenting our recent analysis of DNA methylation quantitative trait loci (mQTLs) in a large collection of fetal and adult brain samples. Although most mQTLs are developmentally stable, a subset are characterized by fetal-specific effects and enriched amongst risk loci identified in a recent large-scale genome-wide association study (GWAS) of schizophrenia, a severe psychiatric disorder with a hypothesized neurodevelopmental component. Novel tools mean that it is now feasible to examine epigenetic variation across the genome in large numbers of samples, and I will give an overview of our recent epigenome-wide association studies (EWAS) of schizophrenia, dementia and other neuropsychiatric disorders, integrating findings with those from recent GWAS analyses. Finally, I will outline some of the issues related to epigenetic epidemiological studies of neuropsychiatric disease and explore the feasibility of identifying peripheral biomarkers of disease phenotypes manifest in inaccessible tissues such as the brain.

Maja Jagodic
Department of Clinical Neuroscience, Karolinska Institutet

Disease Epigenetics

Applied epigenomics: insights into the pathogenesis of Multiple Sclerosis

Abstract:
Multiple Sclerosis (MS) is a chronic inflammatory disease characterized by autoimmune destruction of myelin and neurons in the central nervous system. Today, MS is one of the leading causes of neurological disability in young adults. Although the cause remains unknown, vast epidemiological data establish MS as a complex disease influenced by genetic and environmental factors. Recently, we have demonstrated that epigenetic mechanisms, in particular DNA methylation, can mediate the risk of developing disease, which gives us a possibility to utilize epigenetics to better understand disease pathogenesis. One of the main challenges in understanding and managing MS is our limited knowledge of the processing occurring in the target organ - the brain. Recent progress in development of methods to survey epigenetic modifications opened up possibilities to study brain tissue and mechanisms that underlie neuronal loss in MS patients.

Selected Short Talks & Rapid Fire Talks

Science Day 1

Susan Clark: *DNA hypermethylation encroachment at CpG island borders in cancer* – **Selected Short Talk**

Ralf Jauch: *Enhancing cellular reprogramming by directed factor evolution* – **Selected Short Talk**

Hiroshi Kimura: *Chromatin Integration Labeling Technology (ChILT): an immunoprecipitation-free method for low-input epigenomic profiling*

Alfred Cheng: *Integrative chromatin profiling of primary liver cancers reveals epigenetic vulnerability for effective combination immunotherapy*

Wan Kin Au Yeung: *Histone H3K9 methyltransferase G9a in oocyte is essential for preimplantation development but dispensable for CG methylation protection*

Aniruddha Chatterjee: *Global DNA methylation levels regulates PD-L1 expression in melanoma*

Huating Wang: *MyoD induced enhancer RNA interacts with hnRNPL protein via CAAA motif to activate target gene transcription during myogenic differentiation*

Jian Yan: *Allelic Binding of Human Transcription Factors to Genetic Variants that Are Associated with Type-2-Diebetes PredispositionAllelic Binding of Human Transcription Factors to Genetic Variants that Are Associated with Type-2-Diebetes Predisposition*

Science Day 2

Lu Gan: *Structural cell biology of eukaryotic nuclei* – **Selected Short Talk**

Simon Heath: *Inference of genomic spatial organization from a whole genome bisulfite sequencing sample* – **Selected Short Talk**

Sarah Kimmins: *Epigenomic signatures in sperm associated with Body Mass Index (BMI) and male infertility*

Marcel Schulz: *A supervised method for enhancer identification and linkage to target genes*

Dongsoo (Stephen) Lee: *Characterization of the Molecular Consequences of CIC-knockout and Neomorphic IDH1 R132H Mutation on Transcriptomic and Epigenomic Landscapes*

Chloe Wong: *Genome-wide DNA methylation profiling identifies convergent molecular signatures associated with idiopathic and syndromic forms of autism in post-mortem human brain tissue*

Shravanti Rampalli-Deshpande*: Regulation of peripheral heterochromatin domain organization via histone and non-histone protein methylation*

Chunhui Hou: *Systematic 3D Genome Architecture Analysis in Xenopus tropicalis*

Yann Joly: *The ethics of epigenetics research: An overview of the activities of the IHEC Bioethics Workgroup* – **Selected Short Talk**

Charles Dupras: *Epigenetic Discrimination: Emerging Applications of Epigenetics Calling for Ethical Scrutiny* – **Selected Short Talk**

Sebastian Preissl: *Single nuclei chromatin accessibility analysis reveals epigenetic heterogeneity of mouse brain regions* – **Selected Short Talk**

Susan Clark
Garvan Institute of Medical Research

Science Day 1

DNA hypermethylation encroachment at CpG island borders in cancer

Abstract:
DNA methylation plays a key role in demarcation of regulatory regions, including promoter-associated CpG islands. While CpG islands are typically maintained in an unmethylated state, a proportion of islands are subject to hypermethylation in cancer cells due to as yet elusive means. Using WGBS we identified hypermethylated CpG islands in prostate and breast cancer, including a subset that display methylation encroachment across 5' or 3' CpG island borders. We show that the pattern of H3K4me1 at CpG island borders in normal cells determines the mode of CpG island DNA hypermethylation in cancer. Notably, genetic manipulation of KMT2D results in depletion of H3K4me1 and loss of DNA methylation at the CpG island borders, and conversely enrichment of H3K4me1 results in border hypermethylation. Together these findings identify a unique role for H3K4me1 in shaping CpG island methylation in normal and cancer cells.

Ralf Jauch
The University of Hong Kong

Science Day 1

Enhancing cellular reprogramming by directed factor evolution

Abstract:
To decode sequence-function relationship of transcription factor (TF) mediated cell fate conversions we use structural modeling and quantitative biochemical assays to analyze the formation of TF complexes on regulatory DNA in combination with genomics techniques. By contrasting the dynamic binding and gene regulation of paralogous TFs and engineered factors we begin to appreciate 'enhancer codes' directing cell fate programming. Further, we could identify functionally critical structural elements endowing selected TF to direct this process. We specifically ask how combinations of lineage specifying TFs (including SOX, OCT, PAX and FOX family proteins) work together to guide cell fate conversions in a step-wise manner. Lessons learned from these studies led us to hypothesize that native TFs are not optimized to direct artificial cell fate conversions and that they can be enhanced by directed evolution. To test this concept we generate TF libraries by randomizing selected amino acids and by recombining domains of paralogous genes. Using these libraries we perform pooled library screens, cell selection based on phenotypic read-outs and amplicon sequencing. This way, we identify artificially evolved TFs (eTFs) that program cell fates faster, more efficiently and in a more controlled fashion outperforming their wild-type counterparts. We propose that artificial reprogramming factor evolution presents a general paradigm that can be applied to any biomolecule-driven cell conversion system with utility in regenerative biomedicine.

Hiroshi Kimura
Tokyo Institute of Technology

Science Day 1

Chromatin Integration Labeling Technology (ChILT): an immunoprecipitation-free method for low-input epigenomic profiling

Abstract:
Chromatin immunoprecipitation followed by sequencing (ChIP-seq) has been the standard technique to reveal genome-wide protein–DNA interactions. However, obtaining epigenomic information from limited numbers of cells is still challenging. We have established a new immunoprecipitation-free epigenomic profiling method named Chromatin Integration Labeling Technology (ChILT), which can amplify genomic sequences near the target molecules without lysing cells. Using ChILT, we reliably profiled the distribution of histone modifications and DNA-binding factors in 100 to 1,000 cells. ChILT was also able to be applied for single cell analysis. Thus, ChILT can be a good alternative to ChIP-seq for epigenomic profiling using small numbers of cells.

Alfred Cheng
The Chinese University of Hong Kong

Science Day 1

Integrative chromatin profiling of primary liver cancers reveals epigenetic vulnerability for effective combination immunotherapy

Abstract:
The growing epidemics of obesity and diabetes, which cause nonalcoholic fatty liver disease (NAFLD), have paralleled the increasing incidence of hepatocellular carcinoma (HCC). We have previously elucidated new function of histone deacetylase 8 (HDAC8) in promoting NAFLD-associated hepatocarcinogenesis through interacting with enhancer zeste of homology 2 (EZH2) (Cancer Research 2015;75:4803-16). Blockade of immune-checkpoints such as the programmed death-ligand 1 (PD-L1)/PD-1 axis has elicited antitumor T-cell responses in a broad spectrum of cancers. However, the clinical trials in advanced HCC exhibited objective response only in a small proportion of patients. Given its causal role in oncogenicity, delineating the epigenetic regulation by HDAC8 would provide insights for improved immunotherapeutic strategies. Here, we perform nanoscale chromatin profiling of 10 primary NAFLD-associated HCCs to uncover HDAC8 and EZH2 co-regulated enhancers that are characterized by H3K27ac loss and H3K27me3 gain in tumor tissues. Coupling with their transcriptome profiles and target prediction, we have identified 1,200 repressed enhancers whose target genes are enriched in inflammatory responses, chemotaxis and leukocyte differentiation/migration. Using an orthotopic HCC preclinical model, we found that specific inhibition of HDAC8 by a selective antagonist PCI-34051 exhibited strong anti-tumorigenic effect in immunocompetent C57/Bl6 (but not nude) mice, which was accompanied with increased tumor-infiltrating cytotoxic T lymphocytes. Notably, PCI-34051 treatment significantly enhanced tumor eradication by anti-PD-L1 therapy and promoted long-term survival in all treated mice. Our integrative epigenomics and functional analysis demonstrate that selective chromatin modifications by HDAC8 alter tumor immune surveillance. Rational combinatorial epigenetic immunotherapy can therefore be devised to fully unleash anti-tumor T cell responses, leading to complete remission of HCC.
This work is supported by the RGC CRF (C4017-14G).

Wan Kin Au Yeung
Kyushu University

Science Day 1

Histone H3K9 methyltransferase G9a in oocyte is essential for preimplantation development but dispensable for CG methylation protection

Abstract:
Mammalian histone methyltransferase G9a (also called EHMT2) deposits H3K9me2 on chromatin and is essential for postimplantation development. However, its role in oogenesis and/or preimplantation development is poorly understood. We show that mouse oocytes have large organized chromatin H3K9me2 enriched domains that are rather depleted of CG methylation, which contrasts with their association in embryonic stem and somatic cells. Oocyte-specific disruption of G9a results in reduced H3K9me2 and impaired reorganization of heterochromatin, but only a modest reduction in CG methylation in oocytes. Preimplantation embryos derived from such oocytes show abnormal chromosome segregation and frequent developmental arrest, but surprisingly, their CG methylation is only minimally affected. Furthermore, in both oocytes and 2-cell embryos, G9a depletion has limited impact on the expression of genes and retrotransposons. Our findings illuminate the functional importance of G9a in early development and call into question the proposed role for H3K9me2 in CG methylation protection in zygotes.

