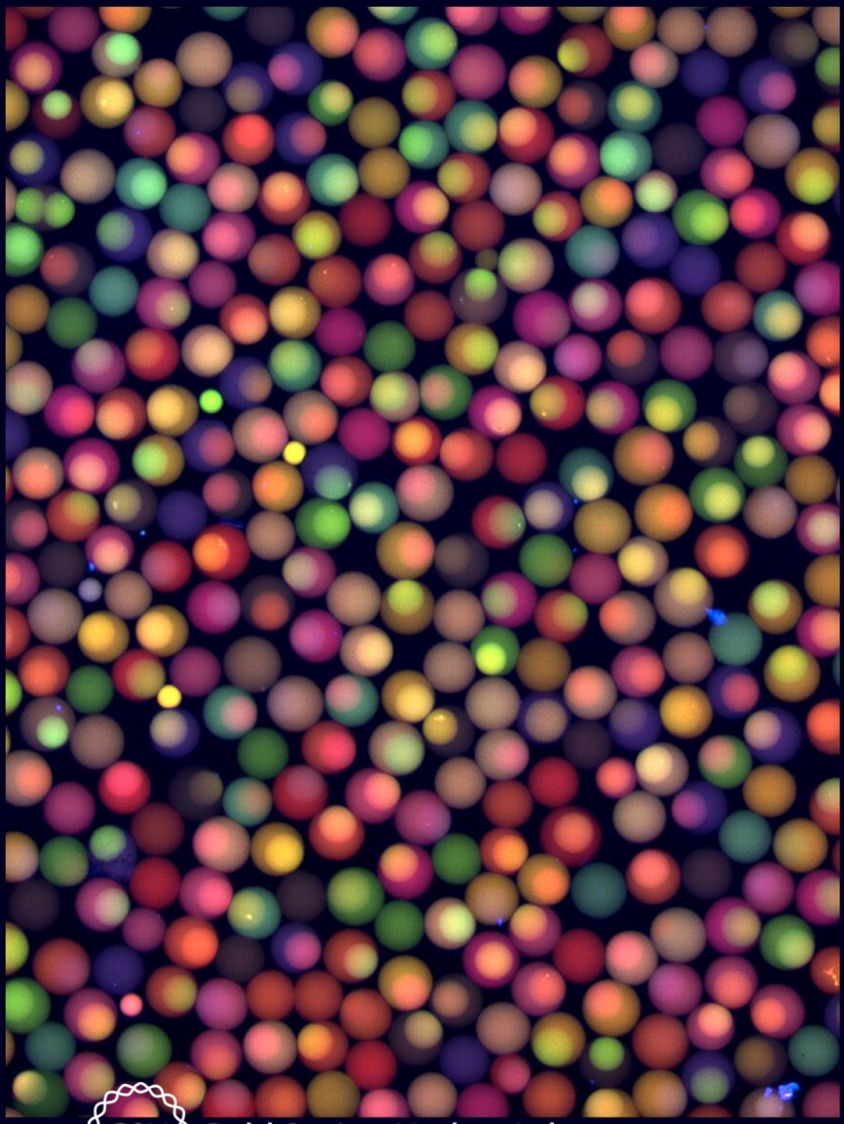
Abstracts of papers presented
at the 2025 meeting on

SINGLE CELL ANALYSES

November 12–November 15, 2025



CSH Cold Spring Harbor Laboratory
MEETINGS & COURSES PROGRAM

Abstracts of papers presented
at the 2025 meeting on

SINGLE CELL ANALYSES

November 12–November 15, 2025

Arranged by

Junhyong Kim, *University of Pennsylvania*
Tatjana Sauka-Spengler, *Stowers Institute for Medical Research & University of Oxford, UK*
Aaron Streets, *University of California, Berkeley*

CSH Cold Spring Harbor Laboratory
MEETINGS & COURSES PROGRAM

Support for this meeting was provided in part by **PacBio** and **10X Genomics.**

Contributions from the following companies provide core support for the Cold Spring Harbor meetings program.

Corporate Benefactors

Regeneron

Corporate Sponsors

Agilent Technologies
Biogen
Calico Labs
New England Biolabs
Novartis Institutes for Biomedical Research

Cover: Optically recognizable barcode beads (ORBBs). Image generated by Boden Eakins, Rodrigo Cotrim Chaves, Bryce Manifold (University of California, Berkeley).

SINGLE CELL ANALYSES
Wednesday, November 12 – Saturday, November 15, 2025

Wednesday	7:30 pm – 8:30 pm	**Keynote Speaker**
Thursday	9:00 am – 12:00 pm	**1** New Methods for Single Cell Analyses
Thursday	2:00 pm – 4:00 pm	**2** New Applications of Single Cell Analyses
Thursday	4:00 pm – 4:30 pm	**Flash Talks**
Thursday	4:30 pm – 5:00 pm	**Plenary I**
Thursday	5:00 pm	*Wine & Cheese Party*
Thursday	7:30 pm – 10:30 pm	**Poster Session**
Friday	9:00 am – 12:00 pm	**3** Spatially-resolved Single Cell Analyses I
Friday	2:00 pm – 4:30 pm	**4** Spatially-resolved Single Cell Analyses II
Friday	4:30 pm – 5:00 pm	**Plenary II**
Friday	6:00 pm	*Cocktails and Banquet*
Saturday	9:00 am – 10:30 am	**5** New Single Cell Analytical Methods
Saturday	10:45 am – 11:30 am	**Panel Discussion**

Lunchtime Workshops (immediately following morning session)
PacBio: Thursday, November 13
10X Genomics: Friday, November 14

Mealtimes at Blackford Hall are as follows:
Breakfast 7:30 am-9:00 am
Lunch 11:30 am-1:30 pm
Dinner 5:30 pm-7:00 pm

Bar is open from 5:00 pm until late

All times shown are US Eastern: Time Zone Converter

Cold Spring Harbor Laboratory is committed to maintaining a safe and respectful environment for all meeting attendees, and does not permit or tolerate discrimination or harassment in any form. By participating in this meeting, you agree to abide by the Code of Conduct.

For further details as well as Definitions and Examples and how to Report Violations, please see the back of this book.

Abstracts are the responsibility of the author(s) and publication of an abstract does not imply endorsement by Cold Spring Harbor Laboratory of the studies reported in the abstract.

These abstracts should not be cited in bibliographies. Material herein should be treated as personal communications and should be cited as such only with the consent of the author(s).

Please note that photography or video/audio recording of oral presentations or individual posters is strictly prohibited except with the advance permission of the author(s), the organizers, and Cold Spring Harbor Laboratory.

Any discussion via social media platforms of material presented at this meeting requires explicit permission from the presenting author(s).

Printed on 100% recycled paper.

PROGRAM

WEDNESDAY, November 12—7:30 PM

Introduction: **Junhyong Kim,** University of Pennsylvania
Tatjana Sauka-Spengler, Stowers Institute for Medical Research & University of Oxford, UK
Aaron Streets, University of California, Berkeley

KEYNOTE SPEAKER

Scott Fraser
Chan Zuckerberg Imaging Institute

"Better single cell analyses by combining AI and multidimensional imaging"

THURSDAY, November 13—9:00 AM

SESSION 1 NEW METHODS FOR SINGLE CELL ANALYSIS

Chairperson: **Jase Gehring,** University of California, Berkeley

Single cell and "single cell-ish" functional genomics for learning at scale
Paul Blainey.
Presenter affiliation: The Broad Institute of MIT and Harvard, Cambridge, Massachusetts.

High-quality-factor metasurfaces for dynamic peptide sequencing and single-cell phenotyping
Jennifer Dionne, Jack Hu, Varun Dolia, Darrell Omo-Lomai, Kai Chang, Nhat Vu, Amanda Kirane.
Presenter affiliation: Stanford University, Stanford, California; Pumpkinseed, Palo Alto, California. 1

Scalable co-sequencing of RNA and DNA in individual cells
Peter A. Sims, Timothy R. Olsen, Romella K. Sagatelian.
Presenter affiliation: Columbia University Irving Medical Center, New York, New York. 2

THURSDAY, November 13—2:00 PM

SESSION 2 NEW APPLICATIONS OF SINGLE CELL ANALYSES

Chairperson: **Boden Eakins,** University of California, Berkeley

THURSDAY, November 13—4:00 PM

FLASH TALKS

THURSDAY, November 13—4:30 PM

PLENARY SPEAKER

THURSDAY, November 13—5:00 PM

Wine and Cheese Party

THURSDAY, November 13—7:30 PM

POSTER SESSION

See p. xii for List of Posters

FRIDAY, November 14—9:00 AM

SESSION 3 SPATIALLY-RESOLVED SINGLE CELL ANALYSES I

Chairperson: **Leon Han,** University of California, Berkeley

Long Cai.
Presenter affiliation: California Institute of Technology, Pasadena, California.

Spatial sequencing of host-microbiome interactions
Iwijn De Vlaminck.
Presenter affiliation: Cornell University, Ithaca, New York. 13

Single cell spatial transcriptomics analysis of immune microenvironment in non-small cell lung cancer
Anni Zhang, Troy Noordenbos, Maximilian Diehn.
Presenter affiliation: Stanford University, Stanford, California. 14

Volumetric DNA microscopy for optics-free three-dimensional spatial genomics
Joshua Weinstein, Nianchao Qian.
Presenter affiliation: University of Chicago, Chicago, Illinois. 15

Photolabile oligonucleotides combined with topologically imposed light gradients enable spatially resolved single-cell transcriptomics and epigenomics
Robert Piscopio, Chieh Wang, Alex Chialastri, Mikolaj Godzik, Weiyue Wang, Lauren J. Li, Kellie Heom, Neha Saxena, Maxwell Z. Wilson, Siddharth S. Dey.
Presenter affiliation: University of California Santa Barbara, Santa barbara, California. 16

FRIDAY, November 14—2:00 PM

SESSION 4 SPATIALLY-RESOLVED SINGLE CELL ANALYSES II

Chairperson: **Xinyi Zhang,** Memorial Sloan Kettering Cancer Center, New York, New York

Integrated transcriptomic and morphological deep learning for spatially informed single-cell analysis of pancreatic cancer progression
Nischal Bhandari, Huaien Wang, Jose Reyes, Sandeep Nadella, James M. Crawford, Gabriele Campanella, Zhen Zhao.
Presenter affiliation: Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. 22

Modeling spatial variation using interpretable neural representations
Uthsav Chitra, Clover Zheng, Cong Ma, Ben Raphael.
Presenter affiliation: Princeton University, Princeton, New Jersey. 23

Prediction of protein subcellular localization in single cells
Xinyi Zhang, Yitong Tseo, Yunhao Bai, Fei Chen, Caroline Uhler.
Presenter affiliation: Broad Institute of MIT and Harvard, Cambridge, Massachusetts; Massachusetts Institute of Technology, Cambridge, Massachusetts. 24

Multi-scale exploratory spatial data analysis in lung adenocarcinoma
Lambda Moses, Bianca Dumitrascu.
Presenter affiliation: Columbia University, New York, New York. 25

FRIDAY, November 14—4:30 PM

PLENARY SPEAKER

Developing instrumented tissues for spatiotemporal omics
Shana Kelley.
Presenter affiliation: Northwestern University, Evanston, Illinois; Chan Zuckerberg Biohub Chicago, Chicago, Illinois. 26

SATURDAY, November 15—9:00 AM

SESSION 5 NEW SINGLE CELL ANALYTICAL METHODS

Chairperson: **Shiyi Yang,** University of California, Berkeley

SATURDAY, November 15—10:45 AM

PANEL DISCUSSION

SATURDAY, November 15—11:30 AM

Closing Remarks

< *Back*

< Back

< Back

< *Back*

HIGH-QUALITY-FACTOR METASURFACES FOR DYNAMIC PEPTIDE SEQUENCING AND SINGLE-CELL PHENOTYPING

Jennifer Dionne[1,2], Jack Hu[2], Varun Dolia[1], Darrell Omo-Lomai[1], Kai Chang[1], Nhat Vu[2], Amanda Kirane[1]

[1]Stanford University, Stanford, CA, [2]Pumpkinseed, Palo Alto, CA

The earth's biosphere is incredibly information-rich, with estimated information transmission rates exceeding those of the technosphere by 9 orders of magnitude (Lingam et al, Life 13, 1850, 2023). Yet, current methods to extract this information are slow and incomplete, hindering our ability to understand the genesis and evolution of biochemical systems, and to optimize their performance. Here, we present nanophotonic methods that may enable unprecedented data about biochemical systems, at rates previously unattainable. First, we describe our lab's Si-photonic "Very-large-scale Integrated high-Q Nanophotonic Pixels". These photonic resonators achieve high-Q factors, subwavelength mode volumes, and controlled dipole-like radiation, simultaneously, with Q-factors from the thousands to millions, and resonator densities exceeding 100M/cm2. These resonators enable an 8-order-of-mangitude enhancement in spontaneous Raman scattering, enabling efficient single-cell to single-molecule Raman spectroscopy.

As a first application of these resonators, we describe our protein sequencing technology, deSIPHR (de-novo sequencing and identification of proteins with high-resolution Raman spectroscopy). We tether peptides from major histocompatibility complex to each resonator, and use dynamic Raman spectroscopy to monitor the cleavage of each amino acid from the distal terminus. We demonstrate >96% accuracy in distinguishing wild type and mutated peptides from adenocarcinoma cell lines, as well as site-specific phosphorylation. We also show how these resonators, combined with computational metadynamics, can be used to identify new amino acid and post-translational combinations, including distinct conformational populations of glycans.

Next, we adapt our metasurfaces to demonstrate subcellular differentiation and functional state characterization of the tumor immune microenvironment (TIME). These metasurfaces are designed to enhance spontaneous Raman scattering in discretized, sub-wavelength, cartesian grid-like "hot-spots" that span full tissue surface-areas, enhancing Raman scattering across the cell surface while minimizing cell heating damage for live cell imaging as well as fixed/frozen cell imaging.

We demonstrate single-cell differentiation and functional state phenotyping of various melanoma and immune cells found in TIME. Our ML models of Raman scattering yield a classification accuracy of >94% across all cell types. By stimulating cells with therapeutics and/or specific antigens, we demonstrate the ability to dynamically monitor cell plasticity, macrophage polarization, and T-cell activation state. Finally, we introduce spatial-omic Raman information acquired from tissue samples, resolving local cell identification and phenotyping. By extending our findings to whole tumor fixed formalin paraffin embedded (FFPE) and historical clinical specimens, our rapid, label-free tumor profiling methodology has potential to advance clinical biomarker efforts and impact the precision treatment decisions of cancer patients.

SCALABLE CO-SEQUENCING OF RNA AND DNA IN INDIVIDUAL CELLS

Peter A Sims[1,2], Timothy R Olsen[3], Romella K Sagatelian[1]

[1]Columbia University Irving Medical Center, Dept. of Systems Biology, New York, NY, [2]Columbia University Irving Medical Center, Dept. of Biochemistry and Molecular Biophysics, New York, NY, [3]Chan Zuckerberg Biohub New York, Advanced Single Cell Technologies Platform, New York, NY

Our understanding of somatic mosaicism and cancer has benefited from whole genome sequencing of individual cells. However, the interpretation of single-cell genomes in isolation is complicated by a lack of information about cellular identity and transcriptional state. Existing tools for jointly profiling DNA and mRNA from individual cells lack the throughput required for comprehensive analysis of complex tumors and tissues. We recently introduced a scalable method for jointly profiling DNA and expression following nucleosome depletion (DEFND-seq) that uses the widely available 10x Genomics droplet microfluidics platform. We have since developed both experimental and computational methods that further advance and improve DEFND-seq, particularly for applications in cancer biology.

MAPPING THE FUNCTIONAL CONSEQUENCES OF GENE AMPLIFICATION WITH SINGLE-CELL MULTIOMICS

Yue Wang[1], Oliver Cope[1], Jingting Chen[1], Caroline Tomblin[2], Joe Dahl[2], Alison Rojas[2], Tia Tate[2], Liz Brunk[1]

[1]University of North Carolina at Chapel Hill, Pharmacology, Chapel Hill, NC, [2]Bioskyrb Genomics, Inc., Durham, NC

Single-cell technologies are giving us a new lens to study the incredible diversity of cells within the same population. Few cells are more diverse than those carrying gene copy number alterations on extrachromosomal DNA (ecDNA). EcDNA are 2-5Mb circles of DNA that encode genes, replicate, and divide into daughter cells. Unlike normal chromosomes, ecDNA is unevenly inherited: when cells divide, some cells might receive ten copies of ecDNA while others inherit a thousand. When these amplified genes encode growth factors, as is common in cancer, they generate strikingly different cell states within the same tumor.

