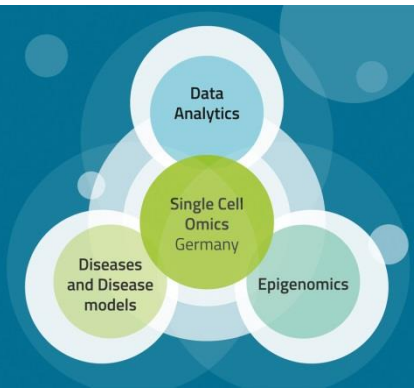
single cell omics germany



Single Cell Omics Germany Workshop:

‚Recent Advances in Single Cell Epigenomics 2022‘

November 16-18, 2022

Überherrn

Program

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

AGENDA

WEDNESDAY, 16 NOVEMBER 2022

11:30 - 12:00 Registration and coffee

12:00 - 12:10 Introduction into SCOG and the scope of the workshop
Jörn Walter (UdS, Saarbrücken)

12:10 - 13:30 Session 1

12:10 - 12:50 NEAT-seq: Simultaneous profiling of intra-nuclear proteins, chromatin accessibility, and gene expression in single cells
Amy Chen (Genentech, San Francisco)

12:50 - 13:30 Mechanisms of cell plasticity in triple negative breast cancers
Céline Vallot (CNRS/Institut Curie, Paris)

13:30 - 14:45 Lunch

14:45 - 16:00 Session 2 – Short talks (15 min each)

Principles of regulatory evolution inferred from early differentiation in primates
Ines Hellmann (LMU, Munich)

Dissecting Ewing sarcoma heterogeneity at a single-cell resolution
Daria Romanovskaia (CeMM, Vienna)

Cell printer assisted single cell ATAC-seq for in-depth epigenetic profiling of rare celltypes
Kathrin Kattler (USAAR, Saarbrücken)

Uncovering dynamic relationships between histone modifications in single cells
Jake Yeung (ISTA, Klosterneuburg)

Annotating scATAC-seq pseudobulk clusters
Aybuge Altay (MPI-MG, Berlin)

16:00 - 16:45 Coffee and Poster session

16:45 - 18:45 Session 3

16:45 - 17:25 Single cell and single molecule computational epigenomics
Maria Colomé-Tatché (Helmholtz Munich)

17:25 - 18:05 Computational epigenomics: from single cells to gene regulatory networks
Ivan Costa (RWTH, Aachen)

18:05 - 18:45 Bayesian modelling of single-cell multi-omics data
Guido Sanguinetti (SISSA, Trieste)

19:30 Dinner

THURSDAY, 17 NOVEMBER 2022

09:00 - 10:20 Session 4

09:00 - 09:40 **Dissecting function in the regulation of developmental trajectories**
Eileen Furlong (EMBL, Heidelberg)

09:40 - 10:20 **Sketching Open and Closed Chromatin, One Cell at a Time**
Giovanni Tonon (IRCCS, Milano)

10:20 - 10:40 Coffee

10:40 - 12:40 Session 5

10:40 - 11:20 **The chromatin landscape during early embryo development at single-cell resolution**
Jop Kind (Hubrecht Institute, Utrecht)

11:20 - 12:00 **Single cell multi-omics landscape of development and ageing**
Wolf Reik (Altos Labs, Cambridge)

12:00 - 13:15 Lunch

13:15 - 14:00 Poster session

14:00 - 15:45 Discussion

15:45 - 16:00 Closing remarks

16:00 End of Meeting (departure)

16:00 - 16:30 Coffee

16:30 - 19:00 EpiScanpy Tutorial

19:30 Dinner

FRIDAY, 18 NOVEMBER 2022

09:00 – 16:00 EpiScanpy Hackathon

16:00 End of Hackathon (departure)

ABSTRACTS – TALKS

SESSION I

Amy Chen (Genentech, San Francisco)

NEAT-seq: Simultaneous profiling of intra-nuclear proteins, chromatin accessibility, and gene expression in single cells

Oligonucleotide-conjugated antibodies have allowed for joint measurement of surface protein abundance along with the transcriptome and/or chromatin accessibility in single cells using high-throughput sequencing. Extending these measurements to gene regulatory proteins in the nucleus would provide a powerful means to link changes in abundance of trans-acting factors to changes in activity of cis-acting elements and expression of target genes. Here, we introduce Nuclear protein Epitope, chromatin Accessibility, and Transcriptome sequencing (NEAT-seq), a technique to simultaneously measure nuclear protein abundance, chromatin accessibility, and the transcriptome in single cells. We apply this technique to profile CD4 memory T cells using an antibody panel for master transcription factors (TFs) that drive distinct helper T cell subsets and regulatory T cells. Using these data, we uncover examples of TFs with regulatory activity gated by three distinct mechanisms: transcription, translation, and regulation of chromatin binding. We also identify regulatory elements and genes associated with each TF, which we use to link a non-coding disease-associated GWAS SNP within a GATA motif to a putative target gene. We observe strong allele-specific chromatin accessibility at the SNP only in cells expressing high levels of GATA3 protein, confirming that the SNP is likely bound and regulated by GATA3. These results illustrate the utility of NEAT-seq for identifying critical steps in regulating expression of proteins of interest, defining candidate enhancer-gene linkages directly regulated by a TF, and validating TF binding and regulation of a specific locus. The combination of regulatory protein quantification with chromatin accessibility and transcript measurements across single cells will also provide a valuable tool for interrogating mechanisms underlying dynamic, TF-driven processes such as cellular differentiation, reprogramming, and transformation.