Aniruddha Chatterjee
University of Otago

Science Day 1

Global DNA methylation levels regulates PD-L1 expression in melanoma

Abstract:
One of the hallmarks of cancer (including melanoma) is its ability to evade the host immune system by up-regulating immune checkpoints. The programmed death-ligand 1 (PD-L1) receptor is one of most important immune checkpoints and is often upregulated in cancer cells. Recent evidence suggests that in melanoma, the patients with PD-L1 expression and absence of tumour infiltrating lymphocytes (TILs) (i.e. "constitutive PD-L1 or PD-L1CON") show worse response rates and have a worse prognosis than patients with PD-L1 expression and the presence of TILs (i.e. "inducible PD-L1 or PD-L1IND"). However, how PD-L1 expression is regulated in melanoma cells remains elusive. Understanding the mechanisms of how PD-L1 is regulated is important for predicting responses for anti-PD-L1 treatment and for developing new combinatorial therapies. We hypothesised that genes and transcription factors involved in pathways that normally regulate PD-L1 expression are either silenced (by hypermethylation) or activated (by hypomethylation) by epigenetic mechanisms in melanoma. As a result, expression or repression of PD-L1 in these cancer cells is determined by their epigenetic status. To address these questions, we have generated whole-genome scale DNA methylomes (using reduced representation bisulfite sequencing, RRBS) and transcriptomes (RNA-Seq) for 12 patient derived melanoma cell lines (six PD-L1IND and six PD-L1CON). We discovered extensive global hypomethylation in the constitutive lines, particularly pronounced in intergenic repeat regions and gene bodies. A high proportion of hypomethylated regions exhibited dichotomous methylation patterns indicating a common regulatory mechanism between the inducible and constitutive groups. RNA-Sequencing data indicated that the hypomethylated state of the PD-L1CON cells was correlated with higher upregulation of the differentially expressed genes at a global-scale and the upregulated genes were associated with several cancer hallmark properties. The upregulated genes exhibited expression signatures of viral mimicry and cytosolic sensing of dsRNA genes similar to what has been observed after DNA methyltransferase inhibitor (DNMTi) treatment in cancers. Further, our analysis reveals that expression of a subset of epigenetic regulator genes were strongly correlated with PD-L1 expression and global methylome status in PD-L1CON and PD-L1IND cell lines. Finally, we show inhibition of DNA methylation resulted in increased PD-L1 transcription in the PD-L1IND cell lines, demonstrating that global hypomethylation mediated changes regulate PD-L1 expression in melanoma. We believe these results are the first to show that DNA methylation levels play a role in regulating PD-L1 on melanoma and suggest they may have important implications for combined treatments targeting DNA methylation (DNMTi) and PD1/PD-L1 (anti-PD1 antibodies).

Huating Wang
The Chinese University of Hong Kong

Science Day 1

MyoD induced enhancer RNA interacts with hnRNPL protein via CAAA motif to activate target gene transcription during myogenic differentiation

Abstract:
Emerging evidence supports active roles of enhancer RNAs (eRNAs) in regulating target gene expression but our understanding of the underlying mechanisms remains incomplete. Here, we study eRNA regulation and function using skeletal myoblast differentiation as a paradigm. We provide a panoramic view of enhancer transcription and dynamics during myogenic differentiation and first categorization of eRNAs by integrating GRO-seq and RNA-seq data. We demonstrate the essential role of master transcription factor MyoD in activating eRNA production. Subsequent in depth dissection of seRNA-1 and -2 uncovered that seRNAs can promote myogenic differentiation in vitro and in vivo. Mechanistically, we found these seRNAs control the transcription of target genes by specifically binding to heterogeneous nuclear ribonucleoprotein L (hnRNPL) and modulate hnRNPL dosage on the target promoter. A CAAA tract on seRNA-1 was further identified to be essential in mediating seRNA-1/hnRNPL binding and function. Disruption of seRNA-hnRNPL interaction attenuates Pol II and H3K36me3 deposition at the target genes, in coincidence with the reduction of their transcription. Furthermore, analyses of hnRNPL binding transcriptome-wide reveal its association with eRNAs is a general phenomenon in multiple cells. Collectively, we propose that eRNA-hnRNPL interaction represents a novel mechanism contributing to target mRNA activation.

Jian Yan
City University of Hong Kong

Science Day 1

Allelic Binding of Human Transcription Factors to Genetic Variants that Are Associated with Type-2-Diabetes Predisposition

Abstract:
Molecular characterization of Genome Wide Association Study (GWAS) reveals key gene regulation and biological mechanism of genetic diseases. To our best knowledge, only a small portion of SNPs alter protein coding while the vast majority of them are located in non-coding regions. Recent studies have linked some risk loci to interfering transcription factor (TF) binding sites and DNA accessibility. However, most risk SNPs have not been systematically characterized for TF binding. Here, we adapted a modified HT-SELEX pipeline using 40-bp human genomic sequences as ligands to quantitatively analyze allelic imbalanced TF binding at 95,886 variants that are most likely associated with type-2-diabetes, termed as SNP-SELEX. In total, we reported allelic binding specificities for 270 distinct human TFs. The SNP-SELEX result showed good correlation with in vivo TF allelic binding and also explained allelic imbalanced DNA accessibility and enhancer activity. In order to expand such analysis to genome wide level, we adapted deltaSVM method to develop prediction models for 125 TFs of their allelic binding that were trained with SNP-SELEX data. We found that the model recovered most of the *in vivo* allelic binding without compromising specificity and outperformed other models. When applying the 125 models to all common human SNPs and intersecting them with known GWAS hits, we could identify candidate TFs that might confer the phenotype. For example, a few TFs, including HLF and MAFG, are found to regulate LDL cholesterol level and confer T2D predisposition. Finally, we performed high resolution in situ Hi-C in HepG2 cells and normal human islet tissues that enabled us to identify the target genes of risk SNPs and their associated TFs. The comprehensive dataset would be a valuable resource to uncover the discipline behind GWAS.

Lu Gan
National University of Singapore

Science Day 2

Structural cell biology of eukaryotic nuclei

Abstract:
Our long-term goal is to understand the relationship between the 3-D arrangement of chromatin and its many functions. Traditional structural and imaging approaches are ineffective because the nucleus (and the cell) is fragile and structurally heterogeneous. We therefore use electron cryotomography as our main tool to determine the structure of nuclei at molecular resolution, in a life-like state. This approach can reveal the locations and in some cases the orientations of nucleosomes and other large macromolecular complexes. As a first step, we have recently shown how HeLa chromatin is arranged at the mono-, di-, and tri-nucleosome levels in the nuclear-envelope proximal heterochromatin. The nucleosomes have asymmetric linker DNA and in many cases are associated with as-yet unidentified macromolecular complexes. Remapping of the average nucleosomes back into their in situ positions and orientations reveals the first levels of higher-order chromatin: di-nucleosomes and tri-nucleosomes. Chromatin at this level follows a zigzag path and is also irregular. These findings are consistent with the emerging theme that nucleosomes in situ do not compact into ordered motifs of any kind. https://www.biorxiv.org/content/early/2018/05/30/334490

Simon Heath
Centro Nacional de Análisis Genómico

Science Day 2

Inference of genomic spatial organization from a whole genome bisulfite sequencing sample

Abstract:
Common approaches to characterize the structure of the DNA in the nucleus, such as the different Chromosome Conformation Capture methods, have not currently been widely applied to different tissue types due to several practical difficulties including the requirement for intact cells to start the sample preparation. In contrast, techniques based on sodium bisulfite conversion of DNA to assay DNA methylation, have been widely applied to many different tissue types in a variety of organisms. Recent work has shown the possibility of inferring some aspects of the three dimensional DNA structure from DNA methylation data, raising the possibility of three dimensional DNA structure prediction using the large collection of already generated DNA methylation datasets. We propose a simple method to predict the values of the first eigenvector of the Hi-C matrix of a sample (and hence the positions of the A and B compartments) using only the GC content of the sequence and a single whole genome bisulfite sequencing (WGBS) experiment which yields information on the methylation levels and their variability along the genome. We train and test our model on 10 samples for which we have data from both bisulfite sequencing and chromosome conformation experiments and our most relevant finding is that the variability of DNA methylation along the sequence is often a better predictor than methylation itself. We then run a prediction on 206 DNA methylation profiles produced by the Blueprint project and use ChIP-Seq and RNA-Seq data to confirm that the forecasted eigenvector delineates correctly the physical chromatin compartments observed with the Hi-C experiment.

Sarah Kimmins
McGill University

Science Day 2

Epigenomic signatures in sperm associated with Body Mass Index (BMI) and Male Infertility

Romain Lambrot, Vanessa Dumeaux, Olusola Sotunde, Karen Lockyear, Natalia Yasovitch, Sophia Zheng, Pamela Kurjanowicz, Trevor Partch, Rose Ghemrawi, Christine Lafleur, Linda Dodds, Amanda MacFarlane, Jacquetta Trasler, Hope Weiler, Clifford Librach, Sergey Moskovtsev and Sarah Kimmins

Abstract:

Introduction: Sperm counts have been declining at an alarming rate with counts being half of what they were 40 years ago. This decline may be attributable to >50% of the population being overweight. In mice and men, the sperm epigenome including histone methylation, has been linked to responses to diet, BMI, fertility and altered reproductive outcomes. Previously we showed that changes in the enrichment of histone methylation in sperm are associated with infertility and poor embryo development. These studies suggest that the sperm epigenome has the potential to be used clinically to assess fertility. Our objective is to compare profiles of histone H3 tri-methylation on lysine 4 (H3K4me3) in sperm from men with differential BMI status (normal vs high) in couples seeking fertility treatment at the CReATe clinic.

Methods: We recruited study participants (n=137) and collected BMI, dietary intake, folate parameters (folate, vitamin B12, homocysteine), vitamin D levels and lifestyle information. Semen quality was analyzed by standard clinical approaches. Men selected for analysis were of either a normal BMI (BMI <25kg/m^2), or overweight and obese (BMI >25 kg/m^2). Men were excluded if they had a DFI>30, were older than 50, smoked and had the TT genotype for the C677T SNP of the MTHFR enzyme. ChIP-seq targeting H3K4me3 was performed on a subset of men including patients with a normal (n=24) or increased BMI (n=24).

Results: All men were folate sufficient (320 – 1090 nmol/L), with 62% having higher levels of RBC folate at >1090 nmol/L. Interestingly, about half of the men in our study exhibited low vitamin D levels (total serum 25-hydroxyvitamin D), and were classified as vitamin D insufficient (32%, 30 to <50nmol/L), or deficient (20.6%, 30nmol/L). Analysis of ChIP-seq data confirmed that we obtain robust and high quality H3K4me3 profiles from sperm. We have identified specific differences in H3K4me3 enrichment between men of normal and increased BMI and fertility status. The aim is to identify epigenomic signatures associated with BMI, fertility and clinical outcomes (e.g. embryo quality, pregnancy).

Conclusions: Lifestyle factors such as obesity may impact the sperm epigenome and thus influence male fertility and clinical outcomes.

Funded by the Canadian Institutes of Health Research

Marcel Schulz
Goethe University Frankfurt

Science Day 2

A supervised method for enhancer identification and linkage to target genes

Abstract:
Understanding transcriptional regulation is a major goal of computational biology. Especially enhancers are essential regulators driving cellular development. Enhancers can be identified experimentally, e.g. using enhancer RNAs, ChIP-seq of Histone Modifications (HMs), or Hi-C experiments. However, experimental linkage of enhancers to genes is challenging. Therefore, several computational methods have been proposed to create tissue-specific enhancer maps from epigenetics data. A common strategy to de-novo link tissue-specific enhancer regions to genes is to unify DNase-hypersensitive-sites (DHS) across several samples. Subsequently, the unified regions are linked to nearby genes; either pure distance based or using a correlation test between the epigenetic signal and the expression of the possible target gene. Also, integrative efforts are made to combine known enhancers in curated databases, such as GeneHancer. Via gene-expression modeling, we show that these approaches are limited in accounting for the distinct regulatory landscape of genes and thus lead to suboptimal enhancer-gene associations. We developed an unbiased, peak-independent, supervised method called STITCH to identify and to link regulatory regions to genes. We apply STITCH on a uniformly reprocessed dataset comprising paired DNase1-seq and RNA-seq data for 215 human samples from IHEC. Within STITCH, we consider the epigenetic-signal of all samples jointly using the minimum description length principle to identify regions exhibiting a signal variation related to the expression of a distinct gene. In contrast to purely peak-based approaches, no sample specific information is lost. STITCH finds associations over large genomic intervals, e.g. 1 mb, leading to an extensive catalog of enhancer-gene interactions. STITCH is compared against GeneHancer and two approaches combining DHS sites in other common ways. Enhancers called by STITCH lead to a better performance of gene-expression models than both GeneHancer regions and peak-based approaches. Additionally, STITCH enhancer predictions cover about 85% of the GeneHancer database, supporting the quality of our predictions. Including data on HMs revealed that the enhancers identified with STITCH and DNase1-seq data are often surrounded by H3K27ac, an established enhancer mark. Furthermore, we show various downstream applications, e.g. how our identified enhancers can be used to link noncoding DNA mutations in cancer to the genes they regulate. STITCH is freely available (https://github.com/SchulzLab/STITCH). Due to an efficient implementation, large data sets comprising hundreds of samples can be processed easily. Thus, we believe that STITCH can pave the way for a better understanding of gene-specific regulation, especially in light of the large amounts of epigenetics data becoming available.