Until now, ecDNA has mostly been studied through bulk sequencing, which averages out this variation, or through imaging that cannot fully capture functional consequences. Only recently have single-cell multiomic technologies emerged that allow us to measure DNA and RNA together in the very same cells. In our study, we combine single-cell imaging with 10x Genomics multiome and BioSkryb ResolveOME to trace how copy number differences shape expression heterogeneity across populations in 6 cancer cell lines. We compare amplifications carried on ecDNA with those embedded as homogeneously staining regions (HSRs), uncovering distinct transcriptional outcomes.

Uniquely, by integrating BioSkryb and 10x platforms, we resolve isoform variants at single-cell resolution and link them to network-level effects, revealing which cancer-driving genes are most functionally important. This approach provides a path to connect DNA structure, isoform regulation, and cellular behavior, with broad implications for biomarkers, therapeutic strategies, and understanding how genomic architecture fuels cancer evolution and drug resistance.

GENOTYPE-SPECIFIC SAMPLE ENRICHMENT FOR RNA SEQUENCING OF RARE CELLS

Rodrigo Cotrim Chaves[1,2], Xiaohan Zhu[1], Aaron M Streets[1,2,3]

[1]University of California Berkeley, Bioengineering, Berkeley, CA, [2]UC Berkeley - UCSF Graduate Program in Bioengineering, Berkeley, CA, [3]Chan Zuckerberg Biohub SF, San Francisco, CA

Sample preparation for single-cell transcriptomic analysis often includes cell-sorting steps for enrichment of specific cell types of interest. However, sometimes the cell type of interest is defined by a genetic signature, and it is difficult to sort cells based on their genotype alone, when there is not a corresponding surface marker. To address this challenge, we have developed Seek-seq, a method in which the molecular contents of single cells are captured in hydrogel beads for PCR-activated sorting, allowing for sample enrichment based on genotype for downstream transcriptomic studies.

Using droplet microfluidic techniques, individual cells are encapsulated and lysed inside polyacrylamide beads. This hydrogel mesh creates a structure to capture and compartmentalize both genomic DNA and mRNA content of single cells, while allowing for diffusion-based buffer exchanges and multi-step reactions to take place. By performing reverse transcription with gel-immobilized polydT primers, cellular transcriptomic information is imprinted onto the hydrogel bead. Beads are then reinjected into droplets for digital droplet PCR (ddPCR) and hydrolysis-probe ddPCR chemistry is used to achieve a highly-sensitive and specific assay targeting a unique genomic locus. A novel probe design is used that produces a bead-immobilized fluorescent signal, allowing for aqueous-phase processing and sorting in conventional flow sorters. Sorted beads retain cellular transcriptomic information, and can be further processed for RNA-seq analysis.

While there are other methods available for genotypic-specific sample enrichment, Seek-seq uniquely leverages the gold standard genotyping assay of hydrolysis-probe ddPCR coupled with high-throughput processing in conventional sorting instruments. Thus, Seek-seq enables comparative transcriptomic analysis of samples with cells of heterogeneous genotypes, such as in infectious diseases or cancer, and is particularly well-suited for enrichment of rare cell genotypes for deep-sequencing studies.

ULTRASENSITIVE AND HIGH THROUGHPUT DETECTION OF VIRAL ISOFORM-OMES IN SINGLE HOST CELLS

Emanuelle Grody[1,2], Jinhee Kim[1], Jakob Harrison-Gleason[1], Ramon Lorenzo-Redondo[1], Elena Martinelli[1], Yogesh Goyal[1,2]

[1]Northwestern University, Feinberg Medical School, Chicago, IL, [2]Chan Zuckerberg Biohub Chicago, Chicago, IL

Alternative splicing is essential for replication of virus genomes that are dense with overlapping regulatory and coding regions. Human Immunodeficiency Virus (HIV) and its primate analog Simian Immunodeficiency Virus (SIV) transcripts exhibit remarkable diversity, with studies revealing more than 50 alternatively spliced isoforms. These isoforms together constitute the viral isoform-ome, the pool of viral RNA inside an infected cell or population. Although various viral isoforms have been implicated in infection progression and outcomes at the cell population level, less is known about the role of the viral isoform-ome in single host cell fate due to limited sensitivity and detection biases of standard single-cell RNA sequencing (scRNAseq) protocols. This limitation hampers our understanding of viral latency, persistence, and the host cell response to infection or therapeutics. To address this gap, we developed SALVEseq (Single-cell Amplified Libraries for Variant isoform Enrichment), a novel technology combining scRNAseq, targeted isoform amplification, and computational analysis. With SALVEseq, we have demonstrated capture of viral isoform-omes and molecular states of single host cells at high throughput during infection and latency. SALVEseq uses multiplexed PCR primers upstream of major splice sites to selectively enrich alternatively spliced viral isoforms while maintaining single-cell resolution through cell barcode preservation. With high specificity—accurate recapitulation of known expression profiles with no false positives in negative samples—SALVEseq recovered viral transcripts in over 4-fold more host cells compared to standard scRNAseq, revealing a previously hidden population of infected cells. Within host cells, SALVEseq recovered over 200-fold more viral splicing events than standard scRNAseq, including previously undocumented splice sites. SALVEseq enabled distinction between cells containing different classes of viral isoforms, and differential gene expression analysis of SALVEseq-identified infected cells revealed novel host genes and pathways, including signaling and senescence programs, that can be functionally linked to infection status. We have further demonstrated the utility of SALVEseq to study splicing in other systems, including liver regeneration. By providing new insights into viral splice programs within single host cells and their relation to infectivity and disease progression, this technology opens new avenues for understanding viral persistence, host-pathogen interactions, and the development of cure strategies for viral infections.

COALESCING SINGLE CELL ANEUPLOIDY AND TRANSCRIPTOMES AT SCALE TO DECODE BREAST CANCER PROGRESSION

Kaile Wang[1,2], Rui Ye*[1,3,4], Shanshan Bai[1], Zhenna Xiao[1], Lei Yang[1,4], Chengling Tang[1], Emi Sei[1], Jinyu Peng[2], Anna Casasent[1], Steven Lin[3,4], Chandandeep Nagi[5], Alastair Thompson[6], Savitri Krishnamurthy[7], Nicholas Navin[1,4,8]

[1]MD Anderson Cancer Center, Systems Biology, Houston, TX, [2]Shanghai Institute of Biochemistry and Cell Biology, University of Chinese Academy of Sciences, State Key Laboratory of RNA Innovation, Science and Engineering, CAS Center for Excellence in Molecular Cell Science, Shanghai, China, [3]MD Anderson Cancer Center, Houston, TX, [4]MD Anderson Cancer Center, Houston, TX, [5]Baylor College of Medicine, Pathology, Houston, TX, [6]Baylor College of Medicine, Surgery, Houston, TX, [7]MD Anderson Cancer Center, Anatomical Pathology, Houston, TX, [8]MD Anderson Cancer Center, Bioinformatics, Houston, TX

*: These authors contributed equally

Understanding epithelial lineages of breast cancer and genotype-phenotype relationships requires direct measurements of the genome and transcriptome of the same single cells at scale. To achieve this, we developed wellDR-seq, a high-genomic resolution, high-throughput method to simultaneously profile the genome and transcriptome of thousands of single cells. We profiled 33,646 single cells from 12 estrogen receptor-positive breast cancers, and identified ancestral subclones in multiple patients that showed a luminal hormone responsive lineage, indicating a potential cell-of-origin. In contrast to bulk studies, wellDR-seq enabled the study of subclone level gene-dosage relationships, which showed near-linear correlations in larger chromosomal segments and extensive variation at the single gene level. We identified dosage-sensitive and dosage-insensitive genes, including many breast cancer genes. We also identified sporadic copy number aberrations in non-cancer cells. Overall, these data reveal complex relationships between copy number and gene expression in single cells, improving our understanding of breast cancer progression.

REGVELO: GENE-REGULATORY-INFORMED DYNAMICS OF SINGLE CELLS

Zhiyuan Hu

Medical Research Institute, Wuhan University, Wuhan, China

RNA velocity has emerged as a popular approach for estimating near-future cell trajectories from single-cell unspliced-spliced mRNA ratios, but it typically assumes that genes act independently. Conversely, gene regulatory network (GRN) methods recover regulatory circuitry from transcription factors to target genes but lack dynamics.

To bridge this gap, we developed RegVelo, a deep-learning framework that couples splicing kinetics with GRN priors. To be specific, RegVelo leverages the Bayesian generative model that embeds a GRN weight matrix into the transcription rate of ODE-based splicing dynamics. The gene- and cell-specific spliced and unspliced abundances are estimated using a parallelizable ODE solver. RegVelo infers latent time, refined GRN weights, and the posterior velocity distribution. This further enables the quantification of intrinsic and extrinsic cell uncertainty.

Upon inferring velocity and GRNs, RegVelo enables in silico TF knockouts by removing the regulon connections and generating the perturbed velocity vector field. The cosine dissimilarity between the perturbed cell velocity and the original one reflects the local changes of each cell, termed the perturbation effect score. By further combining RegVelo's generative model with CellRank2's model-agnostic framework, we predict the effects of perturbations on cell fate decisions and screen key TF drivers of lineage specification.

We benchmarked RegVelo with simulated datasets and seven real-world datasets, including cell cycle, mouse pancreatic endocrinegenesis, and human hindbrain development. RegVelo inference of velocity, latent time, and GRNs outperformed dynamics-only and GRN-only methods across lineage complexity, prior GRN corruption levels, and sequencing depth, based on the evaluation of terminal state identification scores, cross-boundary correctness, and GRN AUROC metric.

Finally, applied to zebrafish neural crest data, RegVelo recapitulated trajectories and faithfully identified key lineage drivers, e.g., *tfec* as an early driver and *elf1* as a novel driver of pigment lineage, subsequently validated in vivo. Using in vivo Perturb-seq data as the ground truth, RegVelo significantly outperformed existing perturbation prediction models. Overall, our work provides a general framework to connect GRNs with developmental dynamics and guide targeted experiments.

DISSECTING GENE REGULATORY MECHANISM USING SINGLE-MOLECULE FOOTPRINTING

Ayelen Lizarraga[1], Kaixuan Luo[2], Xiaotong Sun[2], Diana Vera Cruz[1], Xin He[2], Sebastian Pott[1]

[1]The University of Chicago, Department of Medicine, Chicago, IL, [2]The University of Chicago, Department of Human Genetics, Chicago, IL

Gene regulation is a highly dynamic process tightly controlled in time and space, mediated through interactions of transcription factors (TFs), nucleosomes, and DNA methylation. A large body of work over the last decades has identified and characterized regulatory elements in the human genome and established the individual roles of TFs, nucleosomes, and DNA methylation in controlling gene expression. However, most studies rely on short-read sequencing technologies that fragment DNA and aggregate signal across cell populations. These methods also measure one modality at a time thereby obscuring the dynamic and coordinated nature of gene regulation. Single-molecule footprinting (SMF) approaches, such as Fiber-seq, use methyltransferase-based chromatin mapping to overcome these limitations. SMF simultaneously captures chromatin accessibility, DNA methylation, and TF binding across contiguous 10-20 kb molecules.
Here we present ongoing work focusing on two areas: (1) We applied SMF to identify multimodal quantitative trait loci (QTLs) in a cohort of 30 human lymphoblastoid cell lines, to detect genetic variants associated with novel regulatory modes including accessibility variability, nucleosome positioning patterns, and co-accessibility between distant elements. These single-molecule chromatin QTLs capture allele-specific effects with haplotype resolution, revealing mechanistic relationships invisible to population-based approaches. (2) To understand gene regulatory dynamics at the level of individual molecules, we profiled LPS-stimulated macrophages across a temporal activation time course. These data enable us to test how transcription factors cooperatively remove nucleosomes at regulatory elements and how DNA methylation modulates chromatin accessibility.
Together, SMF reveals previously invisible coordination between regulatory factors and provides mechanistic insights into how genetic variants affect chromatin organization and cellular activation.

CROSS-SPECIES SINGLE-CELL METHYLOME ATLASES IDENTIFIED EPITHELIAL CELLS AS THE DRIVER OF EPIGENETIC AGING IN HUMAN AND MOUSE KIDNEYS

Hyeonsoo Jeong[1], Blue B Lake[1], Dinh Diep[1], Qi Yan[1], Sanjay Jain[2], Kun Zhang[1]

[1]Altos Labs, Institute of Science, San Diego, CA, [2]Washington University School of Medicine, Department of Pathology, St Louis, MO

Chronic kidney diseases (CKD) are resulting from long-term stresses or injuries, as well as maladaptation of multiple cell types in the kidney. Epigenetic mechanisms have been implicated in CKD disease progression, yet the specific roles of individual cell types remain incompletely understood. Here, we generated a cross-species single-cell methylome atlas comprising 34,456 human kidney nuclei from 12 donors (5 healthy, 7 CKD) and 23,807 mouse nuclei from young (4-month) and aged (20-month) animals. Integration with single cell atlases of transcriptome and chromatin accessibility provided a high-resolution multiomic map of kidney cell types, including cell states associated with injury and repair within the proximal and distal tubules. We demonstrated that CKD kidney tissues exhibit accelerated epigenetic aging, which is primarily driven by kidney epithelial cells. Local coordinated DNA methylation in cell type-specific cis-regulatory regions reduced in injured epithelial cell states, and differentially methylated genes were enriched at loci lacking CpG islands. Using an integrated dual-omics approach to simultaneously profile DNA methylation and 3D chromatin conformation in the same cells, we identified chromatin conformation remodeling in injured tubule cells at euchromatin–heterochromatin boundaries, suggesting an erosion of facultative heterochromatin and increased epigenomic instability. Collectively, these findings advance our understanding of the epigenetic and chromatin remodeling events that may underlie non-resolving repair during CKD pathogenesis.

SINGLE-CELL RESOLUTION MAPPING OF NOTCH SIGNALING REVEALS PRINCIPLES OF SELF-ORGANIZATION

Akilandeswari Balasubramanian[1], Tim Dullweber[2], Tiger Lao[1], Zubenelgenubi Scott[1], Allyson Quinn Ryan[1], Shloka Jaisingh[1], Teun Huijben[1], Eduardo Hirata-Miyasaki[1], Rachel Banks[3], Emma Spikol[1], Jacqueline Fernandez[1], Yuening Liu[1], Shalin Mehta[1], Loic Royer[1], Anna Erzberger[2], Adrian Jacobo[1]

[1]Chan Zuckerberg Biohub San Francisco, Quantitative Tissue Morphogenesis, San Francisco, CA, [2]EMBL Heidelberg, Cell Biology and Biophysics Unit, Heidelberg, Germany, [3]Insitro, South San Francisco, CA

How cells coordinate their behaviors to self-organize into organs is a central question in developmental and single-cell biology. The zebrafish lateral line neuromast, composed of mechanosensory hair cells, support cells, and mantle cells, provides a model to study how fate decisions emerge from signaling. During the development and regeneration of neuromasts, support cells terminally differentiate into hair cells through a Notch-mediated process. These cell-fate decisions occur alongside proliferation and death events that alter the complex 3D structure of the organ. How Notch signaling coordinates cell fates in this changing environment remains an open question.

We established a framework for long-term live imaging of neuromast development across four days and, through deep learning–based segmentation and cell tracking, generated single-cell–resolution datasets capturing the population dynamics of all neuromast cells over time. Analysis revealed robust patterns of self-organization, where distinct cell types emerged in reproducible spatial arrangements and the characteristic ratio of cell types was maintained robustly.

Using this imaging data, combined with scRNA-seq (Lush et al., 2019) and Hybridization Chain Reaction (HCR), we developed a mathematical model and simulated Notch signaling dynamics within our datasets, incorporating the real cell–cell contact network. This model can predict and explain the spatiotemporal pattern of support cell differentiation into hair cells.

We validated these predictions experimentally by perturbing the system with small-molecule inhibition of Notch signaling and targeted single-cell ablations. Finally, we generated single-cell maps of Notch receptor and ligand gene expression using HCR combined with deep learning–based segmentation and quantification, and compared these data with the predictions of our model. Together, these results demonstrate that Notch feedback operates as a robust and dynamic cell selection mechanism for differentiation in the neuromast.

This multifaceted approach deepens our understanding of the role of Notch signaling during development and regeneration, and establishes a generalizable single-cell analysis framework for uncovering how cell–cell interactions shape fate decisions and drive robust pattern formation at the tissue scale.