Céline Vallot (CNRS/Institut Curie, Paris)

Mechanisms of cell plasticity in triple negative breast cancers

The dynamic nature of chromatin and transcriptional features are expected to participate to tumor evolution. Our group focuses on the study of the dynamics of histone modifications in cancer cells upon cancer treatment as well as during the initial steps of tumorigenesis. We develop experimental and computational approaches to map histone marks at single-cell resolution, enabling the investigation of the dynamics of chromatin marks in tumor samples (Grosselin et al. Nat Genet 2019; Prompsy et al. Nat Comm 2020).

We have recently combined single-cell epigenomic and transcriptomic approaches to lineage tracing strategies to reveal the initial epigenomic events driving tolerance to chemotherapy in triple-negative breast cancer (Marsolier & Prompsy et al., Nat Genet 2022). We show that the repressive histone mark H3K27me3 is a lock to the activation of a drug-persistent expression program in breast cancers. Under chemotherapy, very few cells can survive the treatment, and these cells have a remodeled repressive epigenome, with targeted loss at key promoters. Using a demethylase inhibitor in combination to chemotherapy, we improve the response rate and delay recurrence both in vitro and in vivo.

We also study mechanisms of cell plasticity in early breast tumorigenesis in vivo. We have recently mapped state transitions during BRCA1-tumorigenesis using mouse and human samples. We discovered that luminal progenitor cells undergo a partial epithelial to mesenchymal transition at the onset of tumorigenesis (Landragin &Saichi, biorxiv 2022).

SESSION II (SHORT TALKS)

Ines Hellmann (LMU, Munich)

Principles of regulatory evolution inferred from early differentiation in primates

Philipp Janssen, Jessica Radmer, Zane Kliesmete, Anita Termeg, Paulina Spurk, Fiona Edenhofer, Mari Ohnuki, Wolfgang Enard, Ines Hellmann

Changes in regulatory sequences have undoubtedly played an important role during human evolution. However, we still know very little about them. This is in part because we do not understand very well how gene regulation is encoded in the human genome. Here, we take an evolutionary approach by generating and analyzing single-cell expression and epigenomic data from four primate species during embryoid body formation to learn more about the "regulatory code". One central insight from recent studies of the evolution of regulatory elements is that they have a high turnover rate even though the regulated processes are very conserved across species. Hence, it has been speculated that compensatory evolution is common in regulatory evolution.

In order to evaluate how common compensatory evolution is for transcriptional regulatory elements (TREs) during primate evolution, we differentiated induced pluripotent stem cells from humans, orangutans, cynomolgus and rhesus macaques to embryoid bodies (EBs) and generated scRNA-seq from day 8 and day 16 and scATAC-seq data for the day 16 samples. In this data we could confidently assign and distinguish at least 10 cell types that were present in at least two species.

Using the scRNA-seq data, we estimate the conservation of gene regulation and regulatory networks during early primate differentiation, and using the scATAC-seq data we infer and compare the active TREs for each gene, species and trajectory. This allows us to link sequence and regulatory evolution, allowing us to gain new insights into the workings of the regulatory code.

Daria Romanovskaia (CeMM, Vienna)

Dissecting Ewing sarcoma heterogeneity at a single-cell resolution

Daria Romanovskaia[1], Marcus Tötzl[2*], Christoph Bock[1,3], Eleni M. Tomazou[2]*

[*]DR and MT contributed equally

[1] CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria.

[2] Children's Cancer Research Institute, St. Anna Kinderkrebsforschung, Vienna, Austria.

[3] Institute of Artificial Intelligence, Center for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna, Vienna, Austria

Ewing sarcoma (a pediatric solid tumor, often referred to as an enhancer disease) is an excellent model for studying the roles of epigenetic deregulation and tumor heterogeneity, given its low mutation rates and its well-defined oncogenic driver (the fusion protein EWS-FLI1). We have previously shown that Ewing sarcoma is characterised by widespread epigenetic reprogramming of gene-regulatory elements downstream of the EWS-FLI1 fusion (Sheffield, 2017).

Here we dissect the Ewing sarcoma epigenetic heterogeneity at a single cell resolution. We explain the extensive clinical heterogeneity by unravelling tumor cell heterogeneity and the interaction of tumor cells with cells of the tumor microenvironment.

We have performed joint single-cell epigenome and transcriptome profiling and spatial transcriptomics for a retrospective cohort of 30 clinically well-annotated frozen patient samples, providing detailed information of cell states and tumor heterogeneity.
We identified epigenetic and transcription-regulatory processes that correspond to the tumor differentiation state and can be associated with relapse and/or metastatic disease. We have characterized changes in the epigenetic landscape of the disease and identified signatures of open chromatin, associated with it. We have described the composition and the epigenetic state of the tumor microenvironment in Ewing sarcoma.
Our in-depth analysis of these data provides a catalogue of the precise cellular identities and compositions of Ewing sarcoma tumors and uncovers alterations in gene regulation that affect the cellular interplay, thus helping to understand the underlying biology of the developmental progression of fusion driven pediatric sarcomas.

Sheffield, N. C. et al. DNA methylation heterogeneity defines a disease spectrum in Ewing sarcoma. Nat. Med. 23, 386–395 (2017).