Dongsoo (Stephen) Lee
University of British Columbia

Science Day 2

Characterization of the molecular consequences of CIC-knockout and neomorphic IDH1 R132H mutation on transcriptomic and epigenomic landscapes

Abstract:
CIC, or Capicua, encodes a transcription factor that mediates repression of genes downstream of RAS/MAPK signalling. CIC is found to harbor deleterious somatic mutations in 50-70% of type 1 low grade gliomas (LGG). Type 1 LGGs are a cohort of tumours molecularly defined by the loss of heterozygosity of chromosome arms 1p and 19q and the gain of neomorphic IDH1/2 mutations. Despite the high frequency of mutations in CIC within this tumour type, CIC's putative tumour suppressive role remains to be elucidated. It is also unclear how CIC may cooperate with neomorphic IDH1/2 to promote gliomagenesis. To comprehensively characterize the molecular consequences of CIC loss, we performed RNA-seq, Whole Genome Bisulfite Sequencing, and ChIP-seq on CIC and 6 different histone modifications on isogenic CIC-wildtype (WT) and CIC-knockout (KO) normal human astrocytes. To also investigate the collective effects of CIC deficiency and neomorphic IDH1 on the transcriptome and epigenome, we generated the same dataset in isogenic CIC-WT and CIC-KO astrocytes possessing the IDH1 R132H mutation. Analysis of differentially expressed genes illustrates the enrichment of oncogenic pathways in specifically the CIC-KO, IDH1-R132H cells, supporting a synergistic relationship between CIC loss and IDH1-R132H in driving tumour progression. Integration of CIC binding sites, differential methylation and chromatin states defined by histone modifications are ongoing to unveil the epigenetic mechanisms underpinning the regulatory changes in these isogenic cell line models.

Chloe Wong
King's College London

Science Day 2

Genome-wide DNA methylation profiling identifies convergent molecular signatures associated with idiopathic and syndromic forms of autism in post-mortem human brain tissue

Abstract:
Autism spectrum disorder (ASD) encompasses a collection of complex neuropsychiatric disorders characterized by deficits in social functioning, communication and repetitive behavior. Building on recent studies supporting a role for developmentally moderated regulatory genomic variation in the molecular etiology of ASD, we quantified genome-wide patterns of DNA methylation in 233 post-mortem tissues samples isolated from three brain regions (prefrontal cortex, temporal cortex and cerebellum) dissected from 43 ASD patients and 38 non-psychiatric control donors. We identified widespread differences in DNA methylation associated with idiopathic ASD (iASD), with consistent signals in both cortical regions that were distinct to those observed in the cerebellum. Individuals carrying a duplication on chromosome 15q (dup15q), representing a genetically-defined subtype of ASD, were characterized by striking differences in DNA methylation across a discrete domain spanning an imprinted gene cluster within the duplicated region. In addition to the dramatic cis-effects on DNA methylation observed in dup15q carriers, we identified convergent methylomic signatures associated with both iASD and dup15q, reflecting the findings from previous studies of gene expression and H3K27ac. Cortical co-methylation network analysis identified a number of co-methylated modules significantly associated with ASD that are enriched for genomic regions annotated to genes involved in the immune system, synaptic signalling and neuronal regulation. Our study represents the first systematic analysis of DNA methylation associated with ASD across multiple brain regions, providing novel evidence for convergent molecular signatures associated with both idiopathic and syndromic autism.

Shravanti Rampalli-Deshpande
Institute for Stem Cell Biology and Regenerative Medicine

Science Day 2

Regulation of peripheral heterochromatin domain organization via histone and non-histone protein methylation

Abstract:
Perinuclear heterochromatin is an emerging nuclear domain for driving many normal and disease processes, including development, cancer and aging. Here we uncover a novel role for Euchromatic histone methyltransferases (EHMTs) in tethering heterochromatin to the nuclear periphery (NP). We identified that the EHMTs regulate factors critical for heterochromatin organization and stability of LaminB1 via its methylation, thereby regulating higher-order chromatin organization at the NP. We observed depletion of EHMTs correlated with loss of methylated LaminB1 and peripheral heterochromatin in aging primary human fibroblasts. Restoration of EHMT expression reverts peripheral heterochromatin defect in aged cells. Collectively our studies elucidated a new mechanism by which EHMTs regulate heterochromatin domain organization and explains its impact on fundamental changes associated with the intrinsic aging process.

Chunhui Hou
Southern University of Science and Technology

Science Day 2

Systematic 3D Genome Architecture Analysis in Xenopus tropicalis

Abstract:
Chromatin architecture is critical in regulating gene expression. The de novo establishment of 3D chromatin structures differs in different model systems around zygotic genome activation (ZGA) during early embryogenesis. We examined the chromatin conformation during Xenopus tropicalis embryogenesis and in terminally differentiated blood and liver cells. Different from the relative fast emergence of topologically associating domains (TADs) at ZGA in other models, TADs and chromatin compartments emerge relatively late after ZGA and consolidate continuously through a very long time window in Xenopus. Knocking down of RNAPII severely delays the embryo development in a morpholino dose-dependent manner. In embryos developed beyond ZGA, TADs seem to be enhanced compared to wild type embryos at the same developmental stage. However, in embryos stalled at ZGA, chromatin remains unstructured. Knocking down of cohesin Rad21 revealed that cohesin is required for the de novo establishment of TADs as in in vitro cell experiment. Cohesin is also indispensable for embryo development. Taken together, our results suggest that transcription is required for de novo chromatin architecture establishment at least through the providing of protein translation templates. Surprisingly, CTCF knocking down affects neither the embryo development nor the de novo establishment of chromatin structures. We also noticed that when TAD weakened, compartment seems to be enhanced. In normal blood and liver cells, the TAD structure is conserved but the compartmentalization is significantly stronger in blood cells whose nucleus size is obviously smaller, which suggests that a reverse correlation exists between nucleus size and chromatin compartmentalization. Our results offer mechanical insights into the de novo establishment of vertebrate genome architecture and reveal unexpected relation between nuclear size and genome compartmentalization.

Yann Joly
McGill University

Science Day 2

The Ethics of Epigenetics Research: An Overview of the Activities of the IHEC Bioethics Workgroup

Yann Joly, on behalf of the IHEC Bioethics Workgroup

Abstract:
The IHEC Bioethics Workgroup is a multidisciplinary, international group of researchers, interested in the relationship between epigenetic science, ethics, society and public policy. It was assembled in 2012 to advise IHEC in the development and update of its governance policies, and to investigate emerging ethical issues raised by epigenetics research. This presentation will present an overview of some of the key research topics investigated by IHEC during the 2013-2019 period. These issues include: 1) epigenetic data sharing and privacy, 2) return of results and incidental findings in epigenetic research, 3) portrayal of epigenetic research in social science academic literature and 4) epigenetic discrimination. The presentation will conclude with a critical assessment of the opportunities and challenges met by ethics groups, embedded within a larger scientific consortium in 'omics' research.

Charles Dupras
McGill University

Science Day 2

Epigenetic Discrimination: Emerging Applications of Epigenetics Calling for Ethical Scrutiny

Abstract:
Over more than two decades, various policies have been adopted worldwide to restrict the use of individual genetic information for non-medical reasons by third parties and prevent 'genetic discrimination'. In this presentation, I will bring attention to the growing interest for individual epigenetic information by insurers and forensic scientists. I will question whether such interest could lead to 'epigenetic discrimination' – the differential adverse treatment or abusive profiling of individuals or groups based on their actual or presumed epigenetic characteristics – and argue that we might already be facing the limitations of recently adopted normative approaches against genetic discrimination. First, I will highlight some similarities and differences between genetic and epigenetic modifications, and stress potential challenges to regulating epigenetic discrimination. Second, I will argue that most existing normative approaches against genetic discrimination fall short in providing oversight into the field of epigenetics. We conclude with a call for discussion on the issue, and the development of comprehensive and forward-looking preventive strategies against epigenetic discrimination.

Reference: Dupras C., L. Song, K. M. Saulnier and Y. Joly (2018) "Epigenetic Discrimination: Emerging Applications of Epigenetics Pointing to the Limitations of Policies Against Genetic Discrimination". Frontiers in Genetics 8(202): 1-6. http://dx.doi: 10.3389/fgene.2018.00202

Sebastian Preissl
University of California, San Diego

Science Day 2

Single nuclei chromatin accessibility analysis reveals epigenetic heterogeneity of mouse brain regions

Abstract:
Transcriptional regulatory regions in the genome including promoters and distal acting enhancers play fundamental roles for development and disease. These genomic regions can be identified by the presence of open chromatin as measured by ATAC-seq (Assay for transposase-accessible chromatin using sequencing). However, heterogeneity of primary tissues poses a significant challenge in mapping the precise chromatin landscape in specific cell types. Therefore, we optimized a combinatorial barcoding-assisted single-cell assay for transposase-accessible chromatin for use on flash-frozen primary tissue samples (single nuclei ATAC-seq, snATAC-seq). We applied the methodology to deconvolute the cellular composition of the mouse forebrain, defined cell-type-specific transcriptional regulatory sequences and inferred potential master transcriptional regulators. In addition, we identified cell-type specific enrichment of gene sequence variants associated with human disease in genome-wide association studies (GWAS). For example, regulatory elements accessible in the microglia population were highly enriched for Alzheimer's disease associated risk variants. Further process optimization including liquid handling robotics enables us now to robustly generate libraries for more than 5,000 single nuclei chromatin accessibility profiles in a single experiment. We use this approach to map the epigenetic heterogeneity in distinct regions dissected from the adult mouse brain. Initial analysis of more than 70,000 nuclei isolated from the primary motor cortex revealed >30 cell populations corresponding to all major neuronal and non-neuronal cell types. I will present the most recent insights from our analysis. This work is supported by NIH Grant MH114831.