COMPREHENSIVE ANALYSIS OF GENE EXPRESSION DYNAMICS ACROSS AN ANIMAL LIFE CYCLE

John I Murray[1], Christopher L Large[1,2], Rupa Khanal[1,2], Junhyong Kim[2]

[1]Perelman School of Medicine, University of Pennsylvania, Department of Genetics, Philadelphia, PA, [2]University of Pennsylvania, Department of Biology, Philadelphia, PA

Single cell approaches to measure gene expression genome-wide across all cells of developing embryos are revolutionizing the study of development. The nematode *Caenorhabditis elegans*, with its invariant and fully mapped cell lineage, provides an ideal system for in toto analysis of embryogenesis. We pioneered the use of whole-organism quantitative imaging measure transcription factor protein dynamics at single cell resolution in *C. elegans* embryos and extended this to genome scale with a lineage-resolved single-cell RNA-seq atlas covering gastrulation through terminal differentiation. We further developed a comparable atlas for the related nematode *C. briggsae* (as evolutionarily distant from *C. elegans* as human is from chicken), enabling systematic comparisons of cell-type-specific gene expression conservation and divergence, and identifying both spatial and temporal changes in gene expression across evolution. These resources are proving invaluable for discovering new developmental regulators and raising questions inaccessible without such comprehensive datasets.

We have now integrated our embryonic atlas with complementary datasets covering cells and nuclei from earlier embryonic stages. This results in a single combined atlas of embryogenesis, providing full gene expression trajectories from fertilization through differentiation, and allowing inference of transcriptional dynamics. Further integrating the embryo atlas with data from larval stages reveals multiple classes of cell identity maturation across the life cycle, including cyclic gene expression programs linked to the molt cycle (initiated in mid-embryogenesis) and stage-specific trends such as reduced expression of transcriptional regulators and elevated expression of cell fate effector genes postembryonically. These analyses emphasize the complexity of not just spatial, but also temporal regulation of gene expression across the organism.

NEXT-GENERATION, SAME-CELL ANALYSES INCORPORATING SINGLE-CELL PROTEOMICS

Amy E Herr[1,2]

[1]Chan Zuckerberg Biohub San Francisco, Investigator, San Francisco, CA,
[2]University of California, Berkeley, Bioengineering, Berkeley, CA

Even while new single-cell sequencing assays are emerging, single-cell genomics and transcriptomics tools have become standardized. As we consider the next wave of tools, including multi-modal tools, we are informed by the successes and challenges of microfluidic approaches in sequencing. Microfluidic technology is quietly operating 'under the hood' in the most powerful tools, exemplified by ubiquitous flow cells and droplet systems. In this talk, I will focus on two areas where we have gained design insight for approaching analysis that includes proteomics.

First, I will describe recent research from my lab that physically links together multiple, independent measurement modalities in a 'single-cell, same-cell' paradigm. Such so-called "joint analyses" are important to directly correlate different – but interrelated – layers of molecular information. These types of joint analyses may play important roles in generative models of cells and cellular systems, owing to low biological and technical noise. Here, I will describe a suite of approaches that allow us to interrogate the nuclear compartment (snATAC-seq) after isolating and archiving the cytoplasmic proteome. By design, our joint-analysis tools allow researchers to ex-post query the proteome of a unique originating cell, as informed by a priori same-cell sequencing. In essence, flipping the order of assay operations on its head. In our case, the proteome is not destroyed in the sequencing assays.

Second, I will share results from our lab's adoption of single-cell mass spectrometry proteomics, for both questions in mammalian and marine cell systems. In breast cancer, our work centers on protein-focused assays capable of whole-organoid profiling with single-cell resolution. Working with patient-derived organoids, we are seeking to understand the scrutinize the hypothesis of highly aggressive – yet rare -- cancer cell sub-populations (cell of origin). In marine symbiotic systems, our work centers on teasing apart the protein signaling pathways and regulatory networks that establish, sustain, and can disrupt coral-algae symbiosis (coral bleaching). While seemingly diverse biologies, a cellular- and even subcellular-level view highlights useful similarities.

SPATIAL SEQUENCING OF HOST-MICROBIOME INTERACTIONS

Iwijn De Vlaminck

Cornell University, Dept. of Biomedical Engineering, Ithaca, NY

Microbe-microbe and host–microbe interactions are central to the functioning of the human microbiome, but few tools are available to measure these interactions spatially within tissues. In this talk, I will introduce multiplexed imaging and spatial RNA sequencing approaches that enable simultaneous mapping of gut microbial communities, host gene expression and host gene regulation at single-micron resolution. By applying these methods, we reveal the longitudinal stability of spatial architectures in the healthy human oral plaque microbiome, demonstrate the location-dependent organization of microbial communities in the mouse intestine, and reveal interactions within and between microbial taxa at short length scales. I will also show how these tools reveal changes in microbial and host cell architectures at microbiome–tumor interfaces. Last, I will discuss recent extensions of these methods to enable spatial mapping of mobile genetic elements and their bacterial hosts in complex microbiomes. Together, these methods provide a new avenue to study host-microbiome interactions, with the potential for significant advances in microbiome science.

SINGLE CELL SPATIAL TRANSCRIPTOMICS ANALYSIS OF IMMUNE MICROENVIRONMENT IN NON-SMALL CELL LUNG CANCER

Anni Zhang[1], Troy Noordenbos[2], Maximilian Diehn[1]

[1]Stanford University, Department of Radiation Oncology, Stanford, CA,
[2]Stanford University, Department of Medicine, Stanford, CA

Lung cancer is the leading cause of cancer-related death, with a 5-year survival rate of only 28%. Non-small cell lung cancer (NSCLC) accounts for ~85% of cases, with adenocarcinoma (LUAD) as the most common subtype. Although immune checkpoint inhibitors (ICIs) have shown promise, only ~20% of patients respond, underscoring the need to understand resistance mechanisms. Tumor cells escape immune surveillance through progressive alterations in immune-cell composition and function, including impaired antigen presentation and reduced T-cell cytotoxicity. By comparing tumor regions with peritumoral normal regions, we can uncover these immune shifts and mechanisms of immunoediting. While single-cell studies have provided insights into transcriptional changes in NSCLC, most lack spatial context as they rely on dissociated cells. Single-cell spatial transcriptomics preserves tissue architecture while resolving cell states and ligand–receptor circuits *in situ*, enabling precise mapping of critical immune features, such as tertiary lymphoid structures (TLSs) and regulatory T cells (Tregs), which are key determinants of ICI responsiveness.

To investigate immune remodeling associated with LUAD tumorigenesis, we applied the Xenium single-cell spatial transcriptomics platform to 12 pretreatment LUAD specimens, including 9 with paired tumor-adjacent normal lung tissue. Using the Xenium Prime 5k gene panel plus 100 customized genes, we captured all major cell types previously identified in paired single-cell RNA sequencing data (scRNA-seq) from the same samples. Comparative analysis of tumor versus non-tumor subregions revealed marked shifts in immune composition and function. Specifically, B cells, plasma cells and T cells, particularly Tregs, were significantly enriched in tumors relative to adjacent normal tissue. We further identified 7 niche subtypes, each defined by distinct compositions of spatially adjacent cell types. Notably, TLS-like niches, containing aggregates of *CXCL13*+ T and *CXCL13*+ B cells were significantly enriched in tumor regions, likely suggesting an active immune response against tumor cells. To dissect functional differences between immune cells in tumor and normal regions, we will integrate scRNA-seq with spatial data using Tangram. Leveraging the greater depth of scRNA-seq, including broader gene coverage and higher transcript counts per cell, will allow us to resolve intracellular signaling pathways and cell–cell communication networks regulating immune–tumor interactions. Collectively, these findings illuminate immune remodeling in LUADs and may inform new strategies to improve durable response to ICIs in NSCLC.

VOLUMETRIC DNA MICROSCOPY FOR OPTICS-FREE THREE-DIMENSIONAL SPATIAL GENOMICS

Joshua Weinstein[1,2], Nianchao Qian[1]

[1]University of Chicago, Department of Medicine Section of Genetic Medicine, Chicago, IL, [2]University of Chicago, Pritzker School of Molecular Engineering, Chicago, IL

Generalizable methods for spatially resolving genomic and transcriptomic information across intact tissues in true three-dimensional context, without reliance on predefined reference genomes, spatial landmarks, or optical imaging, would expand the impact of single cell genomics on the biology idiosyncratic clinical specimens and non-model organisms. We present volumetric DNA microscopy, which encodes molecular proximities directly into DNA sequences through an entirely in situ chemical reaction, followed by standard sequencing. Molecular coordinates are then computationally reconstructed using geodesic spectral embeddings, a scalable manifold-learning method optimized for spatial-genomic datasets containing millions of uniquely identified molecules. Here we demonstrate volumetric DNA microscopy applied to whole zebrafish embryos across development, each yielding datasets comprising approximately 10 million spatially tagged transcripts per organism. The method recovers anatomical compartments, isoform expression patterns, and developmental gene-expression gradients at median local spatial resolutions of ~2 microns -- sufficient to distinguish neighboring cells. As sequencing libraries simultaneously capture cDNA sequences and spatial linkages, each dataset inherently integrates genome-wide expression profiles with a comprehensive three-dimensional map of cellular microenvironments. Volumetric DNA microscopy therefore provides a sequencing-native spatial-genomic platform ideally suited to large-scale analyses of cell-cell interactions in whole mount tissue contexts.

PHOTOLABILE OLIGONUCLEOTIDES COMBINED WITH TOPOLOGICALLY IMPOSED LIGHT GRADIENTS ENABLE SPATIALLY RESOLVED SINGLE-CELL TRANSCRIPTOMICS AND EPIGENOMICS

Robert Piscopio[1], Chieh Wang[2], Alex Chialastri[2], Mikolaj Godzik[1], Weiyue Wang[1], Lauren J Li[1], Kellie Heom[2], Neha Saxena[2], Maxwell Z Wilson[1], Siddharth S Dey[2,3]

[1]University of California Santa Barbara, Department of Molecular, Cellular, and Developmental Biology, Santa Barbara, CA, [2]University of California Santa Barbara, Department of Chemical Engineering, Santa Barbara, CA, [3]University of California Santa Barbara, Department of BioEngineering, Santa Barbara, CA

The spatial organization of cells within tissues is critical for biological function, development, and disease. Spatial transcriptomics enables gene expression profiling within intact tissue, yet most current methods are limited by low resolution, inflexible formats, and lack of compatibility with multiomic analysis. Few techniques can simultaneously capture the transcriptome and epigenome at single-cell resolution while preserving spatial context. To overcome these limitations, we present scSTAMP-seq (single-cell Spatial Transcriptomic And Multiomic Profiling), a light-based method that labels live or fixed cells with cholesterol-conjugated, photolabile oligonucleotides that incorporate into cell membranes. Patterned light exposure enables us to “stamp” the position of spatial coordinates into cells prior to sample dissociation and single-cell sequencing, enabling flexible, high-resolution encoding of spatial information across fields of view using discrete or gradient-based illumination patterns. scSTAMP-seq is modular, demonstrated by seamlessly integrating it with both plate- and droplet-based scRNA-seq, as well as a multiomic workflow that jointly profiles gene expression, DNA methylation, and chromatin accessibility from the same cells. In a human gastruloid model, the method reconstructed radial germ layer expression patterns and revealed spatial distributions of chromatin regulators. Collectively, these results establish scSTAMP-seq as a scalable and adaptable platform for spatially resolved, multiomic single-cell analysis, bridging spatial biology and single-cell genomics with high precision.

PHOTOTHERMAL CHEMICAL MICROSCOPY: BOND-SELECTIVE SINGLE CELL IMAGING AT SINGLE BIOMOLECULE SENSITIVITY

Jianpeng Ao (presenter), Ji-Xin Cheng

Boston University, Biomedical Engineering, Boston, MA

Thus far, imaging biomolecules inside a living cell largely relies on fluorescence, whereas the fluorescent labels often perturb the structure and function of host molecules. The labeling approach further inhibits discovery-driven studies of human specimens. While coherent Raman scattering microscopy allowed high speed label-free chemical imaging, the sensitivity is limited to millimolar level. Vibrational photothermal microscopy was invented to overcome these barriers (Nature Methods, Review, May 2025). In this platform technology, vibrational excitation of chemical bonds quickly relaxes into heat. The subsequent photothermal effect (temperature rise, refractive index modulation, and thermal expansion) is measured with a probe beam with extremely high sensitivity. Of the various modalities, mid-infrared photothermal (MIP) microscopy (Science Adv 2016) has allowed click-free imaging of azide-tagged small molecules in live cells (Science Adv 2024) and mapping of enzymatic activities (Nature Methods 2024). Stimulated Raman photothermal (SRP) microscopy (Science Adv 2023) has allowed in situ imaging of lipid rafts in the cellular environment. Shortwave infrared photothermal (SWIP) microscopy (Nature Photonics 2024) has allowed chemical analysis of single cells inside a tumor spheroid. Very recently, we have reached single biomolecule imaging sensitivity via photothermal modulation of interferometric scattering. This presentation aims to introduce the principle, instrumentation, and broad applications of photothermal chemical microscopy.

HIGH-RESOLUTION CHEMICAL AND TRANSCRIPTIONAL ANALYSIS OF WHOLE-MOUNT TISSUE SECTIONS

Bryce Manifold[1], Leon Han[2], Aaron Streets[1,2]

[1]UC Berkeley, Bioengineering, Berkeley, CA, [2]UC Berkeley, Center for Computational Biology, Berkeley, CA

Single cell spatial transcriptomics (ST) offers revolutionary insight to foundational biology and disease progression by contextualizing inter-cell gene expression heterogeneity within tissue architecture. Meanwhile coherent Raman microscopy has become an attractive method for label-free chemically quantitative microscopy with subcellular spatial resolution. The integration of these two technologies offers a significant potential in shaping the total leveraged information from samples, as well as improving ST throughput and resolution via in silico modeling. In this work we demonstrate the first combination of coherent Raman microscopy with capture-based transcriptomics in spleen. We show that the integration of these datasets mutually augments each technique respectively, offering a new perspective of biological insights.

MESOSCALE WHOLE CELL STRUCTURE ANALYSIS REVEALS SUBPOPULATIONS OF INSULIN SECRETORY GRANULES

Kevin Chang[1], Aneesh Deshmukh [1], Riva Verma[2], Valentina Loconte[3,4,5], Kate White [1,6]

[1]University of Southern California, Department of Chemistry, Los Angeles , CA, [2]University of Southern California, Department of Biomedical Engineering, Los Angeles, CA, [3]University of California San Francisco, Department of Anatomy, San Francisco, CA, [4]Lawrence Berkeley National Laboratory, Molecular Biophysics and Integrated Bioimaging Division, Berkeley, CA, [5]Diamond Light Source , Harwell Science and Innovation Campus, Didcot, United Kingdom, [6]University of Southern California, Department of Quantitative and Computational Biology, Los Angeles, CA

The advent of whole-cell datasets at mesoscale resolution (tens of nanometers to 10 microns) makes it possible to inventory the vast community of organelles within a single cell. While advancements in soft X-ray tomography (SXT) and volume electron microscopy (vEM) technologies have expanded our ability to map cells in their entirety, analytical strategies capable of fully exploiting these data remain underdeveloped. To bridge the disconnect between experimental sophistication and downstream analysis, we adopt the conceptual framework of morphomics - the systematic, unbiased, and quantitative exploration of cellular structure. Inspired by other “omics” methodologies, we use unsupervised clustering to identify subpopulations of insulin secretory granules (ISGs), key sites of insulin processing and storage in pancreatic β-cells. In SXT tomograms, this approach reveals distinct ISG subpopulations based on molecular density, size, and spatial positioning. Across different insulin secretory stimuli conditions, we observe shifts towards more biochemically dense and secretion-ready subtypes, demonstrating that exogenous signals can dynamically remodel ISG subpopulation distributions. To broaden this analytical framework, we extend this methodology to three-dimensional datasets collected using volume electron microscopy (vEM). Integrating vEM subpopulations with SXT subpopulations from primary β-cells uncovers insights inaccessible by either method alone. This holistic strategy allows us to define ISG subpopulations from whole-cell structures, potentially providing a basis for therapeutic approaches that can enrich physiologically beneficial ISG subpopulations. We envision that our framework will be a broadly applicable strategy for organelle subpopulation discovery and the analysis of cellular systems, generalizable across whole cell methods and cell types.