Kathrin Kattler (USAAR Saarbrücken)

Cell printer assisted single cell ATAC-seq for in-depth epigenetic profiling of rare celltypes

Due to recent technological advances the high-throughput analysis of accessible chromatin in large numbers of single cells has become feasible and adds to the understanding of an important layer of regulation in heterogeneous cell populations. However, overall experiment costs are still high, especially when interested in a particular rare cell population which remains underrepresented in the resulting huge datasets.
Thus, we developed a precise and cost-efficient scATAC-seq protocol to specifically analyze rare celltypes using the Cytena f.sight cell printer. This instrument is a very reliable, cost-efficient and gentle cell dispensing system, which greatly reduces premature cell lysis and avoids debris, duplets, and clumps through its sophisticated optical control parameters. Cells can be efficiently selected based on size and roundness. In addition, the f.sight optionally allows the use of fluorescence as an optical selection parameter for gentle selection of e.g. GFP-labelled or antibody stained cells. We developed a f.sight based approach to dispense specifically selected intact (tagmented) nuclei isolated from cell lines or primary tissues resulting in a fast and robust scATAC-seq protocol that can be applied to the analysis of rare cell populations.
In addition, a modified version of the protocol can be applied to manually isolated single cells. We used this manual approach to study the chromatin landscape during early mouse embryogenesis.

Jake Yeung (ISTA, Klosterneuburg)

Uncovering dynamic relationships between histone modifications in single cells

Jake Yeung[1], Maria Florescu[2], Alexander van Oudenaarden[2]
[1] Institute of Science and Technology Austria (ISTA), Klosterneuburg, Austria
[2] Hubrecht Institute, Utrecht, Netherlands

Regulation of chromatin states involves the dynamic interplay between different histone modifications to control gene expression. Recent advances have enabled genome-wide mapping of histone marks in single cells, but most methods are constrained to profile only one histone mark per cell.
Here we present an integrated experimental and computational framework, scChIX (single-cell chromatin immunocleavage and unmixing), to map multiple histone marks in single cells. scChIX multiplexes two histone marks together in single cells, then computationally deconvolves the signal

using training data from respective histone mark profiles. This framework learns the cell type-specific correlation structure between histone marks, and therefore does not require a priori assumptions of their genomic distributions.
Using scChIX, we demonstrate multimodal analysis of histone marks in single cells across a range of mark combinations: two repressive marks, two active marks, and an active plus a repressive mark. In mouse organogenesis, we find that cell type-specific regulation in active chromatin can be accompanied by stable heterochromatin landscapes that are shared across cell types. Applying scChIX to two active marks during in vitro macrophage differentiation, we find H3K4me1 dynamics preceding H3K36me3. Modeling these dynamics enables integrated analysis of chromatin velocity during differentiation. Overall, scChIX quantifies the interplay between histone modifications in single cells.

Aybuge Altay (MPI-MG Berlin)

Annotating scATAC-seq pseudobulk clusters

Aybuge Altay, Yufei Zhang, Martin Vingron
Max Planck Institute for Molecular Genetics, Ihnestraße 63, Berlin, 14195, Germany

Chromatin structure can control the accessibility of potential gene regulatory elements in a dynamic and cell-type specific manner and therefore plays a critical role in gene regulation. Although genome-wide chromatin structure can be measured by technologies like ATAC-seq in 'bulk', measurements in single-cell (sc) resolution (e.g. scATAC-seq) suffer from abounding zeros in the resulting data. Annotating the cell-types in scATAC-seq data remains a challenge mainly due to the lack of marker open regions to characterize cell-types and the sparsity of the scATAC-seq data. To overcome these limitations, we create scATAC-seq pseudobulk clusters by summing up the reads in a scATAC-seq cluster. We then co-embed these pseudobulks with FACS-sorted bulk ATAC-seq in PCA space and annotate each pseudobulk by the closest bulk cell-type. We couple our approach with transcription factor (TF) footprinting analysis and train a classifier with bulk ATAC-seq TF footprinting profiles to predict the cell-types of pseudobulks. This strategy provides a feasible way to overcome sparsity and leverages a large number of characterized ATAC-seq data. Our pipeline noticeably resolves the cell-type annotations when applied to human primary blood and brain data and performs comparably well with the existing methods demonstrating the strength of our strategy.

SESSION III

Maria Colomé-Tatché (Helmholtz Munich)

Single cell and single molecule computational epigenomics

Single-cell epigenomic measurements represent a rich layer of regulatory information that stands between the genome and the transcriptome. This rich data can be used to study single-cell identity, as well as other chromatin-based properties of single-cells.
In this talk I will discuss how single-cell ATAC-seq and single-cell DNA methylation data can be used to study cell identity [1]. I will introduce and compare multiple feature space constructions for epigenetic data analysis and show the feasibility of common clustering, dimension reduction and trajectory learning techniques for both single-cell DNA methylation data and scATAC-seq data.
Studying single-cell DNA methylation heterogeneity using single-cell DNA methylation measurements is however complicated, as experimental protocols are costly and difficult to implement. I will present here an alternative strategy, which involves minION sequencing combined with deconvolution of the

single-molecule methylation signals to reconstruct cell-type methylation profiles. I will show how, using this method, it is possible to deconvolve the methylomes of different cell types from an in-silico mix of cells.

Apart from single-cell identity, another level of genomic information that can be extracted from single-cell data are single-cell copy number variations (CNVs). I will present a novel algorithm, epiAneufinder [2], which exploits the read count information from scATAC-seq data to extract genome-wide CNVs for individual single-cells, and I will show how the obtained CNVs are comparable to the ones obtained from single-cell whole genome sequencing data. Thanks to epiAneufinder it is therefore possible to add a relevant extra layer of genomic information, namely single-cell copy number variation, to every scATAC-seq dataset without the need of additional experiments.