Science Days Posters

Please note	
Odd number:	Oct 27 Science Day 1
Even number:	Oct 28 Science Day 2

	Presenter	Institution	Title
1	Wan Kin Au Yeung	Kyushu University	Histone H3K9 methyltransferase G9a in oocyte is essential for preimplantation development but dispensable for CG methylation protection
2	Shravanti Rampalli-Deshpande	Institute For Stem Cell Biology and Regenerative Medicine	Regulation of peripheral heterochromatin domain organization via histone and non-histone protein methylation
3	Hiroshi Kimura	Tokyo Institute of Technology	Chromatin Integration Labeling Technology (ChILT): an immunoprecipitation-free method for low-input epigenomic profiling
4	Chunhui Hou	Southern University of Science and Technology	Systematic 3D Genome Architecture Analysis in Xenopus tropicalis
5	Huating Wang	The Chinese University of Hong Kong	MyoD induced enhancer RNA interacts with hnRNPL protein via CAAA motif to activate target gene transcription during myogenic differentiation
6	Chloe Wong	King's College	Genome-wide DNA methylation profiling identifies convergent molecular signatures associated with idiopathic and syndromic forms of autism in post-mortem human brain tissue
7	Aniruddha Chatterjee	University of Otago	Global DNA methylation levels regulates PD-L1 expression in melanoma
8	Chloe Wong	King's College	DNA methylome marks of exposure to psychosocial stress during adolescence: Analysis of a novel longitudinal MZ discordant twin study
9	Jian Yan	City University of Hong Kong	Allelic Binding of Human Transcription Factors to Genetic Variants that Are Associated with Type-2-Diabetes Predisposition
10	Marcel Schulz	Goethe University Frankfurt	A supervised method for enhancer identification and linkage to target genes
11	Alfred Cheng	The Chinese University of Hong Kong	Integrative chromatin profiling of primary liver cancers reveals epigenetic vulnerability for effective combination immunotherapy
12	Sarah Kimmins	McGill University	Epigenomic signatures in sperm associated with Body Mass Index (BMI) and male infertility
13	Ralf Jauch	The University of Hong Kong	Enhancing cellular reprogramming by directed factor evolution
14	Sebastian Preissl	University of California, San Diego	Single nuclei chromatin accessibility analysis reveals epigenetic heterogeneity of mouse brain regions

15	Ruey Su	Public Health Agency of Canada, University of Manitoba	High HDAC Activity of Female Cervical Cells is Predictive of an Elevated Level of Immune Activation that is Favorable for HIV-1 Acquisition
16	Charles Dupras	McGill University	Epigenetic Discrimination: Emerging Applications of Epigenetics Calling for Ethical Scrutiny
17	Karen Wing Yee Yuen	The University of Hong Kong	Epigenetic Regulation of Centromere Establishment
18	Yann Joly	McGill University	The ethics of epigenetics research: An overview of the activities of the IHEC Bioethics Workgroup
19	Atsushi Shimizu	Iwate Medical University	Iwate Tohoku Medical Megabank Epi-Genome Cohort: Mission and Progress reports
20	Lu Gan	National University of Singapore	Structural cell biology of eukaryotic nuclei
21	Nina Gasparoni	Saarland University	Single Cell Omics Germany: A network for single cell researchers
22	Simon Heath	Centro Nacional de Análisis Genómico (CNAG-CRG)	Inference of genomic spatial organization from a whole genome bisulfite sequencing sample
23	Nina Gasparoni	Saarland University	LifeTime: Understanding cells to revolutionize healthcare
24	Simon Heath	Centro Nacional de Análisis Genómico (CNAG-CRG)	gemBS - high throughput processing for DNA methylation data from Bisulfite Sequencing
25	Charles Imbusch	German Cancer Research Center (DKFZ)	Deconvolution of DNA methylation Identifies Subgroups Across Sarcoma Subtypes
26	Charles Breeze	Altius Institute for Biomedical Sciences, Seattle	Integrative analysis of mouse developmental regulatory DNA
27	David Bujold	McGill University	IHEC Data Portal 2018 update: community hubs and integration with Galaxy
28	Daniel Zerbino	EMBL - European Bioinformatics Institute	Functional annotation of GWAS regulatory variants powered by Ensembl
29	Hyo Sik Jang	Washington University in St. Louis	Epigenetic dynamics in zebrafish pigment cell fate
30	Gilles Gasparoni	Saarland University	DNA Methylation analysis on purified neurons and glia dissects age and Alzheimer's disease-specific changes in the human cortex
31	Michelle Moksa	University of British Columbia	Technology Development at the Centre for Epigenome Mapping Technologies
32	Jason Hilton	Stanford University	The ENCODE Pipeline Architecture: Reproducible and portable analysis tools for ChIP-seq, RNA-seq, DNase-seq, ATAC-seq, HiC, ChIA-PET, and whole-genome bisulfite experiments
33	Pierre-Etienne Jacques	Université de Sherbrooke	Epigenomic labeling prediction
34	Nakul Shah	Washington University in St. Louis	Widespread transposable element-driven oncogene expression in cancers
35	Sabrina Tam	Hong Kong University of Science and Technology	Dysregulation of Human Endogenous Retroviruses in Systemic Lupus Erythematosus
36	Rashedul Islam	University of British Columbia	Molecular characterization of the role of RUNX1 in Notch signaling in T-cell Acute Lymphoblastic Leukemia (T-ALL)

37	Shohei Komaki	Iwate Medical University	Enrichment of iMETHYL, a multi-omics genome browser, by eQTL/eQTM/mQTL analyses
38	Thomas Juettemann	EMBL - European Bioinformatics Institute	EpiRR – a registry of reference epigenomes
39	Wenjie Sun	Genome Institute of Singapore	Transcriptomic and Epigenomic Study of Anti-depressant Drug Action in Primate Brain
40	Wilson Tan Lek Wen	Genome Institute of Singapore	Genetic Determinants of Enhancer Epigenetics in Human Hearts
41	Yang Gao	The University of Hong Kong	DNA Methylation Patterns in Normal Tissue Correlate more Strongly with Breast Cancer Status than Copy-Number Variants
42	Abdul Rahman Salhab	Saarland University	A comprehensive analysis of 195 DNA methylomes reveals shared and cell specific features of partially methylated domains
43	Kai-Wei Chang	National Taiwan University	Role of monoallelic expression in autistic spectrum disorders
44	Kibaick Lee	Korea National Institute of Health	A systematic analysis of epigenomic pattern in human pancreatic islets from healthy and T2D donors
45	Sheng Zhong	University of California, San Diego	RNAs as proximity labeling media for identifying nuclear speckle positions relative to the genome
46	Mingqiang Wang	The Chinese University of Hong Kong	Hybrid De Novo Assembly of Branchiostoma belcheri Beihai Amphioxus Genome
47	Zhonghua Liu	The University of Hong Kong	Large-Scale Mediation Effect Signal Detection in Genome-wide Epigenetic Studies
48	In-Uk Koh	Korea National Institute of Health	Korea Epigenome Project (KEP): Epigenome datasets for chronic disease researches
49	Vikas Malik	Guangzhou Institutes of Biomedicine and Health-Chinese Academy of Sciences	The cis-regulatory logic of reprogramming competent and incompetent POU transcription factors
50	Lihua Hu	Nanchang University	miR-3940-5p enhances homologous recombination after DSB in Cr(VI) exposed 16HBE cell
51	Guiping Hu	Peking University	Cr(VI)-induced methylation and down-regulation of DNA repair genes and its association with markers of genetic damage among workers and 16HBE cells

1. **Wan Kin Au Yeung:** ***Histone H3K9 methyltransferase G9a in oocyte is essential for preimplantation development but dispensable for CG methylation protection***
   -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

2. **Shravanti Rampalli-Deshpande:** ***Regulation of peripheral heterochromatin domain organization via histone and non-histone protein methylation***
   -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

3. **Hiroshi Kimura:** ***Chromatin Integration Labeling Technology (ChILT): an immunoprecipitation-free method for low-input epigenomic profiling***
   -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

4. **Chunhui Hou:** ***Systematic 3D Genome Architecture Analysis in Xenopus tropicalis***
   -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

5. **Huating Wang:** ***MyoD induced enhancer RNA interacts with hnRNPL protein via CAAA motif to activate target gene transcription during myogenic differentiation***
   -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

6. **Chloe Wong:** ***Genome-wide DNA methylation profiling identifies convergent molecular signatures associated with idiopathic and syndromic forms of autism in post-mortem human brain tissue***
   -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

7. **Aniruddha Chatterjee:** ***Global DNA methylation levels regulates PD-L1 expression in melanoma***
   -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

8. **Chloe Wong:** ***DNA methylome marks of exposure to psychosocial stress during adolescence: Analysis of a novel longitudinal MZ discordant twin study***
   Radhika Kandaswamy, Eilis Hannon, Georgina Mansell, Ben Williams, Joe Burrage, Susanna Roberts, Andrea Danese, Jonathan Mill, Louise Arseneault, Helen Fisher, Chloe Wong.

   An emerging body of evidence has shown that individuals exposed to psychosocial stressors have different epigenetic fingerprints compared to those exposed to no/minimal stressful events. Limited conclusions can be drawn from previous findings as these relied on adult retrospective reports of stress or trauma, unusual clinical groups and relatively small samples. The current study proposes a powerful and sensitive longitudinal design involving discordant MZ twins to ascertain a 'purer' impact of psychosocial stress on the epigenome controlling for genetic variation, age, sex and shared environmental exposures. We generated genome-wide DNA methylation (DNAm) profiles for 118 MZ twin pairs from the E-Risk longitudinal study with no exposure to severe stress in childhood, but discordant for stress during adolescence (N=62), concordant for exposure to stress during adolescence (N=28), and no exposure to stress during adolescence (N=28) using buccal

DNA collected at ages 5,10 and 18 and the Illumina Infinium EPIC array. Methylation profiles from age-18 blood derived DNA were also generated and compared to those of age-18 buccal derived DNA to interrogate the impact of psychosocial stress on multiple tissue sources. All data pre-processing and downstream statistical analyses were performed using R statistical packages. The longitudinal epigenetic trajectories associated with exposure to adolescent psychosocial stress were interrogated using linear regression with cluster-robust standard errors. Additionally, the interaction term was added to the model to explore the interaction between time and exposure on the DNAm profile of twins. Preliminary analysis of the repeated measures buccal samples revealed site-specific differential DNAm associated with exposure to adolescence psychosocial stress at multiple CpG sites located across the epigenome including genes involved in stress response pathway and inflammation. Tissue specific differential methylation at CpG loci associated with stress exposure was also evident in our blood and buccal epigenome analysis. Findings from this study point towards a tissue specific stress-related EWAS signature, thereby, contributing to the understanding of the biological mechanisms underlying psychosocial stress exposure during key stages of development.

9. **Jian Yan:** ***Allelic Binding of Human Transcription Factors to Genetic Variants that Are Associated with Type-2-Diabetes Predisposition***
   -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

10. **Marcel Schulz:** ***A supervised method for enhancer identification and linkage to target genes***
    -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

11. **Alfred Cheng:** ***Integrative chromatin profiling of primary liver cancers reveals epigenetic vulnerability for effective combination immunotherapy***
    -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

12. **Sarah Kimmins:** ***Epigenomic signatures in sperm associated with Body Mass Index (BMI) and male infertility***
    -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

13. **Ralf Jauch:** ***Enhancing cellular reprogramming by directed factor evolution***
    -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

14. **Sebastian Preissl:** ***Single nuclei chromatin accessibility analysis reveals epigenetic heterogeneity of mouse brain regions***
    -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

15. **Ruey Su:** ***High HDAC Activity of Female Cervical Cells is Predictive of an Elevated Level of Immune Activation that is Favorable for HIV-1 Acquisition***

    The presence of inflammation at the vaginal mucosae, an interface of host cells and commensal microbes is tightly associated with increased susceptibility to sexually

transmitted infection. Microbial metabolites, such as lactate, butyrate and H_2O_2 have been shown to alter cellular histone deacetylase (HDAC) activity and affect regulation of gene expression. This study evaluated the role of HDAC activity in cervical mononuclear cells (CMC) of Kenyan HIV-seronegative female sex-workers (FSWs) in regulating pro-inflammatory genes to affect cellular susceptibility to acquisition of HIV-1. Higher HDAC activity in CMC was found in FSWs who were highly susceptible to HIV-infection, and had elevated pro-inflammatory cytokine/chemokine in the matching cervicovaginal lavage (CVL). These pro-inflammatory chemokine, including CXCL9, CXCL10, IL-1β, IL-6, IL-8, IFN-γ, IL-12, CCL2, CCL3, MIP1β and TNF-α were key mediators of leukocyte-recruitment and cellular activation, based on pathway analysis data. Furthermore, we showed in vitro that reducing HDAC activity in CMC from the FSWs resulted in significant decreases (>90%) in HIV-replication, due to, at least in parts, the impaired up-regulation of IRF-1, which is the key transcription regulator to productive HIV-1 infection and trans-activation of IFN-γ stimulated genes, such as IL-12 and CXCL10. This study further evaluated the potential contribution of vaginal H2O2 in modulating baseline mucosal immunity, by inhibition of cellular HDAC activity. CVL from matched FSWs were assayed for baseline H_2O_2 concentrations. CVL H_2O_2 level was lower in the HIV-susceptible FSWs, inversely correlated with the HDAC activity in CMC of all samples examined ($r= -0.9$, $p<0.001$) and with the CVL levels of α- and β-chemokine that mediate the recruitment of immune cells. Together, these suggest that the baseline CVL H_2O_2 level and CMC HDAC activities may be one of the key influences over the vaginal susceptibility to HIV-1 infection via regulating the homeostasis of cytokine/chemokine production and hence, the recruitment and activation of leukocytes to and at the mucosal site, in the absence of infection or immune activation.