INTERPRETABLE REPRESENTATION LEARNING FOR SINGLE CELL TRANSCRIPTOMICS

Bianca Dumitrascu

Columbia University, New York, NY

Single-cell RNA-seq enables the study of cell states across diverse biological conditions, such as aging, drug treatments, and tissue injury. However, disentangling shared and condition-specific transcriptomic patterns remains a significant computational challenge, particularly in settings with missing data or complex experimental designs. In this talk, I will introduce new deep generative frameworks designed to disentangle these transcriptomic signals, allowing for robust integration, cross-condition prediction, and biologically interpretable insights. Using real and simulated scRNA-seq datasets, we demonstrate that disentanglement can uncover shared wound healing patterns and distinct changes in cell behavior, including age-dependent immune responses and drug-modulated extracellular matrix remodeling. Finally, I will discuss open problems towards a pipeline for synthetic self-organizing systems.

MULTI-MODAL SINGLE-CELL ANALYSIS: SELF-SUPERVISED VISION MODELS REVEAL HIDDEN MORPHOLOGY-TRANSCRIPTOME RELATIONSHIPS ACROSS DIVERSE BIOLOGICAL SYSTEMS

Shreya S Deshmukh[1], Mehdi Mohseni[1], Shan Sabri[1], Richard Yau[1], Yining Chen[2], Makenzie Sacca[1], Nirvan Rouzbeh[1], Olaia Vila[1], Gaetano Viscido[2], James Walker[3], Joseph Ecker[3], Andrea Califano[2], Pier F Gherardini[1], Gary Schroth[1]

[1]Cellanome, Inc., R&D, Foster City, CA, [2]CZ Biohub New York, New York, NY, [3]Salk Institute for Biological Sciences, Department of Plant Biology, La Jolla, CA

Single-cell transcriptomics has revolutionized our understanding of cellular heterogeneity, yet morphology, a fundamental determinant of cellular function, remains disconnected from sequencing. We demonstrate how DINOv2, a self-supervised vision transformer, can extract high-dimensional morphological features from images of cells in Cellanome's CellCage™ enclosures and integrate them with transcriptomics from the same single cells to reveal hidden cellular phenotypes.
We first established the discovery potential of DINOv2 with images from six cell lines, showing that model embeddings cluster by identity, confirming the method captures meaningful morphological differences. We then showed that the generalization ability of a Mask R-CNN-based segmentation model depends on the similarity, in DINOv2 embedding space, between cells in the training and test sets. Further, the embeddings were robust under common image augmentations (e.g. rotation, blur), supporting their integration with transcriptomic data.
Applying this framework across diverse biological systems revealed striking morphology-transcriptome relationships, invisible to single-modality analyses. In cultured fibroblasts, clustering of DINOv2 embeddings identified four morphologically distinct populations corresponding to various states of adhesion. Fully-adhered and elongated fibroblasts were enriched for ECM proteins (CCN2, ST3GAL5, LOX), weakly adhered cells expressed inflammatory chemokines (CCL20, IL24), and intermediate states showed mixed adhesion-stress signatures. These phenotypes were undetectable by transcriptomic clustering alone.
In plant protoplasts, DINOv2 distinguished three populations: healthy round cells enriched for photosynthesis-related genes, stressed parabolic cells with elevated oxidative phosphorylation, and small guard cells involved in abscisic acid signaling. Finally, using a THP-1 monocyte differentiation system, DINOv2 revealed morphological heterogeneity correlating with polarization pathways. Specific morphological phenotypes – capturing cell shape, texture, and pseudopodia – were enriched in M1 and M2 conditions and upregulated in chemokine signaling, macrophage activation, and cytoskeletal reorganization genes.
Our results demonstrate that DINOv2 features capture functionally relevant morphological variation that reflects underlying expression programs. This approach complements transcriptomic analysis, enabling discovery of morphology-defined cellular subpopulations and mechanistic insights into the gene expression programs driving distinct morphological states. In conclusion, integrated multi-modal datasets generated with Cellanome's technology provide new insights into cell state transitions, stress response, and differentiation processes across biological systems, establishing a powerful paradigm for understanding cellular heterogeneity beyond transcriptomics alone.

INTEGRATED TRANSCRIPTOMIC AND MORPHOLOGICAL DEEP LEARNING FOR SPATIALLY INFORMED SINGLE-CELL ANALYSIS OF PANCREATIC CANCER PROGRESSION

Nischal Bhandari[1], Huaien Wang[1], Jose Reyes[2], Sandeep Nadella[1], James M Crawford[3], Gabriele Campanella[4], Zhen Zhao[1,3]

[1]Cold Spring Harbor Laboratory, Cancer Center, Cold Spring Harbor, NY, [2]Memorial Sloan-Kettering Cancer Center, Cancer Biology and Genetics Program, New York, NY, [3]Zucker School of Medicine at Hofstra/Northwell, Department of Pathology, Manhasset, NY, [4]Mount Sinai, Artificial Intelligence and Human Health, New York, NY

Spatial transcriptomics preserves tissue architecture for single-cell biology, and when integrated with morphology-based deep learning, expands the potential of single-cell analysis for histopathologic diagnosis. In a pilot study of 8 tissue sections from 2 pancreatic cancer patients and 1 non-cancer donor, we combined high-resolution spatial transcriptomics (10x Genomics Visium HD and Xenium) with a novel H&E-based morphological foundation model to interrogate single-cell biology across the spectrum of neoplastic progression. This approach reconstructed a continuum of epithelial transitions, from normal ductal and acinar cells through reactive and atypical intermediates to neoplastic precursors and pancreatic ductal adenocarcinoma, captured at single-cell resolution while preserving spatial context. Intermediate cell populations with progenitor-like transcriptional programs were attempted, providing translational validation of tumor-initiating populations suggested by mouse models. Spatial mapping further revealed distinct stromal programs that co-evolve with epithelial changes, highlighting fibroinflammatory niches spatially linked to tumor progression. Notably, morphological features aligned to 8-micron transcriptomic spots correlated with gene expression and distinguished reactive, atypical, and neoplastic states (AUC 0.69–0.71) without additional training. By integrating molecular profiles, morphology, and spatial context within human specimens, this framework illuminates the cellular and microenvironmental dynamics of early pancreatic tumorigenesis, offering new opportunities for biomarker discovery, diagnostic precision, and therapeutic intervention.

MODELING SPATIAL VARIATION USING INTERPRETABLE NEURAL REPRESENTATIONS

Uthsav Chitra[1,2,3], Clover Zheng[1], Cong Ma[1,4], Benjamin J Raphael[1]

[1]Princeton University, Dept. of Computer Science, Princeton, NJ, [2]Johns Hopkins University, Dept. of Computer Science, Baltimore, MD, [3]Johns Hopkins University, Data Science and AI Institute, Baltimore, MD, [4]University of Michigan, Dept. of Computational Medicine and Bioinformatics, Ann Arbor, MI

Recent spatial omics technologies measure mRNA, protein, metabolite, and other modalities at thousands of locations within tissue sections, revealing spatial patterns of cell types and molecular activity. However, current datasets are often sparse and incomplete due to technological and cost constraints. We developed multiple machine learning approaches to overcome this sparsity by modeling the latent geometry of individual tissue slices and by integrating measurements across multiple samples/modalities. GASTON and Multi-GASTON utilize deep neural networks to learn continuous spatial coordinate(s) that describe the major axes of spatial variation across individual or multiple tissue slices, enabling the identification of spatial expression gradients for individual transcripts. We apply these methods to analyze spatial variation in cell types and gene expression in normal tissues and derive gene expression gradients within tumor microenvironments.

PREDICTION OF PROTEIN SUBCELLULAR LOCALIZATION IN SINGLE CELLS

Xinyi Zhang*[1,2], Yitong Tseo*[2], Yunhao Bai[1], Fei Chen[1,3], Caroline Uhler[1,2]

[1]Broad Institute of MIT and Harvard, Cambridge, MA, [2]Massachusetts Institute of Technology, Cambridge, MA, [3]Harvard University, Cambridge, MA

The subcellular localization of a protein is important for its function, and its mislocalization is linked to numerous diseases. While atlas-scale efforts have been made to spatially profile thousands of proteins in multiple cell lines, the number of possible combinations of proteins and cell lines far exceeds what have been measured even in the largest subcellular localization datasets. Furthermore, given the observed variability of protein localization not just across cell lines but also across single cells within a cell line, a measurement in an existing atlas might not accurately reflect localization in a new experiment with different biological context that might change protein localization. However, existing computational models for protein localization prediction either miss cell-type specificity or cannot generalize to unseen proteins or unseen cell lines. Here we present a method for **P**rediction of **U**nseen **P**roteins' **S**ubcellular localization (PUPS). PUPS combines a protein language model and an image inpainting model to utilize both protein sequence and cellular images. These two ingredients allow our model to predict the localization of any protein in a particular single cell: The protein sequence input allows generalization to unseen proteins and the landmark stains allow cell-type-specific predictions that capture single-cell variability and can generalize to unseen cell lines. Experimental validation shows that PUPS can predict protein localization in newly performed experiments outside the Human Protein Atlas used for training. The ability of PUPS to generalize to unseen proteins and cell lines enables us to assess the variability in protein localization across cell lines as well as across single cells within a cell line, which has been difficult given that existing atlases only cover a small subset of all protein and cell line pairs. We identified that the proteins with the most variable ratios between nucleus and cytosol across cell lines are associated with transcription, cell differentiation and chromatin regulation, whereas high single-cell variability in protein localization within a cell line is mainly associated with cell division, transcription, double-strand break repair and apoptosis. Collectively, PUPS provides a framework for predicting differential protein localization across cell lines and single cells within a cell line, including changes in protein localization driven by mutations.

*Authors contributed equally.

MULTI-SCALE EXPLORATORY SPATIAL DATA ANALYSIS IN LUNG ADENOCARCINOMA

Lambda Moses, Bianca Dumitrascu

Columbia University, Department of Statistics, New York, NY

Spatial phenomena, from weather to tissue morphology, occur at various length scales. Spatial -omics technologies also measure gene expression at a wide range of spatial scales. Transcriptome-wide technologies with lower resolution such as Visium are often paired with single cell resolution technologies that require a limited gene panel such as Xenium in the same study as they provide complementary information. However, the same data analysis methods are often applied to data from technologies with different scales without considering implications of scales. To explore such implications, we apply exploratory spatial data analysis (ESDA) methods such as the univariate Moran's I and bivariate Lee's L on Visium and Xenium data from adjacent sections of lung adenocarcinoma (LUAD) samples. These samples come from a comprehensive study spanning multiple patients through stages of LUAD. We have considered how and why ESDA results may change with scale and have identified scale-dependent gene programs and pathways that may relate to advancement of LUAD. This multi-scale ESDA approach has been implemented in an R package Wayfarer, to be submitted to Bioconductor.

DEVELOPING INSTRUMENTED TISSUES FOR SPATIOTEMPORAL OMICS

Shana Kelley[1,2]

[1]Northwestern University, Chemistry, Evanston, IL, [2]Chan Zuckerberg Biohub Chicago, N/A, Chicago, IL

We are developing engineering-driven technologies to enable direct, high-resolution measurement of molecular and cellular activity within human tissues. By embedding arrays of microsensors and sampling probes in tissue environments, we aim to continuously monitor biochemical and biophysical signals and define how their perturbation drives inflammation and immune dysregulation. These tissue-integrated platforms are coupled with multi-omics profiling and AI-based analysis to generate comprehensive, dynamic models of inflammatory processes.

Our initial efforts focus on human skin tissue, which provides an accessible and physiologically relevant model to dissect inflammatory dynamics. While advances in genomics and single-cell biology have revealed much about the behavior of individual cells, understanding how billions of cells interact within intact tissues remains a major challenge. Our integrated platforms are enabling the first holistic and real-time measurements of immune cell activity within tissue, allowing us to identify molecular tipping points that lead to pathological inflammation. Ultimately, this approach will inform new strategies to predict, prevent, and control inflammation-driven diseases.

PERMUTATION ENHANCES THE RIGOR OF SINGLE-CELL DATA ANALYSIS

Jingyi Jessica Li[1,2]

[1]Fred Hutchinson Cancer Center, Biostatistics Program, Seattle, WA,
[2]University of California, Los Angeles, Department of Statistics and Data Science, Los Angeles, CA

Ensuring the reliability and accuracy of single-cell data analysis is critical, particularly in visualizing complex biological structures and addressing data sparsity. This talk introduces two novel statistical methods—scDEED and mcRigor—that leverage permutation-based techniques to enhance the rigor of these analyses. scDEED (Xia et al., 2024, Nature Communications) addresses the challenge of evaluating the reliability of two-dimensional (2D) embeddings produced by visualization methods like t-SNE and UMAP, which are commonly used to visualize cell clusters. These methods, however, can sometimes misrepresent data structure, leading to erroneous interpretations. scDEED calculates a reliability score for each cell embedding, comparing the consistency between a cell's neighbors in the 2D embedding space and its pre-embedding neighbors. Cells with low reliability scores are flagged as dubious, while those with high scores are deemed trustworthy. Additionally, scDEED provides guidance for optimizing t-SNE and UMAP hyperparameters by minimizing the number of dubious embeddings, significantly improving visualization reliability across multiple datasets. mcRigor focuses on enhancing metacell partitioning in single-cell RNA-seq and ATAC-seq data analysis, a common strategy to address data sparsity by aggregating similar single cells into metacells. Existing algorithms often fail to verify metacell homogeneity, risking bias and spurious findings. mcRigor introduces a feature-correlation-based statistic to measure heterogeneity within a metacell, identifying dubious metacells composed of heterogeneous single cells. By optimizing metacell partitioning algorithm hyperparameters, mcRigor enhances the reliability of downstream analyses. Moreover, mcRigor allows for benchmarking and selecting the most suitable partitioning algorithm for a dataset, ensuring more robust discoveries. scDEED and mcRigor demonstrate the power of permutation-based approaches in refining single-cell data analysis, providing researchers with tools to achieve more accurate and reproducible insights into complex cellular processes.

ENTROPY METRICS FOR BIOLOGICAL EXPLORATION IN HIGH-DIMENSIONAL scRNAseq DATA

Nikhila P Swarna[1], Lior Pachter[1,2]

[1]California Institute of Technology, Division of Biology and Biological Engineering, Pasadena, CA, [2]California Institute of Technology, Department of Computing and Mathematical Sciences, Pasadena, CA

We introduce ember (entropy metrics for biological exploration), a method that leverages information entropy to quantify variation in high-dimensional, multivariate single-cell RNAseq. Applying ember to 8 tissues from 8 individuals for each of the 8 founder mouse strains, we discovered that genetic background impacts the cell type specificity of a marker gene. More broadly, we learn that tissue specificity dominates genotype specificity in mice. In the developing mouse kidney, ember uniquely identifies genes specific to non-consecutive developmental stages. We also show that the generated entropy metrics can be used in tandem with unsupervised and supervised machine learning algorithms to discover clusters of genes that share specificty and infer gene regulatory networks within the data. Together, these applications establish ember as a powerful tool and provide a roadmap for elucidating the impact of genetic variation using a mouse model.

HEIMDALL: A MODULAR FRAMEWORK FOR TOKENIZATION IN SINGLE-CELL FOUNDATION MODELS

Ellie Haber*[1], Shahul Alam*[2], Nicholas Ho*[2], Renming Liu[2], Evan Trop[2], Shaoheng Liang[2], Muyu Yang[2], Spencer Krieger^[2], Jian Ma^[2]

[1]1, Machine Learning Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, [2]2, Ray and Stephanie Lane Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA

Foundation models trained on single-cell RNA-sequencing data (scRNA-seq) have rapidly emerged as powerful tools for single-cell analysis. Yet their performance critically depends on how cells are tokenized into model inputs – a design space that remains poorly understood. Here, we introduce HEIMDALL, a comprehensive framework and open-source toolkit for systematically evaluating tokenization strategies in single-cell foundation models (scFMs). HEIMDALL decomposes each scFM into modular components: a gene identity encoder (F_G), an expression encoder (F_E), and a "cell sentence" constructor (F_C). This modularization enables controlled ablations that disentangle the effects of tokenization from confounding factors such as model size or pretraining corpus. Using a transformer trained from scratch, we evaluate tokenization strategies on cell type classification in two challenging transfer-learning settings: gene generalization, simulating limited gene overlap across species or technologies, and cross-tissue generalization, testing annotation transfer across tissue types. Our results show that tokenization choices exert little influence in-distribution but become decisive under distribution shifts, with F_G driving the largest performance gains. HEIMDALL further demonstrates how existing strategies can be recombined to improve generalization. By standardizing evaluation and providing an extensive library, HEIMDALL establishes a foundation for reproducible, systematic exploration of single-cell tokenization and accelerates the development of the next generation of scFMs.