[1] A. Danese, M.L. Richter, D.S. Fischer, F.J. Theis and M. Colomé-Tatché. EpiScanpy: integrated single-cell epigenomic analysis. Nature Communications, 12, 5228 (2021).

[2] A. Ramakrishnan, A. Symeonidi, P. Hanel, M. Schubert, Maria Colomé-Tatché. epiAneufinder: identifying copy number variations from single-cell ATAC-seq data. doi: https://doi.org/10.1101/2022.04.03.485795

Ivan Costa (RWTH, Aachen)

Computational epigenomics: from single cells to gene regulatory networks

Single cell sequencing combined with open chromatin (ATAC) is a powerful tools to uncover epigenetic and regulatory changes associated with cellular and malignant transformations. Moreover, the integrative analysis of transcripts, open chromatin and their DNA sequence from multiomics protocols allow the inference of gene regulatory networks controlling cell differentiation. However, this data has challenging characteristics such as high dimension, high sparsity and low count distributions. We will describe a suite of computational methods for tackling these computational challenges. This includes methods for dimension reduction and denoising of sparse scATAC-seq matrices methods for learning integrative embeddings and for inferring enhancer based gene regulatory networks from single cell multiomics data. These were used to dissect differentiation processes associated with fibrosis in kidney diseases and heart myocardial infarction.

Li, Z., et al. scOpen: chromatin-accessibility estimation of single-cell ATAC data. Nature Communications, 12(1):6386 (2021) Cheng, M., et al. MOJITOO: a fast and universal method for integration of multimodal single cell data. Bioinformatics, 32: i282–i289 (2022). Kuppe, C., et al.,. Nature, 608:766–777 (2022).

Li, Z., et al., scMEGA : Single-cell Multiomic Enhancer-based Gene Regulatory Network Inference. bioRxiv, doi:10.1101/2022.08.10.503335 (2022).

Guido Sanguinetti (SISSA, Trieste)

Bayesian modelling of single-cell multi-omics data

Joint work with Christos Maniatis, Catalina Vallejos and Nour El Kazwini

The last decade has witnessed the irresistible rise of single-cell RNA-seq, a technology which can measure simultaneously the mRNA expression levels of thousands of genes in thousands of individual cells. This has enabled an unconfounded measurement of cell states, and the discovery of new cell populations. More recently, scRNA-seq has been paired with other techniques (such as BS-seq or ATAC-seq) to measure simultaneously multiple molecular features in the same cell. The promise of these technologies is the ability to understand gene regulation at the single cell level, but, so far, the high levels of noise have prevented significant inroads in that direction. In this talk I argue that some of these effects might be remedied by taking a Bayesian statistical perspective to the modelling of noise in the data. I show that a simple hierarchical model can lead to improved estimates of correlations between genomic features, a taxing task with standard tools (Maniatis et al, PLoS CompBio 2022). I

then move on to present a more structured Bayesian model which aims to infer biological pathways (topics) from joint scRNA-seq scATAC-seq protocols data.

SESSION IV

Eileen Furlong (EMBL, Heidelberg)

Dissecting function in the regulation of developmental trajectories

Expression states are regulated by transcription factors (TFs), which drive both the diversity and progression of cell lineages. TFs function through enhancers, non-coding elements that recruit different combinations of TFs to regulate a specific expression pattern in both time and tissue. Transcription factors thereby regulate developmental progression and cellular diversity, yet characterising their functional input remains challenging. We are exploring the extent to which combining single cell regulatory genomics with both natural sequence variation (perturbations in cis) and TF loss-of-function mutants (perturbations in trans) can functionally dissect enhancer activity and developmental phenotypes, at both a cellular and molecular level. We recently showed that single cell ATAC-seq we can predict tissue-specific enhancers, identify cell types and follow their trajectories during embryogenesis[1]. We have now substantially extended this to study context-specific effects of natural sequence variation, in addition to dissecting the functional input of TF mutants. By sampling regulatory changes over a dense time-course of both mesoderm development and the entire embryo in Drosophila, we could reconstruct developmental trajectories of all major lineages, uncover the TFs involved and the enhancers they regulate[2,3]. We developed a nuclear genotyping strategy to systematically assess recessive lethal mutations from embryo pools with mixed genotypes. Applying this approach to four TF mutants revealed different outcomes on both cell state and enhancer function[2]. We have also shown how deep learning on a dense time-course can be used to learn and predict developmental time, allowing the formation of continuous trajectories and to zoom in the scale of minutes[3]. I will discuss how we are combining these approaches as a framework to reassess mutant phenotypes, and identify new phenotypes at both a cellular and molecular level.

[1] Cusanovich DA, Reddington JP, Garfield DA, et al. (2018). The cis-regulatory dynamics of embryonic development at single cell resolution. *Nature*, Mar 22;555(7697):538-542

[2] Secchia S, et al. (2022). Simultaneous cellular and molecular phenotyping of embryonic mutants using single-cell regulatory trajectories. *Dev Cell*. Feb 28;57(4):496-511.e8.

[3] Calderon D, Blecher-Gonen R, Huang X, Secchia S, et al. (2022). The continuum of Drosophila embryonic development at single-cell resolution. *Science* Aug 5;377(6606):eabn5800.