16. **Charles Dupras:** ***Epigenetic Discrimination: Emerging Applications of Epigenetics Calling for Ethical Scrutiny***

-> Abstract available in the Selected Short Talks and Rapid Fire Talks section

17. **Karen Wing Yee Yuen:** ***Epigenetic Regulation of Centromere Establishment***

The centromere is a specialized genomic region that directs chromosome segregation during cell division. The centromere assembles the kinetochore, which binds to microtubules emanating from opposite spindle poles. Most functional centromeres in eukaryotes are epigenetically marked by the centromere-specific histone H3 variant CENP-A. Due to its crucial function, the existing centromere is usually maintained stably at the same position through mitotic cell cycles and generations. However, new centromeres, known as neocentromeres, can form on ectopic regions when the original centromere is inactivated or lost during chromosomal rearrangements, and also on exogenously introduced DNA, forming artificial chromosomes. Moreover, CENP-A, as an essential protein for cell division, is overexpressed in many cancer cells, potentially causing ectopic, new centromeres. However, how centromere is established or inactivated is not clear and is technically challenging to explore. Injecting DNA into the germline of Caenorhabditis elegans leads to formation of artificial chromosomes in embryos, which establish de novo centromeres and segregate autonomously within a few cell cycles, much more efficiently

than human artificial chromosomes. Using this in vivo, real-time model, we elucidated that histone acetylation and transcription create an open chromatin environment that favors centromere establishment, consistent with the results in human artificial chromosomes. We could also modify the DNA composition injected and determine whether sequence features affect de novo centromere formation.

18. **Yann Joly:** ***The ethics of epigenetics research: An overview of the activities of the IHEC Bioethics Workgroup***
    -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

19. **Atsushi Shimizu:** ***Iwate Tohoku Medical Megabank Epi-Genome Cohort: Mission and Progress reports***
    Atsushi Shimizu, Kozo Tanno, Hideki Ohmomo, Shohei Komaki, Tsuyoshi Hachiya, Ryohei Furukawa, Yuh Shiwa, Yoichi Sutoh, Kanako Ono, Ryo Otomo, Natsuko O. Shinozaki, Kotaro Otsuka, Kotaro Oyama, Mamoru Satoh, Ryujin Endo, Yasushi Ishigaki, Akimune Fukushima, Jiro Hitomi, Kiyomi Sakata, Kuniaki Ogasawara, Kenji Sobue, Makoto Sasaki.

    The Great East Japan Earthquake (GEJE) that occurred on March 11, 2011, caused serious damage to the Pacific coastal areas of Japan. The Iwate Tohoku Medical Megabank Organization (IMM) was established with the aim of reconstructing the medical system in disaster-struck areas, supporting disaster victims, and realizing the potential of personalized medicine using epigenomic information. The IMM recruited 31,806 participants in Iwate prefecture between May 2013 and March 2015. Questionnaire forms, physiological function test results, and biological specimens were collected from the participants. Of the IMM participants, 82.6% lived in the tsunami-devastated area, 45.4% experienced the death of close relatives, and 27.3% suffered damage. To establish the system of collecting and utilizing epigenome data in a cohort study, we studied the transportation method of blood samples, intra-individual dynamics of the epigenome, and blood cell composition correction method. Furthermore, we performed genome-wide epigenome analysis for monocytes, CD4-positive T lymphocytes and neutrophils sorted from blood samples of about 100 participants for understanding the diversity of the human blood epigenome. The result demonstrated that many of loci which have large epigenome diversity were highly associated with known epigenome markers. Thus, we designed a custom SureSelect Methyl-Seq probe-set which targets CpG sites with large inter-individual variations. After analyzed 384 PBMC samples with the probe-set, we identified several susceptibility CpG loci associated with BMI and serum levels of pepsinogen II. We are planning to further analysis for an association between epigenome and disease by adding a greater number of samples.

20. **Lu Gan:** ***Structural cell biology of eukaryotic nuclei***
    -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

21. **Nina Gasparoni:** ***Single Cell Omics Germany: A network for single cell researchers***

22. **Simon Heath:** ***Inference of genomic spatial organization from a whole genome bisulfite sequencing sample***
    -> Abstract available in the Selected Short Talks and Rapid Fire Talks section

23. **Nina Gasparoni:** ***LifeTime: Understanding cells to revolutionize healthcare***

24. **Simon Heath:** ***gemBS - high throughput processing for DNA methylation data from Bisulfite Sequencing***

    DNA methylation is essential for normal embryogenesis and development in mammals and can be captured at single base pair resolution by whole genome bisulfite sequencing (WGBS). Current available analysis tools are becoming rapidly outdated as they lack sensible functionality and efficiency to handle large amounts of data now commonly created. We developed gemBS, a fast high-throughput bioinformatics pipeline specifically designed for large scale BS-Seq analysis that combines a high performance BS-mapper (GEM3) and a variant caller specifically for BS-Seq data (BScall). gemBS provides genotype information and methylation estimates for all genomic cytosines in different contexts (CpG and non-CpG) and a set of quality reports for comprehensive and reproducible analysis. gemBS is highly modular and can be easily automated, while producing robust and accurate results. gemBS is released under the GNU GPLv3+ license. Source code and documentation are freely available from www.statgen.cat/gemBS.

25. **Charles Imbusch:** ***Deconvolution of DNA methylation Identifies Subgroups Across Sarcoma Subtypes***

    The role of aberrant DNA methylation and modulation of gene expression is well established in cancers. Sarcomas are cancers of mesenchymal origin and are very heterogeneous in nature. Studying DNA methylation from bulk tumors is complicated for heterogeneous tumor types as different proportions of tumor, stromal and immune cell components may confound molecular classification. In this study, we applied MeDeCom a reference-free deconvolution method to The Cancer Genome Atlas (TCGA) sarcoma DNA methylation data to decompose into latent methylation components (LMC). Additionally, we integrated matching transcriptome data to refine each LMC. We were able to identify subtype specific LMCs, for example for Synovial Sarcoma and uterine and extra-uterine Leiomyosarcoma. Moreover, we discovered a subset of tumor samples that shared an immune cell infiltration signature. We used an independent in-house sarcoma cohort to validate and further investigate the identified molecular profiles. Our workflow shows that integration of methylome-based deconvolution and transcriptome holds a great potential in the discovery of novel biomarkers that may ultimately contribute to our understanding of sarcoma biology and may aid in risk stratification.

26. **Charles Breeze:** ***Integrative analysis of mouse developmental regulatory DNA***

    The fetal stage is a critical phase in mammalian biology, corresponding to key windows in the development of the main organs and tissues. However, relatively little is understood

about the regulation of most mammalian genes during fetal development. To address this challenge, we mapped DHSs across 9 developmental timepoints and 15 tissues in mouse, spanning days 9.5-21 of mouse fetal development. We uncovered 1.8 M DHSs, including 0.5 M DHSs that are specific to fetal development. The majority of the DHSs are tissue-specific and of the tissue-restricted DHSs, a substantial proportion are restricted to specific developmental timepoints. We used these data to characterise TF network formation during development, highlighting key regulators of organ formation in mammals, and characterising the dynamics of the takeover of the chromatin landscapes by these regulators. To compare these regulatory programs to those found during human development, we aligned mouse and human developmental stages using a TF-anchored network, and uncovered regulatory regions and networks linking early development and human disease. We also found strong links between DHS developmental persistence and evolutionary conservation, highlighting a key subset of TFs that present a higher conservation status. Our results constitute a powerful resource for understanding mammalian biology and its relationship to human regulation in development and disease.

27. **David Bujold:** ***IHEC Data Portal 2018 update: community hubs and integration with Galaxy***

The IHEC Data Portal (epigenomesportal.ca/ihec) is the integrative online resource to navigate through datasets produced by the International Human Epigenome Consortium. With over 10,000 human datasets and an average of 250 unique sessions weekly, it is the central access point to visualize and obtain IHEC datasets. In order to increase the quality and accessibility of epigenomic data from IHEC and the community as a whole, we have added many features over the last year, to be included in the October 2018 release of the Portal. First, we are now offering methods to integrate datasets from the community directly into the IHEC Data Portal. By filling an IHEC Data Hub document (e.g. with the Metadator, epigenomesportal.ca/metadator), any lab can publish their epigenomic datasets and make them available in the Portal for download, visualization and analysis, for instance to assess comparability with existing IHEC data. Next, a tightly coupled instance of the Galaxy Framework is also available. By selecting datasets from IHEC and community hubs, users can launch anonymous Galaxy sessions, and import their own datasets in for further analysis along with IHEC data. New services have also been added, such as a reporting tool on datasets and tracks usage by the public. Lastly, many new datasets will be added in the next release, mostly for the hg38 reference assembly, and existing reference epigenomes will include the metadata improvements efforts of the epiMAP re-analysis project. The IHEC Data Portal is hosted by GenAP (genap.ca), and funded under the CEEHRC, by the CIHR and by Genome Quebec, with additional support from Genome Canada. The correlation matrix computation approach was developed by the Université de Sherbrooke, and funded by NSERC. The computing and networking infrastructure, and part of the software development, are provided by Compute Canada and CANARIE.

28. Daniel Zerbino: ***Functional annotation of GWAS regulatory variants powered by Ensembl***

Ensembl is one of the world's leading sources of information on the structure and function of the genome. It brings together genome sequences, genes, non-coding RNAs, known variants, and other data to create an up-to-date, comprehensive and consistent resource. The Ensembl annotations are widely used for the analysis and interpretation of genome data using tools such as the Ensembl Variant Effect Predictor (VEP), which can quickly annotate the known variants of an individual and report on the potential effects of each. However, the analysis of individual variants remains challenging when analyzing the results of genome wide association studies (GWAS). In a majority of cases, the associated variants are not likely causing coding variants, rather potential regulatory variants with weak phenotypic associations. Here, we present how weak associations with causal genes can be detected through a genome-wide and multi-layered integrative analysis. Ensembl's Regulatory Build synthesizes epigenomic datasets produced by large-scale projects such as ENCODE, Roadmap Epigenomics or BLUEPRINT. The resulting regulatory annotation defines biochemically active regions across 68 human cell types and 79 mouse cell types, assigning them a function wherever possible. To support the Regulatory Build, Ensembl maintains the International Human Epigenome Consortium's (IHEC) Epigenome Reference Registry (EpiRR), where large epigenomic consortia collect the metadata describing their datasets across multiple species. Regulatory elements are chiefly of interest because of their effect on gene expression. To gain further insight, Ensembl is developing a database of cis-regulatory interactions that attach them to their target genes; as a first step all of the GTEx summary eQTL data is incorporated and can be accessed and viewed. Having brought all this data together, it is possible to develop advanced functional analysis methods without being constrained by the scale of the data, as we demonstrate with our post-GWAS analysis platform, Postgap. This algorithm compares human GWAS results, as stored in public archives or in a private individual study, to a collection of genomic annotations, many of which are stored in Ensembl, producing a list of putative causal genes along with linked evidence. It optionally integrates a Bayesian colocalization algorithm that stochastically tests possible sets of causal variants at each GWAS peak. Postgap can be run locally or precomputed results can be explored on the Open Targets Genetics Portal.