* Authors contributed equally
^ Corresponding authors

ATLAS-BASED APPROACH FOR ANALYZING HIGH-RESOLUTION SPATIAL TRANSCRIPTOMICS WITHOUT FEATURE SELECTION

H. Robert Frost

Dartmouth College, Biomedical Data Science, Hanover, NH

Advances in spatial transcriptomics (ST) have transformed tissue biology, with high-resolution assays (e.g., 10x Visium HD, Xenium) now profiling thousands of genes across millions of spatial locations. However, resolution growth (e.g., ~5k 55 μm spots in Visium vs. ~10M 2 μm spots in Visium HD) has created a severe gap between data size and the scalability of analysis pipelines. Current strategies address this by reducing both locations (via sketching/aggregation) and genes (via feature selection) before PCA-based dimensionality reduction, often paired with on-disk storage (e.g., BPCells). Downstream tasks like clustering are then performed on the reduced embedding. While feasible, this approach risks discarding biologically important signals, since genes excluded before dimensionality reduction cannot be recovered. Moreover, even with these approximations, nonlinear analyses using deep learning remain impractical for high-resolution ST data.
A key opportunity lies in the rapid growth of single-cell atlases such as the Human Cell Atlas and CELLXGENE. These resources already contain >100M human scRNA-seq profiles, with projects like the CZI Billion Cell initiative poised to expand by an order of magnitude. While widely used for label transfer and foundation models (e.g., scGPT, scFoundation, CellFM), their potential for improving computational efficiency in ST pipelines remains underexplored.

We are developing atlas-based dimensionality reduction approaches that eliminate the need for feature selection while enabling efficient gene-level analyses through reduced rank reconstruction. This talk will highlight three strategies:
1) Atlas-guided linear dimensionality reduction: Using low-dimensional embeddings from atlas scRNA-seq data (e.g., Tabula Muris, Tabula Sapiens) to initialize truncated SVD (e.g., IRLBA) reduces iteration counts by nearly two orders of magnitude. This enables full-gene linear dimensionality reduction of high-resolution ST data without feature selection, and potentially without sketching.
2) Atlas-guided nonlinear dimensionality reduction: A novel SVD–autoencoder fusion model pretrained on atlas data and fine-tuned on ST achieves computationally feasible nonlinear embedding and reconstruction without feature selection.
3) Foundation model–based dimensionality reduction: Transformer-based FMs trained on tens of millions of scRNA-seq profiles encode deep transcriptional knowledge but are typically too large to integrate efficiently. By using FM gene embedding matrices to initialize truncated SVD, we achieve the speed of linear methods while generating embeddings that capture richer biological variation.

Together, these approaches leverage atlases and foundation models to overcome computational bottlenecks in high-resolution ST, enabling scalable analyses without compromising biological resolution.

MENINGIOMA CELL REPROGRAMMING AND MICROENVIRONMENT INTERACTIONS UNDERLIE BRAIN INVASION

Ayush Aggarwal*[1,2,3], Mark W Youngblood*[4], Thiebaud Picart[1,5], Hinda Najem[4], Martha A Cady[1,3,6], Stephen T Magill[4], Craig M Horbinski[4,7], James P Chandler[4], Amy B Heimberger[4], Kanish Mirchia[2], Shawn L Hervey-Jumper[1,8], David R Raleigh[1,2,3]

[1]University of California San Francisco, Department of Neurological Surgery, San Francisco, CA, [2]University of California San Francisco, Department of Pathology, San Francisco, CA, [3]University of California San Francisco, Department of Radiation Oncology, San Francisco, CA, [4]Northwestern University, Department of Neurological Surgery, Chicago, IL, [5]University Claude Bernard Lyon, Department of Neurosurgery, Lyon, France, [6]University of California San Francisco, Tetrad Graduate Program, San Francisco, CA, [7]Northwestern University, Department of Pathology, Chicago, IL, [8]University of California San Francisco, Weill Institute for Neurosciences, San Francisco, CA
*authors contributed equally

Background
Meningioma brain invasion encumbers surgical resection and increases the risk of tumor recurrence, but the molecular mechanisms underlying this process are poorly understood.
Methods
To identify molecular and cellular features of brain-invasive meningiomas, we (1) analyzed bulk RNA sequencing data from 199 meningiomas, including 33 brain-invasive tumors, (2) analyzed patient-matched single-cell RNA sequencing data of spatially mapped meningioma samples from the tumor core or brain-tumor interface (BTI), and (3) performed spatial transcriptomic sequencing of brain-invasive meningioma samples. Multiplexed immunofluorescence (IF) was used to validate bioinformatic spatial expression patterns. Functional interactions between meningioma cells and neurons were studied in meningioma/neuron co-cultures using confocal microscopy, multi-electrode array recordings, and live cell calcium imaging.
Results
Transcriptomic analyses showed conserved enrichment of TGM2, S100A11, ZYX, and PDGFRA at the BTI across bulk, single-cell, and spatial RNA sequencing datasets. The expression of these genes at the BTI was confirmed using multiplexed IF. Single-cell bioinformatic and imaging analyses further demonstrated enrichment of macrophages at the BTI. Co-culture assays showed neuronal hyperexcitability and increased proliferation of meningioma cells, suggesting functional communication between meningioma cells and the tumor microenvironment may contribute to meningioma growth in cases with brain invasion.
Conclusions
Meningioma brain invasion is defined by molecular remodeling of tumor cells and functional interactions within the tumor microenvironment.

DEFINING THE CELLULAR NICHES OF AMYLOID PLAQUES AND NEUROFIBRILLARY TANGLES IN ALZHEIMER'S DISEASE USING HIGH-RESOLUTION SPATIAL TRANSCRIPTOMICS

Ekaterina Aladyeva[1], Nicholas Sweeney[2], Hoang-Tuong Nguyen-Hao[3], Sidney Lenz[1], Jacqueline Kaczaral[1], Hongjun Fu[2], Greg T Sutherland[3,4], Oscar M Harari[1]

[1]The Ohio State University, Department of Neurology, College of Medicine, Columbus, OH, [2]The Ohio State University, Department of Neuroscience, College of Medicine, Columbus, OH, [3]The University of Sydney, Sydney Brainomics, School of Medical Sciences and Charles Perkins Centre, Faculty of Medicine and Health, Sydney, NSW, Australia, [4]The University of Sydney, New South Wales Brain Tissue Resource Centre, School of Medical Sciences and Charles Perkins Centre, Faculty of Medicine, Sydney, NSW, Australia

Alzheimer's disease (AD) is a leading cause of dementia worldwide, with projections indicating a dramatic increase due to the growing aging population. Neurodegeneration in AD is linked to the misfolding of amyloid beta and phosphorylated tau proteins, resulting in extracellular amyloid-β plaques and intracellular neurofibrillary tangles (NFTs) — the two hallmark pathologies of AD. In a healthy brain, microglia, the resident immune cells, detect and remove protein accumulations before they aggregate and cause neuroinflammation, but these mechanisms are disrupted in AD patients. Understanding the precise regional and cellular distribution of these pathologies, along with their interactions with glial cells, is crucial for comprehending disease progression and developing effective therapies.

In this study, we performed high-resolution spatial transcriptomics using the 10X Visium HD platform on post-mortem human brain tissue from AD patients to define distinct cellular niches associated with amyloid-β plaques and NFTs. The study analyzed four samples from multiple brain regions at different disease stages. We developed a comprehensive pipeline combining multiplexed immunofluorescent staining image analysis with matching spatial transcriptomics data to detect AD-associated protein accumulations.

Our approach successfully identified specific cell populations and their transcriptional states, which are spatially correlated with amyloid plaque and NFT pathologies. Notably, we discovered two distinct microglia populations expressing different neuroinflammation pathways in proximity to amyloid beta plaques.

These findings provide novel insights into the landscape of AD and may inform future therapeutic approaches targeting specific cellular states associated with disease pathology. In our future work, we plan to integrate control samples to perform spatial differential expression analysis, thereby enabling a more comprehensive investigation into the association between distinct microglial populations and disease progression.

INTEGRATED HIGH-RESOLUTION SINGLE-CELL AND SPATIAL TRANSCRIPTOMICS OF CRANIOFACIAL DEVELOPMENT

Bremy Alburquerque[1], Ana S Gonzalez-Reiche[1], Tochukwu T Iyke-Nzeocha[2], Rajay Kamath[2], Carlos Quintana Martinez[2], Cody Frazer[2], Mithun Saha[2], Frederick R Foster[3], Joan T Richtsmeier[3], Susan M Motch Perrine[3,4], Kazuhiko Kawasaki[3], Meng Wu[1,2], Greg Holmes[1], Ethylin W Jabs[1,2], Harm van Bakel[1]

[1]Icahn School of Medicine at Mount Sinai, New York, NY, [2]Mayo Clinic, Rochester, MN, [3]Pennsylvania State University, University Park, PA, [4]Pennsylvania State College of Medicine, Hershey, PA

Suturogenesis and mandibular morphogenesis are fundamental processes in craniofacial development, essential for the coordinated growth of the skull and brain. Genetic dysregulation in associated pathways is among the most common causes of congenital anomalies. For example, premature fusion of skull bones, or craniosynostosis, is observed in ~1 in every 2,500 births. Micrognathia, characterized by a small, underdeveloped jaw, is estimated to occur in ~1 in every 1,500 births. Pathogenic variants in signaling pathways such as FGF, BMP, WNT, and SHH have been implicated in these conditions. However, their effects on specific cell populations, spatial organization, and the interactions that govern normal versus aberrant craniofacial development are not fully understood.

We performed single-cell RNA sequencing on coronal suture tissues from a mouse model of Apert syndrome, the *Fgfr2*$^{+/S252W}$ mouse, and from mandibular tissues of wild-type embryos, spanning embryonic day 12.5 (E12.5) to postnatal day 0 (P0). Several of these time points were also analyzed using Visium HD spatial transcriptomics. We developed a novel analysis pipeline that integrates single-cell and spatial data, enabling cell-cell communication analysis and spatiotemporal expression profiling at cellular resolution. The workflow supports interactive selection of specific regions of interest, facilitating analysis across a wide range of spatial scales. We illustrate this at the smaller, cellular scale in bone growth fronts at the coronal suture, and at the larger, multi-cellular scale in bone and cartilage of the developing mandible.

Using this approach, we characterized transcriptomic changes during the early stages of premature suture closure in the *Fgfr2*$^{+/S252W}$ Apert syndrome mouse, revealing the impact of these changes on cell-cell communication within the spatial context. Additionally, our integrated analyses uncovered gene expression signatures of major cell populations in their dynamic spatial domains during the development of mandibular symphyseal secondary cartilage. Spatiotemporal trajectory analysis revealed a chondrogenic process distinct from the differentiation in primary cartilages. This work highlights the utility of combining single cell and spatial transcriptomics to study craniofacial development and disease.

TREE RECONSTRUCTION GUARANTEES FROM CRISPR-CAS9 LINEAGE TRACING DATA USING NEIGHBOR-JOINING

Sebastian Prillo[1], Kevin An[1], Wilson Wu[1], Ivan Kristanto[4], Matthew G Jones[5], Yun S Song[1,3], Nir Yosef[1,2]

[1]University of California, Berkeley, Department of Electrical Engineering and Computer Sciences, Berkeley, CA, [2]Weizmann Institute of Science, Department of Systems Immunology, Rehovot, Israel, [3]University of California, Berkeley, Department of Statistics, Berkeley, CA, [4]University of California, Berkeley, Department of Molecular and Cell Biology, Berkeley, CA, [5]Stanford University, Center for Personal Dynamic Regulomes, Stanford, CA

CRISPR-Cas9 based lineage tracing technologies have enabled the reconstruction of single-cell phylogenies from transcriptional readouts. However, developing tree-reconstruction algorithms with theoretical guarantees in this setting is challenging. In this work, we derive a reconstruction algorithm with theoretical guarantees using Neighbor-Joining (NJ) on distances that are moment-matched to estimate the true tree distances. We develop a series of tools to analyze this algorithm and prove its theoretical guarantees. When the parameters of the data generating process are known and there is no missing data, our results align with established results from common evolutionary models, such as Cavender-Farris-Neyman and Jukes-Cantor. However, to account for the realistic case where the parameters of the data generating process are not known and there is missing data, we develop new theory that shows for the first time that it is still possible to obtain reconstruction guarantees in the CRISPR-Cas9 case and in other models of evolution. Empirically, we show on both simulated lineage tracing data and on real data from a mouse model of lung cancer the improved performance of our method as compared to the traditional use of NJ.

IDENTIFICATION OF SEX-DIFFERENTIAL AND HORMONE-SENSITIVE NEURONAL POPULATIONS IN THE AGING MOUSE BRAIN

Rebecca L Anderson, Veronika Kulik, Simón(e) Sun, PhD, Jessica Tollkuhn, PhD

Cold Spring Harbor Laboratory, Neuroscience, Cold Spring Harbor, NY

Neurodegenerative diseases like Alzheimer's and Parkinson's are sex-biased in their incidence and progression. In women, menopause results in a rapid loss of estrogen which coincides with the onset of neurodegeneration. Estrogen is known to have a neuroprotective effect against the emergence of Alzheimer's symptoms, but the molecular mechanisms contributing to this protection are not yet fully understood. Sex steroid hormones such as estrogen bind to nuclear receptor transcription factors which regulate gene expression to produce sex differential biology and disease risk. We sought to identify sex differences in neural aging trajectories and identify cell types that are vulnerable to either aging or the loss of ovarian hormones. We performed snRNAseq in three brain regions with high levels of ERα and PR that are implicated in sex-differential brain function. The BNST is implicated in social and motivated behaviors, while the VMH and arcuate modulate energy balance, food intake, and physical activity. Rodents do not undergo menopause and with aging enter a persistent estrus state with high estradiol and progesterone. To provide a better mouse model of human female aging, we used the ovotoxic chemical 4-vinylcyclohexane diepoxide (VCD) to recapitulate the human menopausal transition in female mice. Additional samples were generated from young adult (p50) and 18-month aged male and female mice. To increase the power for identifying sex DEGs, our lab uses a Esr1-Cre crossed to the Sun1-GFP nuclear reporter line, which permits FANS sorting of ERα+ lineage neurons, followed by snRNA-seq. Our initial results show that our overall strategy of selecting for ERα+ lineages is successful: we detect robust sex differences in gene expression using pseudobulk analysis (DESeq2, adjp<0.1). These sex DEGs vary across cell types. Our data represent the first description of sex differences in gene expression in the arcuate nucleus and in aged mice. We observe distinct subsets of neuronal cell types across brain regions that are strongly affected by the loss of ovarian hormones. In our ongoing analyses we are investigating convergent and divergent changes in gene expression in biologically aged compared to reproductively aged mice, with the long term goal of identifying mechanisms of vulnerability or resilience to brain aging.

THE IMPACT OF HIV-1 ON T CELL CHROMATIN ACCESSIBILITY THROUGH SINGLE-CELL MULTIOMICS

Daniel A Armendariz, Hyein Back, Ya-Chi Ho

Yale University, Microbial Pathogenesis, New Haven, CT

The persistence of HIV-1 in long-lived CD4+ T cell reservoirs remain a major obstacle to a cure. Our previous work has shown that HIV-1 is enriched in specific T cell subsets, including cytotoxic CD4+ T cells in the blood and BACH2-high tissue resident memory (TRM) T cells in the gut. It remains unknown whether this enrichment reflects a preferential infection of pre-existing cells, or a virus-driven reprogramming of the host cell landscape. Previous studies have shown that HIV-1 viral proteins, such as Viral Protein R (Vpr), can drive TRM differentiation and migration from blood to tissue. However, the direct impact of viral proteins on T cell fate decisions at the epigenetic level is not well understood.

We hypothesize that HIV-1 infection, particularly through Vpr, shapes T cell fate decisions by remodeling the chromatin accessibility landscape. To test this, we infected CD4+ T cells from three HIV- donors with one of four lentiviral reporter constructs: a full-length HIV-1 reporter (HIV-GKO), a VPR-deficient mutant (HIV-dVpr), a Vpr-only vector, and an empty vector control. Three days after infection, infected cells versus the uninfected memory cell control were enriched by flow cytometry for single-cell multiomics. Cells were stained with a universal antibody cocktail for characterization of immune cell populations. We then applied the DOGMA-seq assay to profile chromatin accessibility, gene expression, and surface proteins in single cells. Samples were mapped using the Cell Ranger pipeline followed by filtering for high quality cells. We identified 32,491 high quality cells with a mean of 2,643 genes and 2,997 fragments in each cell. Across our conditions, we detected: 5,021 HIV-dVpr cells, 6,654 HIV-GKO cells, 1,438 uninfected cells, 18,724 empty vector control cells, and 654 Vpr-only cells.