Giovanni Tonon (IRCCS Milano)

Sketching Open and Closed Chromatin, One Cell at a Time

Francesca Giannese, Martina Tedesco, Dejan Lazarevic, Valentina Giansanti, Dalia Rosano, Silvia Monzani, Irene Catalano, Elena Grassi, Eugenia R. Zanella, Oronza A. Botrugno, Leonardo Morelli, Paola Panina Bordignon, Giulio Caravagna, Andrea Bertotti, Gianvito Martino, Luca Aldrighetti, Sebastiano Pasqualato, Livio Trusolino, Davide Cittaro and Giovanni Tonon

Recent efforts have succeeded in surveying open chromatin at the single-cell level, but high-throughput, single-cell assessment of heterochromatin and its underlying genomic determinants remains challenging. We engineered a hybrid transposase including the chromodomain (CD) of the heterochromatin protein-1a (HP-1a), which is involved in heterochromatin assembly and maintenance through its binding to trimethylation of the lysine 9 on histone 3 (H3K9me3), and developed a single-cell method, single-cell genome and epigenome by transposases sequencing (scGET-seq), that, unlike single-cell assay for transposase-accessible chromatin with sequencing (scATAC-seq), comprehensively

probes both open and closed chromatin and concomitantly records the underlying genomic sequences. We tested scGET-seq in cancer-derived organoids and human-derived xenograft (PDX) models and identified genetic events and plasticity-driven mechanisms contributing to cancer drug resistance. Next, building upon the differential enrichment of closed and open chromatin, we devised a method, Chromatin Velocity, that identifies the trajectories of epigenetic modifications at the single-cell level. Chromatin Velocity uncovered paths of epigenetic reorganization during stem cell reprogramming and identified key transcription factors driving these developmental processes. scGET-seq reveals the dynamics of genomic and epigenetic landscapes underlying any cellular processes.

SESSION V

Jop Kind (Hubrecht Institute, Utrecht)

The chromatin landscape during early embryo development at single-cell resolution

Isabel Guerreiro[1], Franka Rang[1*], Yumiko Kitazawa[2], Femke Groenveld[1], Ramada van Beek[1], Antoine HMF Peters[2], Jop Kind[1]*

[1] Hubrecht Institute, Utrecht, Netherlands

[2] Friedrich Miescher Institute for Biomedical Research (FMI), Basel, Switzerland

During embryonic development one cell divides and differentiates resulting in a complex organism made up of multiple cell types. Cell fate is largely determined by when and where specific genes are activated or silenced during embryonic development. This process of differential gene regulation can be achieved in several ways, including histone modifications and nuclear localization. For example, chromatin regions located at the nuclear periphery, termed lamina associated domains (LADs), are heterochromatic regions that impact genome organization and play a role in gene regulation. The very first cell fate decision occurs during the first few days of development and is crucial for correct embryonic development. The emergence of single-cell and low-input technologies has revealed heterogeneity and massive chromatin rearrangement occurring during the first stages of development. However, the role of these features in regulating gene expression as well as the interplay between the various chromatin marks remains elusive. In this study, we have employed DamID-based techniques to study LAD, chromatin accessibility and histone mark distribution in early embryos in single cells. Both epigenetic and gene expression readouts from the same cell were successfully obtained in several early developmental stages. This approach has revealed major LAD variability following the first cell division, while silencing histone marks and accessible chromatin remained remarkably similar among cells. Cell-cell LAD variability does not seem to affect transcription. In addition, a large number of genomic regions that tend to locate at the nuclear periphery in all cell types, are strikingly not located at the lamina in 2-cell embryos but are instead enriched in H3K27me3. This suggests an alternative silencing strategy during early development. Upon knock-out of maternal Eed and concomitant loss of maternal H3K27me3 during early embryogenesis, these regions relocate to the lamina and adopt a similar nuclear location as in ESCs and most other cell types. Thus, after fertilization, H3K27me3 appears to delay localization of genomic regions to the nuclear periphery. This study brings insight into the atypical regulatory mechanisms at play during early development by directly linking epigenetics and transcription in single cells.

Wolf Reik (Altos Labs, Cambridge)

Single cell multi-omics landscape of development and ageing

Diljeet Gill[1], Aled Parry[1], Jasmin Taubenschmid-Stowers[1], Maria Rostovskaya[1], Ricard Argelaguet[1,4], Stephen Clark[1,4], Christel Krueger[1,4], Felix Krueger[1,4], Fatima Santos[1] & Wolf Reik[1,2,3,4]

[1] Epigenetics Programme, Babraham Institute, Cambridge CB22 3AT

[2] Centre for Trophoblast Research, University of Cambridge, CB2 3EG

[3] Wellcome Trust Sanger Institute, Cambridge CB10 1SA

[4] Altos Labs Cambridge Institute, Cambridge CB21 6GP

Epigenetic information is relatively stable in somatic cells but is reprogrammed on a genome wide scale in germ cells and early embryos. Reprogramming is essential for imprinting, the return to naïve pluripotency, the erasure of epimutations, and for the control of transposons. Following reprogramming, epigenetic marking occurs prior to and during lineage commitment in the embryo. The epigenome changes in a potentially programmed fashion during the ageing process; this epigenetic ageing clock seems to be conserved in mammals.

Our work addresses the mechanisms and consequences of global epigenetic reprogramming in the germ line and at zygotic genome activation. Using single cell multi-omics techniques, we are beginning to chart the epigenetic and transcriptional dynamics and heterogeneity during the exit from pluripotency and initial cell fate decisions leading up to gastrulation. We discovered priming of enhancers prior to lineage decisions as well as acute epigenetic remodeling of enhancers at the time of lineage commitment. We are also interested in the programmed degradation of epigenetic information during the ageing process, how this might be coordinated across tissues and individual cells, and how this process potentially could be reversed.