29. Hyo Sik Jang: ***Epigenetic dynamics in zebrafish pigment cell fate***

Resolving the genetic and epigenetic determinants that drive specific cell fate decision in complex organisms has been a long-standing goal in developmental biology. During zebrafish embryogenesis, a population of multipotent embryonic cells, called neural crest, is responsible for the migration and production of biologically unique cell types, such as neurons, bones and pigment cells. Zebrafish pigment cell differentiation, in particular, provides an attractive model for studying cell fate progression as a single neural crest progenitor engenders all three morphologically distinct pigment types: black melanocytes, yellow xanthophores and shiny iridophores. Nontrivial classical genetic and transcriptomic approaches have revealed essential molecular mechanisms and gene-regulatory circuits that drive neural crest-derived cell fate decisions. However, how the epigenetic landscape

contributes to pigment cell differentiation is poorly understood. In this study, we investigate how the DNA methylation and chromatin dynamics shape the gene-regulatory network to establish cell identity across various developmental stages of zebrafish melanocyte and iridophore differentiation. We report that pigment cell differentiation is characterized mostly by loci-specific demethylation events and a balanced number of opening and closing chromatin regions, which can be shared or cell-type specific between melanocytes and iridophores. By identifying shared dynamics, we can predict the epigenetic landscape of a melanocyte-iridophore intermediate progenitor. Motif enrichment in the epigenetically dynamic regions, or potential cis-regulatory elements, revealed putative transcription factors that are responsible for driving pigment cell identity. Through this effort, in the relatively uncharacterized iridophores, we define a network of transcriptions factors that are predicted to bind to regulatory elements directly upstream of genes linked to small molecule biosynthesis, lipid metabolism and guanine synthesis cycle, which are essential for iridophore function. On-going work focuses on functionally validating cis-regulatory regions, genetically dissecting the putative transcription factor network, and identifying potential pioneering transcription factors in iridophores. In summary, we chart the global changes in the epigenetic landscape during neural crest differentiation into melanocytes and iridophores to identify epigenetic determinants of pigment cell fate.

30. **Gilles Gasparoni:** ***DNA Methylation analysis on purified neurons and glia dissects age and Alzheimer's disease-specific changes in the human cortex***

Gilles Gasparoni, Sebastian Bultmann, Pavlo Lutsik, Theo Kraus, Matthias Riemenschneider, Hans Kretzschmar, Armin Giese, Heinrich Leonhardt, Jörn Walter

Alzheimer's disease (AD) is a wide-spread age-related neurodegenerative disorder with severe decline of memory and cognitive functions. Neuropathologically, AD is classified into Braak stages according to spread of disease marks across brain regions. Despite intense research, molecular causes for AD and its relationship to healthy aging are not well understood. Epigenetic changes have been discussed to promote a deeper understanding of disease-associated dysregulation. Existing epigenetic studies were based on complex tissue samples and reported only small absolute methylation changes that were difficult to be reproduced across studies. Here we followed a novel approach and report the first genome wide (Illumina 450K methylation array) epigenetic AD-analysis based on physically separated neuronal and non-neuronal (glia) nuclei.

31. **Michelle Moksa:** ***Technology Development at the Centre for Epigenome Mapping Technologies***

M Moksa, Q Cao, M Wong, E Su, A Lorzadeh, A Carles, M Bilenky, A Moussavi, S Wang, A Tam, A Mungall, E Chuah, K Tse, K Mungall, R Moore, S Jones, M Marra, M Hirst

Quantitative epigenomic measurements require highly specialized molecular and computational expertise and access to large-scale sequencing and computational resources. The Centre for Epigenome Mapping Technologies (CEMT) is one of two epigenomic mapping centres in Canada contributing to the International Human

Epigenome Consortium (IHEC). The objective of CEMT is to develop and provide access to standardized epigenomic mapping technologies to local, national and international collaborators. In addition to providing production scale library construction, sequencing, and bioinformatics capabilities in the context of an accredited laboratory environment, CEMT continuously optimizes existing and develops new experimental protocols and computational tools to reduce input material requirements, improve the accuracy and sensitivity of epigenomic measurements, and realize cost savings. Here we present highlights of our recent developments including significant improvements in read mappability and CpG recovery in our single cell methylation pipeline (scPBAL; Hui et al., Stem Cell Reports 2018; Knapp et al., Nat Cell Biol. 2018), production deployment of full epigenomic profiling from frozen biopsied tissue (Siu et al., J Endocrinol. 2017), and optimization of meDIP-seq and hmeDIP-seq. Current areas of focus include the development of strand specific methylation measurements in single cells and the development and implementation of CUT&RUN-sequencing.

32. **Jason Hilton:** ***The ENCODE Pipeline Architecture: Reproducible and portable analysis tools for ChIP-seq, RNA-seq, DNase-seq, ATAC-seq, HiC, ChIA-PET, and whole-genome bisulfite experiments***

Consortia like ENCODE need uniform bioinformatics pipelines to avoid technical artifacts that can confound subsequent integrative analyses of experiments done in different labs or at different times. Such artifacts can be introduced when primary data are analyzed using different software versions, parameters, or run platforms. The ENCODE Data Coordinating Center (DCC) have developed a general containerization and deployment architecture for complex bioinformatics pipelines that produces portable, internally-validating compute environments that run identically in cloud, HPC, and local platforms for the analysis of core ENCODE experiments. Docker, Workflow Description Language (WDL), and continuous integration technologies are used to compose commonly-used bioinformatics tools and purpose-built code into portable, tested workflows. These analysis pipelines produce the uniform ground-level annotations ENCODE applies to the epigenome and transcriptome from ChIP-seq, RNA-seq, DNase-seq, ATAC-seq, HiC, ChIA-PET, and whole-genome bisulfite experiments. By virtue of their portability, both the pipelines and the pipeline architecture itself are of value to other consortia and investigators who need to compare their results to ENCODE's or deploy their own tools in portable ways. Through collaborations with the Roadmap Epigenomics project and the International Human Epigenomics Consortium (IHEC), the ENCODE pipelines are used to produce data and annotations that are compatible across consortia. The ENCODE DCC pipelines and architecture are free and open-source and are documented at https://encode-dcc.github.io/wdl-pipelines/. The DCC codebase is at https://github.com/ENCODE-DCC. ENCODE analyses are distributed through the ENCODE Portal at https://www.encodeproject.org/.

33. Pierre-Etienne Jacques: *Epigenomic labeling prediction*

Jonathan Laperle, Simon Hébert-Deschamps, Pierre-Etienne Jacques

Processing files generated by high-throughput sequencing experiments usually involves quality control (QC) steps conducted to ensure the quality of both the raw data and the alignment to the reference genome. Now that large international consortia such as the International Human Epigenome Consortium (IHEC) have generated thousands of high-quality datasets in a large variety of assays and cell types, it becomes possible to develop tools validating that the signal in a user dataset globally corresponds to the expected "experimental descriptors" such as the right biological assay conducted on the appropriate (or similar) cell type. This is an important task as thousands of epigenomic datasets are now being produced each year outside of large consortia, and that failures or problems can be difficult to detect without extensive investigation. We are developing the Epigenomic Labeling Predictor (EpiLaP) tool that uses machine learning methods over large collections of public epigenomic datasets to predict experimental descriptors. These descriptors can then be used to screen new datasets to detect errors such as mislabeling of samples or other technical artifacts in the data. It can also expand the annotation of datasets to other biological insights not present in the original metadata, as it is often the case with data from public repositories such as GEO. EpiLaP is leveraging our expertise of working with the IHEC datasets acquired during the development of our epigenomic Efficient Correlator (epiGeEC) tool (https://epigeec.genap.ca/galaxy). We will present our most recent results on the development of EpiLaP that will only become more powerful as public datasets grow in number.

34. Nakul Shah*: Widespread transposable element-driven oncogene expression in cancers*

The elucidation of mechanisms behind oncogene activation has been a long-standing goal in cancer biology. Genetic mutation, gene amplification, and chromosomal rearrangement are three classic genetic mechanisms that drive cancer progression and identity, but they provide an incomplete explanation for oncogene activation. Transposable elements (TEs) make up half of the human genome and are a rich genetic resource of regulatory sequences. Cryptic regulatory elements within TEs can be epigenetically reactivated in cancer to influence oncogenesis in a process termed onco-exaptation. However, the prevalence and impact of TE onco-exaptation events across cancer types are uncharacterized. Here, we analyzed 7,769 tumors and 625 normal datasets from 15 TCGA cancer types, identifying 260 TE cryptic promoter activation events involving 174 oncogenes across 3,554 tumors. We report widespread, TE-derived cryptic promoters that boost oncogene transcription in all cancer types. Furthermore, we interrogated the AluJb-LIN28B candidate: the genetic deletion of the TE eliminated oncogene expression, while dynamic DNA methylation-controlled promoter activity, illustrating the necessity and sufficiency of a TE for oncogene activation. Collectively, our results characterize the first global profile of TE onco-exaptation and introduce this prevalent phenomenon as another fundamental mechanism for promiscuous oncogene activation and ultimately tumorigenesis in cancer.

35. **Sabrina Tam:** ***Dysregulation of Human Endogenous Retroviruses in Systemic Lupus Erythematosus***

Systemic Lupus Erythematosus (SLE) is a multi-systemic autoimmune disease. Patients with SLE show a wide range of symptoms and are often misdiagnosed. Currently, the underlying cause of the disease is unclear. Intriguingly, dysregulation of Human Endogenous Retroviruses (HERVs), which are ancestral viral transcripts integrated into the human genome, is strongly correlated with this autoimmune diseases. Envelope (Env) or Group-specific antigen (gag) protein from HERVs could serve as neoantigens and cause cross-activation of lymphocytes, known as molecular mimicry. Recently, it was reported that CD4+ T lymphocytes, derived from SLE patients, show significantly lowered DNA methylation levels at HERV-E family long terminal repeats (LTRs), concomitant with increased of HERV-E gag transcripts. This alteration of epigenetic modification may be involved in the development of the autoimmune response. However, the precise mechanisms of how HERVs can trigger the SLE phenotypes remain elusive. My research examines the epigenetic profile of Peripheral Blood Mononuclear Cell (PBMC) in SLE patients and whether HERVs and immune response genes are differentially expressed. I hypothesized that dysregulation of HERVs in SLE patient T cells and B cells contribute to the misrecognition of self-antigen as foreign. Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) and transcriptome sequencing (RNA-Seq) were performed to delineate the genome-wide chromatin states and transcriptomic changes that occur between disease and control individuals' PBMC, CD4+ T and CD8+ T cells. Identification of HERV elements that potentially drive autoimmune responses in SLE could shed lights on the molecular mechanisms of this and other autoimmune diseases.