Our integrated analysis revealed that HIV-1 Vpr altered the T cell chromatin accessibility landscape coinciding with similar transcriptional changes. We identified increased chromatin accessibility and RNA transcription of the migration and tissue residency genes ITGA1 (integrin α1, CD49a) and ITGA4 (integrin α4, CD49d). Concurrently, HIV-1 Vpr resulted in heightened accessibility and transcription of canonical exhaustion markers including TOX, PDCD1, and TIGIT.

Together these findings provide a single-cell map of HIV-1 Vpr reprogramming of the T cell chromatin accessibility landscape. HIV-1 Vpr may thus bias T cells towards a migration and exhaustion phenotype, contributing to the immune dysfunction of chronic HIV-1 infection.

CHRONOSEQ: MINUTE SCALE AUTOMATED scRNA-SEQ FOR REAL TIME PROFILING OF PERTURBATIONS

Kanishk Asthana[1,2], Alexander Jambor[1,2], Wei Wang[1,3]

[1]University of California San Diego, Chemistry and Biochemistry, La Jolla, CA, [2]University of California San Diego, Bioinformatics and Systems Biology PhD Program, La Jolla, CA, [3]University of California San Diego, Cellular and Molecular Medicine, La Jolla, CA

Current single-cell RNA sequencing technologies achieve temporal resolution of ~30 minutes at best, missing rapid transcriptional changes that occur within minutes of cellular perturbations. We developed **ChronoSeq, a novel droplet-based technology enabling unprecedented 7-minute temporal resolution for single-cell transcriptomics** through innovative time-tagged DNA barcoded beads and automated microfluidic workflows. ChronoSeq works by automatically co-injecting beads with a unique time-tag for each sample of cells taken from a cell culture kept inside the device.

ChronoSeq incorporates time-specific sequence tags into modified Drop-seq beads, coupled with a blockage-resistant microfluidic device supporting 12 unique time-tags across extended 2+ hour experiments. We applied ChronoSeq to profile inflammatory responses in K562 cells stimulated with TNF-α, capturing the rapid NF-κB signaling cascade previously undetectable at conventional temporal scales. Our data revealed distinct activation kinetics among inflammatory genes. Application of Optimal Transport analysis reveal a gradient of activation depending on differentiation towards Erythroid phenotype.

Additionally, we performed mixed-cell perturbation experiments using K562 and differentiated THP-1 cells with LPS stimulation to examine cell-type-specific inflammatory responses and Cell-To-Cell communication dynamics. Moreover, to dissect transcriptional versus translational dynamics, we conducted puromycin translation-blocking experiments following TNF-α stimulation in K562 cells, revealing pure transcriptional kinetics without translational feedback.

ChronoSeq maintains high data quality with <5% doublet rates while dramatically improving temporal resolution. This technology enables precise mapping of perturbation-induced transcriptional cascades, inference of regulatory causality, and characterization of cell-type-specific response dynamics previously inaccessible to single-cell genomics, opening new avenues for understanding cellular perturbation responses.

OPTIMISATION OF MULTI-REGION PROFILING OF BRAIN TISSUE USING EXPLANTED SEEG ELECTRODES

Petra Bachanová[1], Charles J Mazof[2], Barbara C Jobst[3], George J Zannazi[4], Chun-Chieh Paul Lin[4], Jennifer Hong[5]

[1]Dartmouth College, Integrative Neuroscience, Hanover, NH, [2]Dartmouth College, Geisel School of Medicine, Hanover, NH, [3]DHMC, Department of Neurology, Lebanon, NH, [4]DHMC, Department of Pathology, Lebanon, NH, [5]DHMC, Department of Neurosurgery, Lebanon, NH

Background
Craniotomies are a critical source of brain tissue for studying molecular pathology in epilepsy. Yet surgical access to fresh human brain specimens is inherently limited to tissue removed as part of treatment, restricting analyses to pathologic zones without internal controls from normal tissue.
In drug-resistant epilepsy, clinical management increasingly relies on implanted intracranial electrodes to localize seizure-generating regions (Gavvala, 2024). Several groups have recently leveraged these electrodes, following removal, to profile adherent cellular material, revealing spatial gradients in gene expression linked to epileptogenic potential (Dwivedi et al., 2025) and mosaic expression of deleterious gene variants in paediatric epilepsy (D'Gama et al., 2025). Unlike resection, electrode sampling yields material from 10–20 distributed brain regions in the same individual, enabling spatial analysis. However, given its novel nature, the method requires further validation.

Methods
In this ongoing study, we collected cells adherent to explanted electrodes for cytology and RNA extraction. To date, 2 patients (10-12 electrodes each) with sufficient RNA yield (>500 ng total) were selected for bulk sequencing. Cell type composition was inferred with CIBERSORTx using a custom reference matrix we derived from integrated single-nucleus RNA-seq data across 17 surgical epilepsy patients (Buchin et al., 2022; Chen et al., 2023; Kumar et al., 2022; Lerond et al., 2023; Miller et al., 2023). To characterise the neuroimmune compartment, we included deconvolution with a publicly available immune reference. Relative cell type abundance was further validated by cytology in an additional sample using cytospin followed by H&E staining.

Results
Cytology confirmed the presence of neuronal and glial populations, alongside some red blood cell contamination. Bulk transcriptomics deconvolution identified higher abundance of mural cells, interneurons and oligodendrocytes, along with monocytes and NK cells. The higher relative abundance of inflammatory cells may reflect low levels of immune reactivity to electrode implantation.

Conclusion
This method has great potential as a novel approach for profiling fresh brain tissue; however, further optimization and validation are needed to improve robustness and reproducibility. While currently limited to bulk RNA profiling, we are working towards achieving single-cell resolution.

THE FluidFM® TECHNOLOGY FOR TEMPORALLY RESOLVED SINGLE CELL ANALYSIS, GENOME ENGINEERING AND BEYOND

Ludovica Rizzo[1], Kerda Keevend[1], Mira Schaupper[2], Adèle Kerjouan[1], Susanne Mailand[1], Anabel Migenda Herranz1[1], Mantas Survila[2], Torsten Reda[2], Yvonne Goepel[2], Rocco Lafleur[1], Maria Milla[1], Dario Ossola[1], Melanie Horn[1], Tobias A Beyer[1]

[1]Cytosurge, R&D, Glattbrugg, Switzerland, [2]Lexogen, R&D, Vienna, Australia

FluidFM® is an emerging technology that integrates microfluidics with atomic force microscopy (AFM) allowing gentle manipulations of single cells. The technology consists of a force sensitive hollow cantilever with a sharp pyramidal tip (Nanosyringe) combined with a microfluidics pressure control unit. The Nanosyringe can either inject or withdraw liquids (0.2-5pl) into/from cells. This truly unique combination enables novel experimental designs to study a wide range of outstanding questions in biology. Here, we present the latest developments of FluidFM® in the fields of single cell profiling and genome engineering.

Determining the transcriptome of a single cell to characterize it is the cornerstone of countless applications in modern biology (e.g. cancer research, immunology and development). To obtain the transcriptome the cell must be lysed and therefore informs only on a static snapshot of gene expression. This limitation precludes the direct capture of directionality in cell fate transitions and plasticity and relies on correlation and inference. By using FluidFM® to extract cytoplasmic biopsies from cells without killing them, we can transform scRNA-seq from a static endpoint into a dynamic, temporal measurement. This method named Live-seq links temporal gene expression data directly to phenotypic states, enhancing the precision of cellular characterization.

In the context of genome engineering, FluidFM® is used to deliver CRISPR/Cas ribonucleoprotein (RNP) complexes directly into the nucleus of single cells. This approach, coined CellEDIT, provides unprecedented control on the amount and stoichiometry of RNP delivered to individual cells enabling multiplex editing and increases precision. By delivering the RNP directly to the nucleus of single cells, our approach eliminates cell toxicity caused by chemical and physical transfection agents as well as the use of viral vectors. Currently, we are expanding the CellEDIT workflow to co-deliver DNA repair templates to generate point mutations and insertions of large transgenes into distinct genomic loci. We envision that our workflow will enable us to move from genetically engineered to truly genetically designed cell lines.

In summary, we present the FluidFM® OMNIUM, a semiautomated platform that enables manipulation of single cells with unprecedented gentleness. n the future, we envision combining the current applications (Live-seq and CellEDIT) and broadening the range of single cell biopsies to measure the proteome and metabolome of single cells. This will allow researchers to gain a detailed understanding of the systems biology of single cells.

DECONTAMINATING MITOCHONDRIAL MUTATION INFORMATION IN MULTIPLEXED, HIGH-THROUGHPUT SINGLE-CELL DATA

Timothy M Bi[1], Yu-Hsiang Chen[1], Mengchuan Zhuang[1], Christine B Peterson[2], Pavan Bachireddy[1]

[1]University of Texas MD Anderson Cancer Center, Department of Hematopoietic Biology and Malignancy, Houston, TX, [2]University of Texas MD Anderson Cancer Center, Department of Biostatistics, Houston, TX

Recent innovations in single-cell (sc) approaches have enabled the concurrent capture of cell state information and corresponding mitochondrial single nucleotide variants (mtSNVs) at scale, empowering in vivo lineage tracing in humans with applications from deconvoluting donor-recipient origin in stem cell transplant to dissecting leukemic subclonal evolutionary dynamics. However, recent reports have shown that inclusion of low-quality, spurious mtSNVs can influence downstream analyses and data interpretability. We generated paired cell state and mtSNV data via DOGMA-seq for five patients with high-risk myelodysplastic syndromes and one donor across two experiments. Experiment 1 included cells from three patients, while Experiment 2 included cells from two patients and the donor (batched using cell-hashing antibodies). Surprisingly, we observed a nearly 10x higher rate of shared somatic mtSNVs intra-experimentally versus inter-experimentally; moreover, 94.4% of germline mtSNVs in one individual were called as somatic mtSNVs in at least one other individual by mgatk in the same experiment. We reasoned that cell permeabilization could introduce ambient mtDNA from high-frequency variants, resulting in spurious calls of low-frequency somatic mtSNVs. Indeed, analysis of batched MAESTER-seq sc mtDNA data, which does not involve cell permeabilization, showed that germline variants from one individual were 9.1x less prevalent in other individuals as compared to DOGMA-seq data. Based on these observations, we tested and confirmed that a high ratio of mean counts in empty versus cell-containing droplets could distinguish contaminating mtSNVs from true mtSNVs, with improved separation of low-frequency variants following logit transformation of means. Notably, application of the expectation-maximization (EM) algorithm robustly separated putative true and contaminating variants across a wide range of posterior probability thresholds. We are developing MitoClean, a user-friendly analysis package, to facilitate the automated detection and removal of likely contaminating variants. Finally, we validated MitoClean on the four patients with available paired whole genome sequencing data and found an average false negative rate of 1.9%. In conclusion, MitoClean enables the removal of ambient mtDNA contamination introduced by cell permeabilization to enhance reliable and interpretable mtSNV-based analyses from sc data.

AUTOMATED EXTRACTION AND EVALUATION OF CELL-TYPE MARKER GENES

A. Sina Booeshaghi, Nithya Appannagaari, Aaron Streets

UC Berkeley, Department of Bioengineering, Berkeley, CA

Classical cell-type marker genes are abundant, localized, and act as binary indicators. This definition likely arose from the qualitative nature of measuring markers (e.g. fluorescence or morphology), and has proven useful. Classical markers let scientists identify and sort cells, estimate noise, and validate findings. New, quantitative, approaches for measuring gene expression have promised to refine the definition of a marker gene through large-scale data generation. Yet no rigorous definition exists.
Nonetheless a pragmatic definition has emerged: a marker gene is differentially expressed, in a one-vs-all statistical test, filtered by p-value and log-fold change. This approach is ubiquitous but equates differential expression with marker selection; two distinct tasks. And conflating them creates contradictions. For example, UCP1 is a classical marker for brown adipocytes but is not differentially expressed in single-cell data. If a classical marker is not differential, is it still a marker gene? This question highlights the importance of understanding marker selection. Databases like CellMarker (26,915 literature-derived markers) and BGee (41 datasets across 1,849,395 cells) compile extensive lists of markers but don't fully connect expert-identified markers to the data itself making it challenging to evaluate selection strategies.
In this work, we developed mrid, a tool that extracts expert-annotated markers from single-cell studies and links them to the underlying data using large language models (LLM) in order to study marker selection. We curated a database of 1,987 cell-type marker genes extracted from text, figures, and differentially expressed genes (DEG) tables across seven papers, mapped to Uberon and Ensembl identifiers. Using this resource, we found that markers reported in text and figures typically ranked within the top ~60-75 by LFC, indicating that selected markers are strong but not always the strongest. We then evaluated mrid's performance in extracting and linking markers to the underlying data. Accuracy ranged from high precision/recall (0.90/0.58) to near zero depending on how cell types were described in the paper and whether their descriptions were consistent. Encoding the marker-generation workflow improved results only marginally.
Overall, published markers are generally strong, but not consistently top-ranked. Differential expression alone does not fully explain their selection; narrative choices and prior biology likely also play roles. With modest customization, LLMs can extract markers from single-cell studies, enabling automatic curation and evaluation. Taken together, our results demonstrate the utility of LLMs in curating and refining the definition of a marker gene.

SINGLE-CELL STUDY OF THE MOLECULAR LANDSCAPE AT THE TUMOR-BRAIN INTERFACE IN GLIOBLASTOMA

Meghana Budankayala, Liangqi Frank Xie, Jennifer Yu

Cleveland Clinic, Infection Biology, Cleveland, OH

Glioblastoma (GBM) is a highly aggressive and incurable primary brain cancer characterized by profound cellular heterogeneity and rapid progression, with a median patient survival ~15 months. While conventional treatments primarily focus on targeting the tumor mass, nearly 90% of recurrences arise near the surgical resection margin, underscoring the importance of the complex tumor microenvironment (TME) in shaping GBM recurrence. An emerging paradigm highlights that neurons and tumor cells at the tumor-brain interface engage in bidirectional electrochemical communication, shaping neuronal excitability, tumor proliferation and infiltration. However, the precise role of neural activity in regulating GBM cell state and plasticity remains poorly understood, leaving patients at persistent risk for recurrence. To address this knowledge gap, we initiated a first-in-human clinical trial (NCT05565118) that uniquely combines advanced electrocorticography with custom subdural grids and stereotactically placed depth electrodes at and around the tumor margin with spatially matched biopsies, which undergo a spatially resolved single-nucleus RNA sequencing (snRNA-seq) and epigenetic profiling with single-cell ATAC sequencing, enabling a direct correlation between regional electrical activity and tumor-responsive cellular states.
We hypothesize that increased regional electrical activity at the tumor margin fosters a pro-tumorigenic microenvironment that promotes activity-responsive GBM cell state and tumor invasion. Preliminary snRNA-seq analysis of two patients (15,471 cells from five anatomically distinct regions) revealed striking spatial heterogeneity. Cell-cell communication analysis identified extensive bidirectional crosstalk between neurons and malignant astrocytes, whereas robust autocrine signaling was detected in malignant astrocyte populations within the intra tumor region, suggesting self-reinforcing loops that sustain malignant cell proliferation and survival. Strikingly, cell state analysis showed that the PT regions showed a pattern of enrichment of OPC-like states, consistent with invasive features whereas intra-tumoral regions are enriched in MES-like and AC-like states associated with metabolic and inflammatory adaption. These preliminary findings support a spatially organized tumor architecture in which neural–tumor interactions at the periphery reinforce invasive states and recurrence potential. By extending spatial transcriptomic and epigenomic analyses to additional patients, we aim to map activity-dependent cellular states and predict neuron–tumor communication networks underpinning recurrence. These novel insights will inform precision therapeutic strategies to improve GBM outcomes.