ABSTRACTS – POSTERS

1. Carlos Alfonso Gonzalez: Sites of Transcription Initiation Drive mRNA Isoform Selection
2. Aybuge Altay: Annotating scATAC-seq pseudobulk clusters
3. Javier Bregante: Human cholangiocytes expanded long-term as non-cystic organoids acquire bi-potency under hepatocyte inducing conditions
4. Anna Danese: Comparison of the classic and generalized telegram models to describe transcriptional bursting
5. Tobias Hohl: Understanding the interplay between chromatin state and transcription factor networks during neuronal differentiation
6. Narges Jafari: Effects of Tet3 Isoforms on Gene Expression and chromatin Architecture of Early Mouse Embryos
7. Kathrin Kattler: Cell printer assisted single cell ATAC-seq for in-depth epigenetic profiling of rare celltypes
8. Gabriele Malagoli: Comprehensive analysis of long non-coding RNAs in breast cancer using topic modeling
9. Maria Richter: Single-cell metabolic profiling reveals subgroups of primary human hepatocytes showing heterogeneous responses to drug challenge
10. Daria Romanovskaia: Dissecting Ewing sarcoma heterogeneity at a single-cell resolution
11. Melissa Saichi: Luminal progenitors undergo partial epithelial-to-mesenchymal transition at the onset of basal-like breast tumorigenesis
12. Katharina Schmid: Identification of genetic variants that impact gene co-expression relationships using large-scale single-cell data

P01 Carlos Alfonso Gonzalez (MPI-IE, Freiburg)

Carlos Alfonso Gonzalez: Sites of Transcription Initiation Drive mRNA Isoform Selection

The generation of distinct messenger RNA isoforms through alternative splicing and alternative 3' end formation influences the expression and function of genes, often in a cell-type specific manner. Here, we quantitatively assess the regulatory relationships between transcription initiation and co-transcriptional processing steps, particularly 3' end formation. Applying multiple long-read-sequencing approaches to obtain an assembly accurately representing even the longest mRNA isoforms from end-to-end, we quantify mRNA isoform choice in Drosophila and human tissues, including the transcriptionally complex nervous system. We find that in Drosophila brains as well as in human cerebral organoids, 3' end site choice is globally influenced by the site of transcription start. We define a subset of TSSs, “dominant promoters” that impose a transcriptional constraint to predetermine splice and polyadenylation variants, which are characterized by specific epigenetic signatures. In vivo deletion of dominant promoters disrupted the 3' end expression landscape. Our study demonstrates the crucial impact of transcription initiation site choice on the regulation of transcript diversity and tissue identity.

P02 Aybuge Altay (MPI-MG Berlin)

Annotating scATAC-seq pseudobulk clusters

Abstract available in SHORT TALKS section

P03 Javier Bregante (MPI-MG Berlin)

Human cholangiocytes expanded long-term as non-cystic organoids acquire bi-potency under hepatocyte inducing conditions

Javier Bregante [1], Flaminia Kaluthantrige [1,2], Dylan Liabeuf [3], Yohan Kim[1], Daniel Stange [3], Meritxell Huch [1]
[1] Max Planck Institute of Molecular Cell Biology and Genetics, Dresden 01307, Germany
[2] Department of Physiology Development and Neuroscience, University of Cambridge, Cambridge CB2 [3] DY, UK; GlaxoSmithKline, Stevenage SG1 2NY, UK.
[3] Department of Visceral, Thoracic and Vascular Surgery, Medical Faculty and University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

Human liver ductal epithelium is morphologically, functionally and transcriptionally heterogeneous. However, understanding the dynamics within human ductal epithelium has been hampered by the absence of an in vitro system that fully mimics its heterogeneity. We developed an organoid system that allows long-term expansion of human cholangiocytes, while acquiring a bipotent state. To further dissect the transcriptional heterogeneity of this refined organoid system, we employed single-cell RNA sequencing (sc-RNA-seq). We described five transcriptionally different cell states (ductal progenitor, bipotent ductal progenitor, proliferative ductal, mature ductal and mature hepatocyte) that mimic the transcriptional heterogeneity described in vivo. Next, through Velocity and diffusion analysis, we unveiled that human liver organoids arise from a non-proliferative ductal progenitor population, which is capable to transitioning to either a proliferative state or to a mature state, via an intermediate bipotent state. Finally, we aimed to recapitulate cholangiocyte-to-hepatocyte differentiation that occurs upon severe liver injury, and we observed that the acquisition of a bipotent state facilitates the generation of a more mature and functional hepatocyte state. Overall, our study provides an in vitro system that allows gaining better understanding of the cellular heterogeneity and trajectories between cells within the human ductal epithelium.

P04 Anna Danese (LMU, Munich)

Comparison of the classic and generalized telegram models to describe transcriptional bursting

Anna Danese, Rijesh Pal Shrestha, Stefan H. Stricker, M. Colomé-Tatché

Transcription is a discrete process with stochastic properties, resulting in the expression of RNA in bursts. The classical telegram model (CTM) often used to describe transcriptional bursting is a two-state model where promoters switch between on and off states given exponentially distributed dwell times. It was first implemented to describe transcriptional dynamics using scRNA-seq by Larsson et al. (2019). The model captures part of the diversity of transcription regulation. Larsson et al. showed the importance of the core promoter type as well as the presence of enhancers on the transcription dynamics.