36. **Rashedul Islam:** ***Molecular characterization of the role of RUNX1 in Notch signaling in T-cell Acute Lymphoblastic Leukemia (T-ALL)***

Rashedul Islam, Catherine Jenkins, Misha Bilenky, Luolan Li, Alireza Lorzadeh, Michelle Moksa, Annaick Carles, Vincenzo Giambra, Sonya Lam, Catherine Hoofd, Miriam Belmonte, Xuehai Wang, Andrew Weng and Martin Hirst

The key hematopoietic transcription factor *RUNX1* is recurrently mutated (15% of cases) in T-ALL leading to a reduction of its DNA-binding affinity and ability to associate with DNA binding partners. Such loss of function mutations suggest that RUNX1 acts as a tumor suppressor during normal T-cell development. However, wild type RUNX1 can also be overexpressed in T-ALL and a subset of pro-oncogenic NOTCH1 target genes are coordinately regulated by RUNX1 and NOTCH1 in human and mouse T-ALL. The molecular mechanisms driving the pathogenic transcriptional signatures associated with RUNX1 alone and in partnership with NOTCH1 are largely unknown. To address this, we knocked down NOTCH1 by pharmacologic inhibition (NOTCH1-KD) and RUNX1 (RUNX1-KD) by lentiviral shRNAs in a T-ALL cell line (KOPTK1) and performed RNA-seq and ChIP-seq for active (H3K4me1, H3K4me3, H3K27ac, H3K36me3) and repressive (H3K27me3, H3K9me3) histone marks in control and knockdown samples. We observed a significant loss of H3K27ac density following RUNX1 knock-down with ~50% of the 52,427 H3K27ac marked regions showing a >= 2-fold loss of H3K27ac density. Genomic regions that lost H3K27ac

density were enriched with RUNX1 and P300 binding sites and associated with genes involved in cell cycle, Notch and other T-cell signaling pathways. 55% (5,895/10,633) of the RUNX1 and NOTCH1 co-occupied genomic regions showed a gain (>= 2 fold) of H3K27me3 density upon NOTCH1-KD and loss of H3K27ac upon RUNX1-KD. This is consistent with a model where RUNX1 acts in cooperation with NOTCH1 to establish and maintain H3K27ac and NOTCH1 evicts H3K27me3 to drive transcriptional activation of RUNX1+NOTCH1 regulated genes. At least 64 NOTCH1 target genes (e.g., HES4, DTX1, MYC etc.) were co-regulated with RUNX1 through synergistic modification of H3K27 residue. In addition, our analysis revealed a mechanistic link between RUNX1 and its role in disrupting the G1-S checkpoint by driving expression of CDC25A. Collectively our analysis provides mechanistic understanding of the pro-oncogenic role RUNX1 in T-ALL.

37. Shohei Komaki: ***Enrichment of iMETHYL, a multi-omics genome browser, by eQTL/eQTM/mQTL analyses***

Shohei Komaki, Yuh Shiwa, Ryohei Furukawa, Tsuyoshi Hachiya, Hideki Ohmomo, Yoichi Sutoh, Mamoru Satoh, Kenji Sobue, Makoto Sasaki and Atsushi Shimizu

DNA methylation (DNAm) is a key transcriptional regulator and can affect subsequent phenotypes. Upon exploring associations between DNAm variations and phenotypes, numerous DNAm signatures associated with diseases or other traits have been reported. However, few public databases on DNAm are available, and epigenetic analyses of DNAm have analyzed a small portion of methylation sites (CpG sites) in the human genome. To contribute to epigenetic studies by providing comprehensive information regarding DNAm, we established a genome browser, iMETHYL. In iMETHYL, users can browse summary statistics of multi-omics data including DNAm (~24 million CpGs), genomic variation (~9 million SNVs), and gene expression (≥14,000 genes). Multi-omics data were obtained from whole-genome bisulfite and RNA sequencing of each CD4+ T lymphocyte, monocyte, and neutrophil collected from 100 Japanese subjects. Using the comprehensive multi-omics data, we further conducted multi-omics analyses for each cell type [expression quantitative trait locus (cis-eQTL), expression quantitative trait methylation (cis-eQTM), and methylation quantitative locus (cis-mQTL) analyses] and made their corresponding coefficients, P-values, and r-squares accessible on iMETHYL. iMETHYL can be considered a reference in epigenetic studies in setting target regions, inferring biological consequences of changes in DNAm, or conducting case-control studies.

38. Thomas Juettemann: ***EpiRR – a registry of reference epigenomes***

Thomas Juettemann, Agathe Delaune, Ilias Lavidas, Myrto Kostadima, Daniel R. Zerbino

The IHEC consortium aims to produce more than 1000 reference epigenomes and integrate them in large analyses such as the EpiMap project. However, the data supporting these epigenomes are stored in a variety of different public archives, depending on the nature of the data and security requirements of the samples. The EpiRR reference registry (https://www.ebi.ac.uk/vg/epirr), acts as a central reporting and accessioning point for this distributed data, providing IHEC partners and third party scientists a single point of entry to locate all IHEC datasets across pubic archives We have

extended EpiRR to store metadata, so as to provide a central location for IHEC metadata too, and opening up the possibility of searching for IHEC datasets in a comprehensive and consistent fashion. In effect, EpiRR can now read the metadata from the public archive submission and distribute it, thus simplifying the the process of IHEC data submissions. Already, we are supporting the EpiMap project by providing regular statistics on the metadata update efforts across all of IHEC.

39. Wenjie Sun: ***Transcriptomic and Epigenomic Study of Anti-depressant Drug Action in Primate Brain***

Fluoxetine, known as an anti-depressant of selective serotonin-reuptake inhibitors group, is frequently used to treat major depressive disorder. However, little is known about the mechanism of the drug and which brain region the drug actions in due to the paucity of unbiased and large-scale studies. In this study, we apply transcriptome and histone acetylome profiling to over 200 postmortem samples from 14 primate brain regions. Hundreds of differentially expressed genes and over 1,000 differentially acetylated regions were identified responding to fluoxetine treatments. Pathway analysis revealed genes involved in neurotransmission affected by the drug. The response to the drug differ in various brain regions.

40. Wilson Tan Lek Wen: ***Genetic Determinants of Enhancer Epigenetics in Human Hearts***

Heat failure (HF) is a leading cause of morbidity and mortality. Regardless of diverse initiating causes, failing heart manifests a consistent signature stress gene expression changes. A HF gene-expression signature implies that common regulatory pathways exist that mediate disease progression. We hypothesize that HF is a genetic, as well as epigenetic disease. Therefore, it is crucial to map the epigenetic landscape in failing human hearts can lead to novel biomarkers discovery for therapeutic purposes. Non-coding genetic variants in heart failure, particularly in regulatory region such as enhancer, are poorly characterized as the majority were undetected by Genome-wide association studies (GWAS). Hence, we have performed H3K27ac chromatin immuno-precipitation (ChIP-seq) on 70 human hearts to map the active cardiac enhancers. Subsequently, we have performed variant-calling using Genome Analysis Toolkit (GATK) to identify Single Nucleotide Polymorphisms (SNP) within these enhancer. To identify novel regulatory causal SNPs in enhancers, we have performed histone acetylation quantitative trait loci (haQTL). As enhancers are typically located kilobases away from genes, we have also generated chromosome conformation capture (hi-C) libraries from human hearts to study the role of haQTL variants to the regulation of interacting genes. Our results will demonstrate the power of transcriptional enhancers in primary tissue to reveal the underlying mechanism of heart failure.

41. **Yang Gao:** ***DNA Methylation Patterns in Normal Tissue Correlate more Strongly with Breast Cancer Status than Copy-Number Variants***

Normal tissue at risk of neoplastic transformation is characterized by somatic mutations, copy-number variation and DNA methylation changes. It is unclear however, which type of alteration may be more informative of cancer risk. We analyzed genome-wide DNA methylation and copy-number calls from the same DNA assay in a cohort of healthy breast samples and age-matched normal samples collected adjacent to breast cancer. Using statistical methods to adjust for cell type heterogeneity, we show that DNA methylation changes can discriminate normal-adjacent from normal samples better than somatic copy-number variants. We validate this important finding in an independent dataset. These results suggest that DNA methylation alterations in the normal cell of origin may offer better cancer risk prediction and early detection markers than copy-number changes.

42. **Abdul Rahman Salhab:** ***A comprehensive analysis of 195 DNA methylomes reveals shared and cell specific features of partially methylated domains***

Abdulrahman Salhab, Karl Nordström, Gilles Gasparoni, Kathrin Kattler, Peter Ebert, Fidel Ramirez, Laura Arrigoni, Fabian Müller, Julia K. Polansky, Cristina Cadenas, Jan Hengstler, Thomas Lengauer, Thomas Manke, DEEP Consortium and Jörn Walter

Background: Partially methylated domains are extended regions in the genome exhibiting a reduced average DNA methylation level. They cover gene-poor and transcriptionally inactive regions and tend to be heterochromatic. We present a comprehensive comparative analysis of partially methylated domains in human and mouse cells, to identify structural and functional features associated with them.
Results: Partially methylated domains are present in up to 75% of the genome in human and mouse cells irrespective of their tissue or cell origin. Each cell type has a distinct set of partially methylated domains, and genes expressed in such domains show a strong cell type effect. The methylation level varies between cell types with a more pronounced effect in differentiating and replicating cells. The lowest level of methylation is observed in highly proliferating and immortal cancer cell lines. A decrease of DNA methylation within partially methylated domains tends to be linked to an increase in heterochromatic histone marks and a decrease of gene expression. Characteristic combinations of heterochromatic signatures in partially methylated domains are linked to domains of early and middle S-phase and late S-G2 phases of DNA replication.
Conclusions: Partially methylated domains are prominent signatures of long-range epigenomic organization. Integrative analysis identifies them as important general, lineage- and cell-type specific topological features. Changes in partially methylated domains are hallmarks of cell differentiation, with decreased methylation levels and increased heterochromatic marks being linked to enhanced cell proliferation. In combination with broad histone marks, partially methylated domains demarcate distinct domains of late DNA replication.

43. Kai-Wei Chang: ***Role of monoallelic expression in autistic spectrum disorders***

Autistic spectrum disorder (ASD) is a neurodevelopmental disorder characterized by dysfunction in social interactions, communication difficulties, and restricted and repetitive behavior. ASD affects one out of every 59 children, and its broad range of symptom imply no single causal mechanism. Several studies reported the pathogenesis of ASD could involve monoallelic gene expression. In particular, dysregulation of imprinted genes, which is expressed in parent-of-origin manner, have been reported associated with patients with ASD. Nevertheless, how genetic imprinting are involved in ASD, and what type of cortical cells do the effect take place remain largely elusive. By cross-breeding C57BL/6J and CAST/EiJ mice with Cre-loxP-mediated conditional tdTomato reporter system, we investigated gene expression profiles of cortical excitatory neuron (CAMK2A+), inhibitory neuron (GAD2+), and astrocytes (GFAP+), which each cell type of interest was reported associated with ASD in different manner. Our preliminary analysis identified over 200 potentially imprinted genes. The identification of known imprinted genes such as Peg3, Meg3, and H13 support our analysis methods in the identification of imprinted genes. Interestingly, our analysis showed most of the candidate genes exhibit cell-type-specific imprinting, and hence implied the cell type associated roles in cortex. On the other hand, the imprinting candidates showed broad range of functional pathways by gene ontology analysis. Further validation and analysis of their neurobehavioral impacts is currently under progress. The anticipated outcomes will illuminate the roles of the parental-of-origin expressed genes in the context of ASD in cell-type specific manner.

44. Kibaick Lee: ***A systematic analysis of epigenomic pattern in human pancreatic islets from healthy and T2D donors***

Type 2 diabetes (T2D) is a chronic progressive disorder associated with environmental factors which may exert their effect on epigenetic changes. The epigenetic mechanisms include histone modification and DNA methylation, resulting in alteration of gene expression. In this study, to find the T2D-related epigenetic markers, we systematically performed the analysis of histone modification and DNA methylation in human pancreatic islets from healthy and T2D donors. The histone modification patterns were correlated with regulatory elements including promoter and enhancer. From histone ChIP-seq data, we identified 23,998 promoter chromatin states composed of active, weak, poised, and repressive promoters as well as 19,308 enhancer chromatin states by chromatin states analysis using ChromHMM. Furthermore, we found that these promoters regions were associated with 2,433 protein coding genes. Next, we investigated the differential methylated loci(DML) from WGBS analysis and differentially expressed gene(DEG) from RNA-seq analysis, which is based on 2,433 protein coding genes annotated as promoter states. Our results showed that 9 DEGs were associated with DML in its promoter region. Moreover, we found that these methylation levels of CpG in promoter tend to correlate with the gene expression. These results may provide an insight into the understanding of T2D through a systematic analysis of epigenomic pattern and also these identified DEGs may suggest the possibility as T2D-related target gene.