IMPACT OF CLONAL HEMATOPOIESIS ON THE CARCINOGENIC PROCESS OF MULTIPLE MYELOMA

Gayeon Cho*[1], Changhee Park*[2], Gangpyo Ryu*[3,4], Jeongmin Park*[4], Kihyun Kim[5], Sung-Yup Cho[4,6,7], Siddhartha Jaiswal[3], Jong Kyoung Kim[1], Youngil Koh[2,4,8]

[1]POSTECH, Department of Life Sciences, Pohang, South Korea, [2]Seoul National University Hospital, Department of Internal Medicine, Seoul, South Korea, [3]Stanford University School of Medicine, Department of Pathology, Stanford, CA, [4]Seoul National University, Cancer Research Institute, Seoul, South Korea, [5]Sungkyunkwan University School of Medicine, Department of Medicine, Seoul, South Korea, [6]Seoul National University College of Medicine, Medical Research Center, Seoul, South Korea, [7]Seoul National University College of Medicine, Department of Biomedical Sciences, Seoul, South Korea, [8]Genome Opinion Inc., Genome Opinion Inc., Seoul, South Korea

Clonal hematopoiesis (CH), an aging-related phenomenon associated with inflammation and hematologic malignancies. Here, we explore the interaction of CH with terminally differentiated lymphoid malignancy, multiple myeloma (MM). Analysis of CH in clinical cohorts revealed a higher prevalence among MM patients and a lower deep response to proteasome inhibitors, especially in TET2-mutated CH carriers. Single-cell multi-omics analysis confirmed that CH mutations predominantly reside in myeloid lineages. By utilizing the bone marrow samples from MM patients with CH, single-cell transcriptome analyses indicated that TET2-mutated CH promotes enhanced hematopoietic stem cell (HSC) self-renewal and myeloid commitment, leading to inflammatory activation of classical monocytes with upregulated cytokine and chemokine production. This inflammatory milieu establishes an IL1B (monocytes) - CCL28 (mesenchymal cells) - CCR10 (plasma cells) paracrine signaling axis that activates MAPK pathways in MM cells, promoting tumor growth. Analysis of the MMRF dataset confirmed that higher expression of IL1B, CCL28, or CCR10 was each associated with inferior overall survival in MM patients. Furthermore, proteomic validation using UK Biobank data confirmed the enrichment of MAPK, STAT, and integrin pathways in MM patients with CH. Our findings reveal that TET2-mutated CH creates a pro-tumorigenic microenvironment through inflammatory paracrine signaling, identifying the CCR10-mediated axis as a potential therapeutic target for CH-associated MM progression.

HIERARCHICAL CLUSTERING OF SINGLE-CELL RNA-SEQUENCING DATA

Urshila Choubal, Sivan Leviyang

Georgetown University, Department of Mathematics and Statistics, Washington, DC

A major task in scRNA-seq data analysis is the identification of cell types through clustering methods. Among other factors, cellular data is inherently hierarchical and high-dimensional, prompting the use of various techniques to address these challenges. Current pipelines use the Louvain algorithm for clustering, where the number of clusters found is a function of a resolution parameter. Since this parameter is usually selected heuristically, the algorithm often leads to under- or over-clustering which can produce artificial clusters that fail to reflect biological reality. Previous work has shown that hierarchical clustering produces more meaningful cell types; however, commonly used analysis pipelines often overlook this structure by clustering in a single pass and then sub-clustering with external intervention. We aim to develop a hierarchical algorithm that uses Louvain to cluster at every level; however, our goal is to rely on a more principled approach to identify the resolution parameter, thereby reducing user intervention in the clustering process.

CROSS-SPECIES MULTIOMICS ANALYSIS REVEALS SHARED AND DIVERGENT AGE-RELATED BIOLOGICAL SIGNATURES

Erica A DePasquale[1,2], David Redmond[3], Jennifer L Burwinkel[1], Kari A Huppert[1], Ashley E Cast[1], Anna L Peters[1,2], Robert E Schwartz[3], Stacey Huppert[1,2]

[1]Cincinnati Children's Hospital Medical Center, Gastroenterology, Hepatology, and Nutrition, Cincinnati, OH, [2]University of Cincinnati, Pediatrics, Cincinnati, OH, [3]Weill Cornell Medicine, Medicine, New York, NY

Animal models have long served as proxies for studying human biology in both health and disease. Yet, our understanding of the extent to which these models faithfully recapitulate human developmental and physiological processes remains incomplete. Differences in anatomy, physiology, and developmental timing can lead to findings that are robust in model organisms but do not translate well to humans. This challenge is particularly important in pediatric research, where aligning the developmental stages of model organisms with human age ranges is critical for accurate interpretation. In the context of liver disease, where mouse models are commonly used, there is a notable gap in large-scale studies that systematically map mouse post-natal developmental stages to human age ranges. Without this vital information, studies targeting childhood and adult liver conditions risk drawing misleading conclusions from animal data.

To address this gap, we performed a comprehensive cross-species multiomic analysis using single-nucleus (snRNA-seq) and chromatin accessibility (scATAC-seq) profiling. Our dataset is comprised of 20 healthy donor human liver samples spanning key developmental stages: infancy (< 2 years), early childhood (2-6), late childhood (8-12), adolescence (13-17), and adulthood (> 18 years). Simultaneously, we profiled 11 control mouse liver samples across post-natal days P3, P15/16, P60/61, and P120/124, representing a spectrum from neonate to mature mouse. Using Harmony integration and expert-guided cell type annotation, we characterized conserved and species-specific patterns in cell type composition, gene expression, chromatin accessibility, inferred developmental trajectory, and cell-cell communication.

This multi-species, multi-modal liver atlas represents a valuable resource for the biomedical community. It provides a foundation that will enable other scientists to identify the developmentally-equivalent model for their particular translational study goals. Beyond its immediate utility in the liver research community, this atlas highlights the broader importance of developmental context in cross-species comparisons and offers a blueprint for similar efforts in other organ systems. To maximize accessibility and impact of this atlas, our future plans include developing a web-based interface for interactive data exploration.

SPATIALLY RESOLVED SINGLE-CELL PROFILING OF THE TUMOR MICROENVIRONMENT IN COLORECTAL LIVER METASTASES.

Trang Dinh[1,2], Weihong Guo[3], Joseph Willis[4], Zhenghe J Wang[1,2]

[1]Case Western Reserve University, Department of Genetics and Genomic Sciences, Cleveland, OH, [2]Case Western Reserve University, Case Comprehensive Cancer Center, Cleveland, OH, [3]Case Western Reserve University, Department of Mathematics, Applied Mathematics, and Statistics, Cleveland, OH, [4]Case Western Reserve University, Department of Pathology, Cleveland, OH

Colorectal cancer (CRC) remains the second leading cause of cancer-related deaths in the United States, with over 150,000 new cases annually. Nearly half of CRC patients develop colorectal liver metastases (CRLM), the predominant site of distant spread and a major contributor to CRC-associated mortality. To dissect the spatial and cellular heterogeneity of the tumor microenvironment (TME) in metastatic progression, we applied high-resolution spatial transcriptomics (10x Genomics Xenium) to tissue microarrays from 18 patients with matched primary CRC and CRLM samples.

Initial low-resolution clustering of tumor transcriptomes revealed pronounced inter-patient heterogeneity, with tumors from different individuals forming distinct clusters. Within each patient, primary and metastatic tumors often shared a common cluster, suggesting a conserved core tumor identity. However, higher-resolution clustering uncovered intra-patient heterogeneity, with most patients exhibiting distinct tumor sub-clusters between their primary and metastatic sites, indicating transcriptional divergence in the metastatic setting. IL33 was consistently upregulated in CRLM tumors in two-thirds of patients, pointing to a potential role in metastatic adaptation.

The global cellular landscape further revealed a higher abundance of myeloid cells in CRLM compared to primary CRC. Sub-clustering of the myeloid population identified significant enrichment of Angiogenic-like Tumor-Associated Macrophages (Angio-TAMs) in liver metastases, suggesting a pro-angiogenic, vessel-promoting phenotype in the metastatic TME. Surprisingly, the fibroblast subsets between primary and metastatic CRCs are markedly distinct, with fibroblasts from CRLM overexpressing genes involved in angiogenesis and vasculature development. Together, these findings support a broader shift in the stromal-immune landscape upon metastasis.

To deeply characterize these changes, we performed stromal-immune niche analysis to examine how the different cell types organize and interact spatially. Consequently, we identified ten unique multicellular neighborhoods, or niches, that were strongly associated with distinct phenotypes of adjacent tumor cells, suggesting coordinated cellular ecosystems that may support tumor progression and vascular remodeling. Specifically, liver metastatic niches were enriched for co-localized Lipid-Associated TAMs, Angio-TAMs, inflammatory TAMs, and specialized fibroblast subsets, with adjacent tumor cells upregulating the TNF-alpha signaling pathway.

Collectively, our findings underscore the power of spatially resolved single-cell transcriptomics to uncover functional architectures and cellular ecosystems within tumors. This work highlights key spatial features of the liver metastatic TME, offering insights into potential therapeutic targets for disrupting pro-metastatic niches.

OPTICALLY RECOGNIZABLE BARCODE BEADS

Boden B Eakins*[1,2], Rodrigo C Chaves*[1,2], Aaron M Streets[1,2,3]

[1]University of California, Berkeley, Bioengineering, Berkeley, CA, [2]UC Berkeley - UCSF Graduate Program of Bioengineering, Berkeley, CA, [3]Chan Zuckerberg Biohub SF, San Francisco, CA

DNA barcoding enables high-throughput sequencing experiments, allowing thousands to millions of single-cells to be assayed in parallel. Barcodes are typically carried by hydrogel barcode beads, which can be efficiently synthesized and manipulated. However, because bead-based barcodes are not known until sequencing, sequencing results cannot be easily linked to other cellular measurements. We have developed Optically Recognizable Barcode Beads (ORBBs) — a way to identify barcode beads using fluorescent imaging — enabling new strategies for single-cell multi-omics.

The encoding space of fluorescent beads has previously been limited to tens or hundreds due to fluorophore spectral overlap. ORBBs increase the fluorescent bandwidth by having two independently stained regions, a core and a shell, allowing for the generation of $(n^m)^2$ unique beads, where m is the number of fluorophores used and n is the number of distinct fluorescent intensity levels. Prior technologies can only generate n^m unique beads. Using four fluorophores and four intensity levels we can encode 65,536 fluorescently unique ORBBs which are identifiable on a standard epifluorescence microscope with 91% accuracy. This brings the fluorescent encoding space into the throughput range of single-cell sequencing.

We have shown that ORBBs can be fluorescently encoded and DNA barcoded simultaneously via split-pool synthesis. This creates an intrinsic link between the fluorescent signature and the oligos on every ORBB, so the DNA barcode can be identified by imaging the bead. This paradigm opens up new directions for single-cell multi-omics. Deterministically placing known DNA sequencing barcodes, a requirement for many single-cell assays, can be replaced with the random placement of ORBBs and imaging to identify each ORBB. We intend to use this technology for high-throughput, single-cell experiments linking sequencing-based measurements of genotype and gene expression to optical measurements of the downstream cellular phenotype.

* These authors contributed equally.

CELLULAR IMAGING WITH INDEXING (CIWI): A NEW PLATFORM FOR IMAGING AND SEQUENCING SINGLE-CELLS

Boden B Eakins[1,2], Aaron M Streets[1,2,3]

[1]University of California, Berkeley, Bioengineering, Berkeley, CA, [2]UC Berkeley - UCSF Graduate Program of Bioengineering, Berkeley, CA, [3]Chan Zuckerberg Biohub SF, San Francisco, CA

High-resolution microscopy is a powerful method for targeted analyses of single cells. Microscopy methods can detect molecules, such as proteins, lipids, or other metabolites, and localize them in a cell with sensitivity and precision. Single-cell genomic analyses complement microscopy, providing deep molecular profiling of the genome or transcriptome without retaining the spatial distribution of molecules. Here we present a new technological platform, Cellular Imaging With Indexing (CIWI), that enables single-cell, high-resolution microscopy to be linked to single-cell sequencing protocols at a throughput of hundreds of cells per experiment, an order of magnitude greater than previously demonstrated.

CIWI connects a microfluidic imaging chip to a well-plate. The chip, fabricated using thin-chip techniques, can be mounted on any optical system compatible with standard microscope slides. Single cells are trapped, isolated, and locked in place in the chip's imaging chamber to allow for robust image acquisition. Cells in the imaging chamber can be sorted and non-target cell types discarded. Post imaging target cells are indexed into a specific well on a standard 96-well plate; plate-based single-cell sequencing assays can then be linked back to the images.

CIWI was designed for applications where imaging a cell requires seconds or longer and is the rate-limiting step for a high-throughput single-cell assay. Indexing and imaging are parallelized, so that the only downtime on the microscope is the time needed to isolate a cell for imaging, which is 2 seconds per cell plus cell sorting times. Here we demonstrate that we can image and index 96 polystyrene beads in 45 minutes with just over half of that being raw imaging time. Over the course of processing hundreds of beads, we observed 100% indexing efficiency – every bead was found in the expected well. We are currently testing cellular indexing with leukocytes and intend to use CIWI to study how lipids and other metabolites are linked to gene expression.

CLUSTERING AND NEAREST NEIGHBOURS FOR scRNA-SEQ DATA USING GENE-WEIGHTED JENSEN-SHANNON DIVERGENCE

Petur H Einarsson, Pall Melsted

University of Iceland, Faculty of Industrial Engineering, Mechanical Engineering and Computer Science, Reykjavik, Iceland

Clustering cells and identifying their nearest neighbours in single-cell RNA-seq data are common tasks for downstream analyses. As a building block, the key aim is to capture biologically similar cells. Current methods often apply a log1p transformation to the count data and select the most highly variable genes, followed by dimensionality reduction via PCA. To cluster the data, Euclidean distances are most often used in the projected space. Dimensionality reduction greatly reduces the size of the data set and speeds up clustering algorithms. However, this approach may lose information and still be subject to the curse of dimensionality.

We propose using a gene-weighted variant of the Jensen-Shannon Divergence for computing squared distances between cells. The square roots of JSD and wJSD are distance metrics and thus satisfy the triangle inequality. We introduce a maximisation-minimisation algorithm for computing the wJSD centroid of a cluster, which does not have a closed-form solution. In addition, we compute the KNN graph using wJSD as the distance metric. Both algorithms assume row-normalised count matrices as input, thus minimising preprocessing requirements. We derive a new majorization-minimisation algorithm to compute the centroids efficiently.

We evaluated our method by running clustering with k=30 on a dataset with 160,000 cells, expression quantified for 20K genes and antibody expression of 228 proteins, with ground truth cell types derived only from the antibody data. Each iteration took 50 seconds when using Lloyd's algorithm, while Hamerly's algorithm reduces this to 30 seconds in later iterations. The program was implemented in Rust and utilises parallelism to speed up the computation. We computed the exact KNN for the same dataset, utilising the triangle inequality to minimise distance computations.

The wJSD captures the similarity of the cells, even under a simple algorithm like K-means. To measure the quality of the clusters, we use a dataset with antibody expression as ground truth and compare the determined cell types with our clusters. Our clustering method separates B cells, NK cells, CD4T and CD8T cells much better than clustering with the Euclidean distance. Leiden clustering on the wJSD-based KNN graph also yielded promising results. Further analysis of the protein data is needed to better evaluate our algorithm.

SINGLE-CELL LINEAGE TRACING REVEALS EXTENSIVE CELL-STATE PLASTICITY IN GLIOBLASTOMA

Yilin Fan*[1,2], Joshua S Schiffman*[3,4], Toshiro Hara[1,2], Dan A Landau[3,4], Mario L Suvà[1,2]

[1]Massachusetts General Hospital and Harvard Medical School, Department of Pathology and Krantz Family Center for Cancer Research, Boston, MA, [2]Broad Institute of Harvard and MIT, Cambridge, MA, [3]New York Genome Center, New York, NY, [4]Weill Cornell Medicine, New York, NY

A major barrier to effective treatment of glioblastomas is inter- and intra-tumoral heterogeneity. Single-cell profiling of human glioblastomas has shown that the heterogeneity of glioblastoma cells converges to a set of four cell states. Yet it is unclear what the extent of plasticity between cell states is and how cell-state plasticity is regulated. Here, we use a combination of single-cell lineage tracing and transcriptional reporters to study the plasticity of glioblastoma cell states. We introduced a CRISPR/Cas9-based lineage tracer into patient-derived glioblastoma cells and performed high-resolution lineage tracing in 3D spheroid cultures and in orthotopic xenografts in mouse brains. We integrated phylogenetic information with single-cell transcriptomic data profiled from the same cell and robustly inferred transition rates between different states using an analytical framework based on phylogenetic correlations. We found extensive bidirectional transitions between stem-like and differentiated-like states in both 3D spheroid cultures and in orthotopic xenografts. Interestingly, the extent of plasticity between cell states was similar to scrambled controls where cell states on phylogenetic trees were randomly scrambled to simulate complete plasticity, suggesting that glioblastoma cell states are highly plastic both in vitro and in vivo. Next, to identify potential regulators of plasticity, we compared gene expression between sub-lineages with different degrees of plasticity and performed a mini-screen of small-molecule inhibitors of epigenetic modifiers. We found that the histone modifications H3K9 methylation and H3K27 methylation are potential regulators of cell state plasticity. The high levels of cell-state plasticity we observed here in glioblastoma highlight the need to therapeutically target multiple states simultaneously or block state plasticity for effective treatment.