However, the initiation of transcription is a multi-step process that doesn't always follow exponentially distributed dwell times. In this project, we explore the limitations of the CTM to describe certain gene regulation patterns and compare its performance to a generalized telegram model (GTM) that allows arbitrary dwell-time distributions. We compared the performance of both models using scRNA-seq from adult mouse fibroblasts. If they estimate similar transcriptional bursting parameters for most genes, we also found that certain transcription factors regulated by extracellular signaling (i.e. Fos, Atf4) are better described by the GTM. This highlights the importance of a more comprehensive transcriptional dynamic model to capture more of the diversity of transcriptional regulation mechanisms.

P05 Tobias Hohl (MPI-IE Freiburg)

Understanding the interplay between chromatin state and transcription factor networks during neuronal differentiation

Tobias Hohl [1], Camila Fullio [2], Laura Arrigoni [1], Tanja Vogel[2], Thomas Manke [1]

[1] Max-Planck-Institute of Immunobiology and Epigenetics, Freiburg

[2] Institute of Anatomy and Cell Biology, University Freiburg, Freiburg

During embryonal neuronal development, the epigenetic state of chromatin plays a crucial role in the spatiotemporal regulation of gene expression. The removal and deposition of histone modifications contributes to the establishment of different chromatin states which in turn determine DNA accessibility and transcription factor binding to cis-acting regulatory elements. Previously we have shown that the in vitro differentiation of murine embryonic stem cells to neuronal progenitor cells globally increases levels of H3K79me3, a histone modification deposited by DOT1L. Inhibiting DOT1L led to decreased accessibility of SOX2-binding regions as well as negatively affected SOX2 binding to known related enhancers. Therefore, the activity of DOT1L is crucial for the differentiation of neuronal progenitors. To follow up on this work, we want to investigate on cell type specific effects of DOT1L during the development of the primary cortex. We will translate and expand our findings to an in vivo model that helps us generate a comprehensive single-cell multi-omics dataset with a temporal resolution and compare wildtype results with a conditional DOT1L knock-out model. Multimodal and integrative data analysis will let us infer developmental trajectories and cell-type specific transcription factor networks, as well as unravel the role that DOT1L plays in both the differentiation of the numerous cell types involved and the activities of transcription factors that are crucial for the development of the different neuronal cell populations.

P06 Narges Jafari (USAAR, Saarbrücken)

Effects of Tet3 Isoforms on Gene Expression and chromatin Architecture of Early Mouse Embryos

Changes in the DNA methylation landscape are controlled by Tet oxygenases which act as major modifiers of DNA-methylation. Tet3 is an important player during early development and in the brain, mediating the full oxidation from 5mC to 5hmC, 5fC and 5caC. Tet3 triggers replication independent (active) and replication dependent (passive) demethylation processes.

Tet3 is expressed in different isoforms whose role during preimplantation development, in stem cells and in the brain are largely unclear. To better understand the role and dynamics of Tet3 mediated modifications we analysed the expression of various isoforms of Tet3 and generated a series of isoform specific knockouts in mice. In this study we focused on Tet3 oocyte specific isoforms and their effect on expression, methylation and chromatin architecture of the preimplantation mouse embryo. Our observations show in vivo consequences, manifested at different degrees, caused by the loss of individual isoforms as seen by NGS based analyses, immunohistochemistry and phenotypes of the KO mice. We find that the loss of Tet3OO has consequences on gene expression in oocytes and 2-cell embryos and causes changes in 5mC DNA-methylation in repetitive elements of 2-cell embryos.

P07 Kathrin Kattler (USAAR, Saarbrücken)

data Cell printer assisted single cell ATAC-seq for in-depth epigenetic profiling of rare celltypes

Abstract available in SHORT TALKS section

P08 Gabriele Malagoli (Helmholtz Munich)

Comprehensive analysis of long non-coding RNAs in breast cancer using topic modeling

Gabriele, Valle Filippo, Barillot Emmanuel, Caselle Michele, Martignetti Loredana

Single-cell RNA sequencing is a powerful tool to explore cancer heterogeneity. However, the expression of lncRNAs in single cells is still to be studied extensively and methods to deal with the sparsity of this type of data are lacking. Here, we propose a topic modeling approach to investigate the transcriptional heterogeneity of luminal and triple negative breast cancer cells using patient-derived xenograft models of acquired resistance to chemotherapy and targeted therapy.

We show that using an integrative clustering that combines the information coming from mRNAs and lncRNAs treated as disjoint omic layers greatly improves the accuracy of cell classification. Topics associated with specific breast cancer subpopulations show a clear enrichment for pathways involved in subtyping and progression of breast cancer and to sets of lncRNA encoded in the open chromatin regions of breast cancer cell lines. We identified lncRNAs strongly associated with cell clusters already well known in the literature, such as MALAT1 and NEAT1, and highlighted some others that may be clinically relevant.