45. **Sheng Zhong:** ***RNAs as proximity labeling media for identifying nuclear speckle positions relative to the genome***

It remains challenging to identify all parts of the nuclear genome that are in proximity to nuclear speckles, due to physical separation between the nuclear speckle cores and chromatin. We hypothesized that noncoding RNAs including small nuclear RNA (snRNAs) and Malat1, which accumulate at the periphery of nuclear speckles (nsaRNA [nuclear speckle-associated RNA]), may extend to sufficient proximity to the genome. Leveraging a transcriptome-genome interaction assay (mapping of RNA-genome interactions [MARGI]), we identified clusters of nsaRNA-interacting genomic sequences (nsaPeaks). Posttranscriptional pre-mRNAs, which also accumulate to nuclear speckles, exhibited proximity to nsaPeaks but rarely to other genomic regions. Our combined DNA fluorescence in situ hybridization and immunofluorescence analysis in 182 single cells revealed a 3-fold increase in odds for nuclear speckles to localize near an nsaPeak than its neighboring genomic sequence. These data suggest a model that nsaRNAs are located in sufficient proximity to the nuclear genome and leave identifiable genomic footprints, thus revealing the parts of genome proximal to nuclear speckles.

46. **Mingqiang Wang:** ***Hybrid De Novo Assembly of Branchiostoma belcheri Beihai Amphioxus Genome***

Branchiostoma belcheri, also known as Chinese amphioxus, is the closest living invertebrate relative of the vertebrates. It is widely used as a model system for studying evolutionary developmental biology and the origin of vertebrates. Beihai amphioxus is a subspecies of B. belcheri that inhabits in Guangxi Beihai, China. The previously reported amphioxus genome generated by short reads sequencing technologies are highly fragmentary, which hinders the downstream analysis and further applications. Here we report the sequencing and assembly of the 640Mb Beihai amphioxus genome using a hybrid approach that combined Pacific Bioscience Single Molecule Real-Time (SMRT) long reads and Illumina short reads sequencing technology. Specifically, a total of 80 Gb genomic data, including 20X PacBio long reads from Sequel sequencing platform, 66X paired-end reads and 56X mate-pair reads from Illumina HiSeq 2000 sequencing platforms, were generated to achieve a high-quality Beihai amphioxus genome. We performed hybrid de novo assembly, scaffolding, gap filling and polishing bioinformatics pipeline to obtain the assembly result which contains 92,511 contigs and 28,014 scaffolds with the contig N50 of 12 kb and scaffold N50 of 73 kb, respectively. The assembly genome contains 907 (92.7%, out of 978) BUSCO core genes collected from metazoa_odb9 database, with 796 (81.4%) of them being complete. The assembly result contains 15M repetitive sequences, which contains 67% simply repeats, 11% SINEs and 7% low-complexity sequence, accounting for 2.28% of the whole genome. Evidence-driven gene prediction method based on RNA-Seq data has identified 63,612 transcripts and 44,087 protein coding sequences in this Beihai amphioxus genome. We anticipate the Beihai amphioxus genome would improve our knowledge on the genetic diversity of this species,

meanwhile providing a valuable genetic resource for the scientific community to further understand the vertebrate evolution.

47. **Zhonghua Liu:** ***Large-Scale Mediation Effect Signal Detection in Genome-wide Epigenetic Studies***

In genome-wide epigenetic studies, it is often of scientific interest in assessing the mediator role of DNA methylation in the causal pathway from an environmental/a genetic exposure to a clinical outcome. A commonly used approach of mediation analysis consists of fitting two separate regression models: the mediator model and the outcome model, and then the product of coefficient method was used to integrate information from these two association studies and hypothesis testing was performed using the Sobel's test. In this paper, we show that the Sobel's test is overly conservative for the detection of mediation effect in genome-wide epigenetic studies. We emphasize that the null hypothesis of mediation effect testing is composite and hence imposes great statistical challenges for developing a powerful test to detect mediation effects. In this paper, we propose a divide-aggregate test (DAT) for the composite null hypothesis for the detection of mediation effects in genome-wide epigenetic studies. We first divide the composite null parameter space into three disjoint parts, and propose three separate testing procedures for each part. The DAT is then obtained by aggregating the statistical evidences by weighted average from the three parts with the weights estimated as the proportions of true nulls based on the p-values from the mediator and outcome regression models. We further show that the DAT can outperform the Sobel's test and the joint significance test for the detection of mediation effects in genome-wide epigenetic studies. A fast Monte Carlo correction is also proposed for computing the p-value of the DAT method. Simulation studies show that the type I error rates of the DAT method are well controlled and the powers are compared with the Sobel's test and the joint significance test under a range of practical settings. An application to the Normative Aging Study (NAS) identified putative DNA methylation CpG sites as mediators in the causal pathway from smoking behavior to lung functions.

48. **In-Uk Koh:** ***Korea Epigenome Project (KEP): Epigenome datasets for chronic disease researches***

Korea National Institute of Health (KNIH) as an official member of IHEC launched the KEP (Korea Epigenome Project) in 2012 with the aim of activation of epigenomics research in Korea. The project is subject to produce 50 reference epigenome datasets on Korean chronic disease (diabetes, obesity, chronic kidney disease, etc.) related cells with the participation on IHEC. The project has been producing IHEC’s core datasets with DNA methylation, RNA expression, and ChIP-seq of 6 histone modifications for unravelling epigenomic differences among cell-types of diseases. Until 2017, 37 epigenome data have been released through European Genome Archive (EGA) and data portals of IHEC and KEP with informative metadata. The Researchers interested in these KEP data have applied through KEP’s data access committee on the EGA, and utilized Korean reference epigenome datasets. Processed epigenome datasets in BigWig format have been accessed

both in data portal of IHEC and KEP for thousands times worldwide. Especially, the epigenome research groups of Korean scientists had been organized and utilized datasets targeted for Korean chronic diseases in terms of disease-specificity and analytical techniques. By the end of this year, additional 13 IHEC core datasets will be released on the EGA and webpages, and KEP will complete 50+ Korean reference epigenome datasets for the first phase of project. Disease specific differences in Korean epigenome datasets have been analyzed and will be further in details.

49. Vikas Malik: ***The cis-regulatory logic of reprogramming competent and incompetent POU transcription factors***

Vikas Malik, Dennis Zimmer, Yanpu Chen, Laura Glaser, Veeramohan Veerapandian, Andrew Paul Hutchins, Huating Wang, Jiekai Chen, Sebastiaan Meijsing, Sebastian Pott, and Ralf Jauch

Direct DNA dependent interactions between Sox2 and Oct4 are intimately linked to the maintenance and induction of pluripotency. To decode the relevance of this partnership, we profiled genomic occupancy, chromatin state changes and transcriptional outcome of Oct4, Sox2 and reprogramming incompetent Pit-Oct-Unc (POU) factors. Both Oct6 and Oct4 mutant with destroyed Sox2 dimerization interface ($Oct4^{defSox2}$) support somatic silencing and the mesenchymal to epithelial transition reminiscent to the wild-type protein suggesting both processes to be independent of Oct4-Sox2 interaction. Both Oct4-SK and Oct6-SK cocktails facilitate the chromatin opening whilst $Oct4^{defSox2}$ shows reduced opening. However, $Oct4^{defSox2}$ is able to effectively bind chromatin closed in somatic cells that are marked by composite SoxOct DNA elements thus implying that it's rather Oct4-Sox2 cooperative interaction that facilitate chromatin opening rather than Oct4 alone. Additionally, $Oct4^{defSox2}$ gets cleared from SoxOct elements as reprogramming progresses and the activation of pluripotency genes drastically fails. Both these results suggest that the target search of Oct4 is not depending on a Sox2 co-selection mechanism. Yet, the formation of long-lived enhanceosomes requires effectively formed Sox2/Oct4 dimers. In sum, the chromatin state changes during reprogramming are not the predictive of successful reprogramming but it's rather the lineage safeguard role by Oct4-Sox2 cooperative interaction especially at late stage of reprogramming that determines the success of the process.

50. Lihua Hu: ***miR-3940-5p enhances homologous recombination after DSB in Cr(VI) exposed 16HBE cell***

Hexavalent chromium (Cr(VI)) is a well-recognized human carcinogen, yet the molecular mechanisms by which cause human cancer are still not well understood. MicroRNAs (miRNAs), which are small non-coding RNAs, are involved in carcinogenesis and DNA damage repair. Previous occupational population study showed that hexavalent chromium (Cr(VI)) downregulated plasma miR-3940-5p level, and a low miR-3940-5p level was associated with high XRCC2 expression in lymphocytes, indicating that miR-3940-5p maybe play a protective effect in Cr(VI) induced DNA damage. Here we investigated miR-3940-3940-5p expression and its roles in DNA repair in Cr(VI)-treated 16HBE cells. miR-3940-5p change was detected by qRT-PCR. Rad51 foci formation and double strand break

(DSB) were investigated to assess homologous recombination repair (HR) capacity by Immunofluorescent assay and Neutral Comet assay. XRCC2 expression was also evaluated after miRNA oligonucleotides transfection using Western blot. Cr(VI) treatment suppressed miR-3940-5p level in 16HBE cells. miR-3904-5p mimic downregulated XRCC2 expression. As a result, the formation of Rad51-foci was inhibited and DSB repair was prolonged. The results indicate that miR-3940-5p plays a protective effect in Cr(VI) induced DNA damage.

51. Guiping Hu: ***Cr(VI)-induced methylation and down-regulation of DNA repair genes and its association with markers of genetic damage among workers and 16HBE cells***

Hexavalent chromium [Cr(VI)], which have widespread environmental distribution and originate from natural and anthropogenic sources, are common environmental carcinogen. In recent decades, their contamination has increased dramatically because of continuous discharge in untreated industrial effluents. To examine the mechanism of hexavalent chromium [Cr(VI)]-induced carcinogenesis, a cross-sectional study among workers with or without exposure to Cr(VI) as well as in vitro administration of Cr(VI) among 16HBE cells was conducted to explore the associations between Cr(VI) exposure, methylation modification of DNA repair genes and their expression levels, and genetic damage. Results showed that hypermethylation of CpG sites were observed among both occupationally exposure workers and 16HBE cells administrated Cr(VI). DNA damage level (8-hydroxydeoxyguanosine, 8-OHdG and micronucleus frequency) among Cr(VI)-exposed workers were significantly higher than the control group. Among workers, positive correlation was observed between the methylation level of CpG sites in DNA repair genes including CpG6,7, CpG8, CpG9,10,11 of MGMT, CpG11 of HOGG1; CpG15,16,17, CpG19 of RAD51 and (1) indicator of chromium internal exposure blood Cr concentration, as well as (2) genetic damage markers (8-OHdG and micronucleus frequency). Significant negative association between methylation levels of CpG sites in DNA repair genes and corresponding mRNA was also observed among 16HBE cells. This indicated that Cr(VI) exposure can down-regulate DNA repair gene expression by hypermethylation, which leads to enhanced genetic damage. The methylation level of those CpG sites of DNA repair genes can be potential epigenetic markers for Cr(VI)-induced DNA damage.