P09 Maria Richter (LMU, Munich)

Single-cell metabolic profiling reveals subgroups of primary human hepatocytes showing heterogeneous responses to drug challenge

M. L. Richter, E. Sanchez-Quant, M. Colomé-Tatché, C.P. Martinez-Jimenez

Xenobiotics are primarily metabolized by hepatocytes in the liver, and primary human hepatocytes (PHHs) are the gold standard model for the assessment of drug efficacy, safety and toxicity in the early phases of drug development. Recent advances in single-cell genomics have shown liver zonation and ploidy as main drivers of cellular heterogeneity. However, little is known about the impact of

hepatocyte specialization on liver function upon metabolic challenge, including hepatic metabolism, detoxification, and protein synthesis. Here, we investigate the metabolic capacity of individual human hepatocytes in vitro, and assess how chronic accumulation of lipids enhances cellular heterogeneity and impairs the metabolisms of drugs. A phenotyping five-probe cocktail was used to identify four functional subgroups of hepatocytes that respond differently to drug challenge and fatty acid accumulation. These four subgroups display differential gene expression profiles upon cocktail treatment and xenobiotic metabolism-related specialization. Notably, intracellular fat accumulation leads to increased transcriptional variability and diminished the drug-related metabolic capacity of hepatocytes. Our results demonstrate that, upon a metabolic challenge such as exposure to drugs or intracellular fat accumulation, hepatocyte subgroups lead to different and heterogeneous transcriptional responses.

P10 Daria Romanovskaia (CeMM, Vienna)

Dissecting Ewing sarcoma heterogeneity at a single-cell resolution

Abstract available in SHORT TALKS section

P11 Melissa Saichi (CNRS/Institut Curie, Paris)

Luminal progenitors undergo partial epithelial-to-mesenchymal transition at the onset of basal-like breast tumorigenesis

Camille Landragin [1,2,] , Melissa Saichi [1,2,*] , Pacôme Prompsy [1,2] , Adeline Durand [1,2] , Jérémy Mesple [3] , Amandine Trouchet [1,4] , Marisa Faraldo [5,6], Hélène Salmon [3] , Céline Vallot [1,2,4]*

[1] CNRS UMR3244, Institut Curie, PSL University, Paris, France

[2] Translational Research Department, Institut Curie, PSL University, Paris, France

[3] INSERM U932, Institut Curie, PSL University, Paris, France

[4] Single Cell Initiative, Institut Curie, PSL University, Paris, France

[5] CNRS UMR3215, Institut Curie, PSL University, Paris, France

[6] INSERM U934, Institut Curie, PSL University, Paris, France

[*] co-first authors

Luminal progenitors have been proposed to be the cell type of origin of BRCA1 basal-like breast tumors. Yet how luminal cells transform into invasive cancer cells remains poorly understood. With single-cell transcriptomics, epigenomics and spatial multiplex imaging, we characterized the sequence of events leading luminal progenitors with abnormal features to launch umorigenesis. We demonstrate that partial epithelial-to-mesenchymal (EMT) transition occurs at the onset of tumorigenesis in pre-tumoral lesion samples, following a consequent genomic crisis attested by an initial senescence-like process with a disrupted heterochromatin profile. Our pre-tumoral state
signature was also detected in early-stage basal-like tumors, and was associated with a good prognosis. Overall, our study highlights key molecular, genetic and epigenetic events occuring at BRCA1 tumorigenesis, including partial EMT activation, and open large perspectives for early interception of BRCA1 basal-like tumors.

P12 Katharina Schmid (LMU, Munich)

Identification of genetic variants that impact gene co-expression relationships using large-scale single-cell data

*Shuang Li [1,2 *], Katharina T. Schmid [3,4 *], Dylan de Vries [1 *], Maryna Korshevniuk [1], Corinna Losert [3,4], Roy Oelen [1], Irene van Blokland [1,5], BIOS Consortium, sc-eQTLgen Consortium, Hilde E. Groot [5], Morris Swertz [1,2], Pim van der Harst [6], Harm-Jan Westra [1], Monique van der Wijst [1], Matthias Heinig [3,4 **], Lude Franke [1 **]*

[1] Genetics Department, University Medical Center Groningen, the Netherlands

[2] University of Groningen, University Medical Center Groningen, Genomics Coordination Center, Groningen, the Netherlands
[3] Institute of Computational Biology, Helmholtz Center Munich, Munich, Germany
[4] Department of Informatics, Technical University Munich, Munich, Germany
[5] Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands
[6] Department of Cardiology, University Medical Center Utrecht, Utrecht, the Netherlands
* These authors contributed equally ** These authors contributed equally

Expression quantitative trait loci (eQTL) studies have shown how genetic variants (SNPs) affect downstream gene expression. To understand the underlying upstream regulatory processes, single-cell data offers the unique opportunity to reconstruct personalized coexpression networks. By exploiting the large number of cells per individual, we can identify SNPs altering co-expression patterns (i.e.co-expression QTLs, co-eQTLs) using a limited number of individuals. To tackle the large multiple testing burden (SNP-Gene1-Gene2 combinations) associated with a genome-wide analysis, we conducted a co-eQTL meta-analysis across 3 scRNAseq PBMC datasets (177 participants, 1M cells) with a novel filtering strategy, followed by a permutation-based approach. Using this strategy, we identified a robust set of cell-type-specific co-eQTLs for 73 independent SNPs, affecting 946 gene pairs. These co-eQTLs provide novel insights into how disease-associated variants alter regulatory networks. For instance, the type I diabetes (T1D) SNP rs1131017 affects the co-expression between RPS26 and other ribosomal genes. Interestingly, the comparison across cell types showed distinct coeQTLs related to T cell activation specifically within T-lymphocytes, revealing a previously overlooked process that could explain its association with T1D. This example highlights the importance of studying gene regulation at the context-specific level to understand the biological implications of genetic variation. With the expected growth of sc-eQTL datasets, our strategy will soon identify many more co-eQTLs that help elucidate unknown disease mechanisms.