* These authors contributed equally

DESYNCHRONIZATION BETWEEN GENE EXPRESSION AND CHROMATIN STATES REVEALS MECHANISMS DRIVING CELL STATE CHANGE

Connor Finkbeiner[1,2], Dominik Otto[1], Manu Setty[1,2]

[1]Fred Hutch, Basic Sciences, Seattle, WA, [2]University of Washington, Genome Sciences, Seattle, WA

While changes in gene expression and chromatin accessibility are coordinated; they are not synchronous. Regulatory interactions take time and can be complex, leading to a delay, or desynchronization between modalities. For example, in enhancer priming changes in chromatin accessibility precede changes in transcription creating desynchronization between gene expression and chromatin accessibility.
Single-cell multiomic assays (paired RNA+ATAC) offer an opportunity to investigate desynchronized cell states; yet no computational methods exist to systematically identify and characterize them. To fill this gap, we have developed a computational framework to quantify desynchronization between the RNA (expression) and ATAC (chromatin) modalities.
Our approach compares cell state density functions across modalities to identify states where cells predominantly changing in a single modality (i.e. are desynchronized.) To interpret the cause of this desynchronization we developed a desynchronization score for RNA and ATAC features: genes, enhancers, motifs, ect, which quantifies whether they are driving cell state change in a single modality in desynchronized states.
Applying our approach to single-cell multiomic datasets from human fetal retina and hematopoiesis, we identified several patterns.
Fate specification is often detected in the chromatin modality before the expression modality and occurs in desynchronized cell states. In the retina dataset cells appear to commit to the cone fate in ATAC before RNA. We found that enhancers for OTX2, which specifies the cone fate, appear to drive the fate specification in ATAC.
Cell cycle associates with changes in gene expression but not chromatin accessibility. Genes associated with these states show limited changes in enhancer accessibility and enrichment for E2F family transcription factors in their promoters, consistent with cell cycle genes being controlled on the protein level.
Loss of chromatin plasticity can cause desynchronization. In the retina immature ganglion cells retained enhancers for progenitor genes (which are no longer expressed at that time) as well as enhancers for cone genes which were no longer expressed. These enhancers were turned off in desynchronized states.
Understanding how cells' states change over time and identifying the key factors driving these changes are fundamental questions in biology. By identifying desynchronization between modalities, our approach uncovers the mechanisms of cell state change, describes the behavior of key regulatory factors, and is broadly applicable across biological systems, from stem cell commitment to disease progression.

CELL-CELL INTERACTION PROFILING WITH SELF-CROSSLINKING SEMI-PERMEABLE CAPSULES

Jase Gehring[1], Xiaohan Zhu[1], Samarth Jajoo[2], Aaron Streets[1]

[1]University of California, Berkeley, Bioengineering, Berkeley, CA,
[2]University of California, Berkeley, Computing, Data Science, and Society, Berkeley, CA

Advances in high throughput perturbation experiments are driving a push to comprehensively map cellular state space. Current methods typically profile the effect of either pooled cell-intrinsic perturbations (e.g. Perturb-seq) or cell-extrinsic perturbations in bulk by manipulating growth conditions. Here we present preliminary work on a new single-cell functional assay that bridges these two paradigms. Using semi-permeable capsules, liquid-filled microparticles with a hydrogel shell, we co-encapsulate diverse stimulating cells along with target cells and use single-cell genomics to read out functional changes brought about in the target cells as a result of the stimulating cell phenotype. Our assay is conceptually similar to Perturb-seq, but rather than reporting on cell-intrinsic genetic effects we can detect changes in cell state as a result of upstream paracrine signaling from the stimulating cell population. This approach is applicable to studying cell-cell interactions (using stimulating cells of various cell types) or assaying the effect of individual secreted proteins (when stimulating cells have been engineered to express protein variant libraries). We present novel chemistries for fabrication of semi-permeable capsules, characterization of capsule performance, and current progress toward integrating microfluidics, cell engineering, and single-cell genomics into this new single-cell functional assay.

BENCHMARKING SKETCHING METHODS ON SPATIAL TRANSCRIPTOMICS DATA

Ian K Gingerich[1,2], Brittany A Goods[2], H. Robert Frost[1]

[1] Dartmouth College, Department of Biomedical Data Science, Hanover, NH, [2]Dartmouth College, Thayer School of Engineering, Hanover, NH

Spatial transcriptomics (ST) platforms profile hundreds of thousands of locations per section across thousands of genes, creating computational bottlenecks. Sketching selects small representative subsets that preserve structure while reducing runtime and memory. In scRNA-seq, common sketching uses low-dimensional embeddings (e.g., PCA) and prioritizes underrepresented regions: leverage-score sampling, Geosketch, and scSampler, which outperform uniform sampling that over-represents dense regions. But transcriptome-only sketching can distort spatial structure in ST: if high-variability profiles cluster, expression-only sketches over-sample those areas and under-sample others, inflating co-localization, suppressing interactions elsewhere, perturbing gradients and boundaries, and fragmenting or merging domains. We benchmark leverage scores, Geosketch, and scSampler on four ST datasets and matched simulations. Each is applied to three inputs: PCA embeddings, raw coordinates, and spatially smoothed embeddings using a spatial weights matrix; uniform sampling is the baseline. For leverage scores we implement Seurat's approach and a spatially smoothed variant: compute rSVD to obtain a low-rank basis U, multiply by spatial weights W to get U~, then compute leverage from U~ (row-wise squared norms). Sampling without replacement from these probabilities injects local spatial context while retaining transcriptomic signal. Evaluation probes transcriptomic fidelity and spatial integrity: robust Hausdorff distances in expression and coordinate spaces, adjusted Rand index for clustering stability, PCA loading drift, and local cell-type composition MSE; we also examine regional over-/under-representation and edge effects.
Findings:
• Expression-only geometric methods capture global heterogeneity and rare states but over-sample high-variability regions and under-sample homogeneous areas, biasing neighborhood analyses.
• Coordinate-only sampling yields near-uniform coverage but misses transcriptomic extremes and can over-select periphery.
• Leverage score sampling balances both; spatially smoothed leverage further reduces spatial bias, preserves transcriptomic coverage, avoids edge over-selection, and better matches local compositions.
Conclusion: spatially smoothed leverage often matches or outperforms expression-only and coordinate-only sketches, offering a practical, model-agnostic way to couple expression heterogeneity with spatial coverage. For regular dense grids uniform/coordinate sketches may suffice, but heterogeneous in situ assays benefit from spatially informed sketching to enable fast, unbiased analyses of large ST datasets.

BENCHMARKING SPATIAL CELL SEGMENTATION METHODS OF HIV-INFECTED TISSUES

Qijie Guan, Amare Eshetu, Ya-Chi Ho

Yale University, Department of Microbial Pathogenesis, New Haven, CT

Single-cell resolution spatial transcriptomics promises to unravel complex tissue architecture, but its accuracy hinges on the critical computational step of cell segmentation. In dense and complex tissues like human lymph nodes, standard automated algorithms often fail. The Xenium default segmentation, for instance, frequently merges multiple cells into single polygons, while other advanced tools like Segger can overlook entire clusters of densely packed lymphocytes. These inaccuracies compromise the integrity of downstream analyses.

Here, we introduce and validate a semi-automated workflow that employs manual refinement to correct and augment the output from the Segger algorithm. By meticulously curating segmentation boundaries, our method accurately delineates individual cells and captures lymphocytes previously missed in cell-rich areas. We validated this approach across multiple tissue sections from three distinct human lymph node cohorts, incorporating data from both the 480-gene Immuno-Oncology panel and 5K-gene Prime panel to ensure robustness.

Our refined method demonstrably improves segmentation fidelity. While Segger reduced the B and T lymphocyte doublet proportion from the Xenium default (34.21±6.84%) to 29.50±7.93%, our manual refinement further lowered this rate to 23.53±5.40%. Interestingly, despite substantial differences in cell counts, cell area, and transcripts per cell, the overall B and T cell percentages remained deceptively stable across all three methods. This critical finding suggests that high-level cell typing can mask severe underlying segmentation errors. Furthermore, by adding previously omitted cells, our method enhances the precision of spatial calculations, such as transcript to nearest HIV+ cell distance and cell-type annotation.

In conclusion, our manual refinement strategy provides a robust solution to a key bottleneck in spatial transcriptomics analysis. It is an essential step for generating reliable single-cell data in dense tissues, enabling more accurate investigations of cellular composition, gene expression, and spatial interactions within complex microenvironments.

DRIVING SINGLE-CELL-BASED TARGET ID AND VALIDATION VIA A COMPREHENSIVE PBMC IMMUNE scRNA-SEQ ATLAS

Jeremy P Gygi, Joshua Whitener, Madison Stulir, Zongmei Gao, Aridaman Pandit, Zhaleh Safikhani

AbbVie, Cambridge Research Center, Research & Development, Cambridge, MA

Single-cell RNA sequencing (scRNA-seq) has transformed the ability to resolve cellular heterogeneity and uncover disease mechanisms that bulk sequencing often obscures. Notably, some disease-specific signatures are subtle or restricted to rare and understudied cell type populations, such as plasmablasts and plasmacytoid dendritic cells. As such, these signatures may only appear in a subset of studies, underscoring the necessity for integration. However, integrative analysis is often challenged by heterogenous sampling, preservation, and sequencing methods as well as the scarcity of robust preclinical models.

To address these challenges, we present a large-scale integrated single-cell atlas: the PBMC Immune scATLAS comprised of over 2 million carefully curated PBMCs across 399 unique diseased (N = 169, systemic lupus erythematosus – SLE; ulcerative colitis - UC) and healthy donors (N = 230). We used the PBMCs to train a neural network comprised of a variational autoencoder (VAE) and generative adversarial network (GAN) to denoise and harmonize data across studies, platforms, and experimental protocols. This harmonized framework effectively removes undesired batch effects (e.g. technical and experimental noise) while preserving biological signals, enabling cross-dataset integration of immune cell states, cell-cell interactions, responses to perturbations, as well as benchmarking and prioritization of pre-clinical assets.

We applied our PBMC Immune scATLAS to an AbbVie-proprietary internal SLE CITE-seq cohort. The atlas successfully integrated the transcriptional profiles and annotated the cells at both major and minor cell-type resolutions, validated via jointly profiled ADT cell-type-specific markers. Additionally, we identified cell population shifts across both disease status and ethnicity, including shifts in naïve B cells and central memory CD4+ T cells.

To build the atlas, we thoroughly benchmarked a range of state-of-the-art integration methods, including traditional iterative and anchor-based approaches (Harmony, Scanorama), deep generative VAEs (scVI, trVAE, SCALEX), VAE-GAN combinations (iMAP, scCRAFT), and foundational models (scGPT). We assessed the performance of the methods in both preserving biological signal (e.g. cell types, disease-specific populations) and removing batch effects (e.g. study, donor, sequencing technology) across multiple resolutions as well as both supervised and unsupervised gene selection strategies. Specifically, deep learning methods (VAE and VAE-GAN combinations) performed the best across all benchmarking tests with foundational models (scGPT) performing quite poorly, motivating the need to train custom scRNA-seq atlases. Finally, to measure the amount of overfitting, we evaluated the projection of unseen scRNA-seq datasets without retraining, an essential feature for building living reference atlases.

All Authors were or are employees of AbbVie at the time of the study. The design, study conduct, and financial support for this research were provided by AbbVie. AbbVie participated in the interpretation of data, review, and approval of the publication. No honoraria or payments were made for authorship.

STIMULATED RAMAN SCATTERING MICROSCOPY AS AN EFFICIENT TOOL FOR QUANTITATIVE SINGLE-CELL CHARACTERIZATION

Wuji Han[1], Bryce Manifold[2], Bowen Yang[3], Chiara Ceconello[4], Denis Titov[1,3,5], Dario Polli[4], Aaron Streets[1,2,6]

[1]UC Berkeley, Center for Computational Biology, Berkeley, CA, [2]UC Berkeley, Bioengineering, Berkeley, CA, [3]UC Berkeley, Nutritional Sciences and Toxicology, Berkeley, CA, [4]Politecnico di Milano, Physics, Milan, Italy, [5]UC Berkeley, Molecular and Cell Biology, Berkeley, CA, [6]UC Berkeley, Biophysics Graduate Group, Berkeley, CA

Methods for single-cell genomic analysis typically destroy the sample prohibiting for further analyses. This limitation of destructive genomics techniques poses a challenge for multimodal measurements of single-cells. Microscopy has been a well established complement to various biochemical assays because of its high spatial resolution and minimal impact of the sample. Stimulated Raman scattering (SRS) is a microscopy technique that provides a fast and non-perturbative way to visualize a cell's metabolic landscape by probing intrinsic molecular vibrational modes. This enables measurements including lipid and small signaling molecules in vivo, which is challenging for other optical techniques.

Over the past decade, various methods have been developed to profile single cells with stimulated Raman scattering effect. Most of which focus on either sacrificing spatial resolution by doing spectroscopy, or throughput by taking hyperspectral images throughout the fingerprint region. To address these key parameters for single-cell measurements, we explored the potential of two-color SRS imaging for single-cell profiling. With spectral focusing and engineering optimizations, we achieved full tissue high-resolution imaging in less than a day. In addition, we demonstrated computational analyses on the SRS images, such as segmentation, spectral component extraction and single-cell lipid statistics.

Using this detection and analysis pipeline, we applied our method onto multiple systems. Firstly, we studied the effect of multiple common fixatives on over two thousand cultured HeLa cells. Second, we analyzed protein/lipid accumulation of over four hundred whole *C. elegans* organisms with different genotypes, ages, and caloric restriction conditions. Both examples showed the speed and quantification capability of the method. Lastly, we demonstrated the compatibility of two-color SRS with other measurement modalities. Through paired spatial transcriptomic and imaging measurements of mouse tissue slices, we revealed the relation between gene expression and metabolite compositions by analyzing over a hundred thousand cells *in situ*.

A SPATIAL MULTIOMICS CHARACTERIZATION OF THE HETEROGENEITY AND ORGANIZATION OF CELLULAR SENESCENCE IN AGING SKIN

Yeaeun Han[1,2], Dominik Saul[1,3,4], Ana Catarina Franco[1,2], Anthony Lagnado[1,2], Joao Passos[1,2]

[1]Mayo Clinic, Robert and Arlene Kogod Center on Aging, Rochester, MN, [2]Mayo Clinic, Department of Physiology and Biomedical Engineering, Rochester, MN, [3]Eberhard Karls University, Department of Trauma and Reconstructive Surgery, Tübingen, Germany, [4]Bosch Health Campus, Robert Bosch Center for Tumor Diseases, Stuttgart, Germany

Cellular senescence refers to a stable state of cellular arrest, characterized by a senescence-associated secretory phenotype (SASP). Senescent cells accumulate with age in multiple tissues and contribute to tissue dysfunction and age-related diseases, but their identification and characterization in vivo remain a challenge due to their rarity and complexity. Although p16 (*CDKN2A*) and p21 (*CDKN1A*) are used widely as canonical markers of senescence, their expression and function can vary significantly across different tissues and cell types.

To address this complexity, we conducted an extensive analysis of human skin from sun-protected sites across a wide range of ages using a suite of spatially-resolved platforms, including iterative indirect immunofluorescence imaging (4i) proteomics, Akoya Phenocycler (CODEX) , and Xenium transcriptomics. Our analysis revealed that senescent cells are not randomly distributed but aggregate into distinct clusters that increase in both number and size with age.
Within these clusters, we found that p16- and p21-expressing senescent cells present different "senotypes" with divergent localizations ; p16 was enriched in basal melanocytes while p21 accumulated in corneocytes. Each marker was also associated with a distinct, cell type-agnostic SASP profile at both the transcriptomic and proteomic levels. These senescent cell clusters were frequently found adjacent to immune cell infiltrates, and we observed an upregulation of senescence markers within T-cell populations, suggesting active immune crosstalk and potential immune cell senescence. Furthermore, p21 expression correlated with oncogenic transcript signatures in an age-dependent manner.

Our integrated multiomics approach demonstrates the diverse yet ordered spatial mapping of cellular senescence, providing a critical framework of understanding the functional consequences of senescent cells in aging at both the cellular and possibly the tissue level.

GFRAL-EXPRESSING CELLS IN THE DVC RESPONSE TO CIRCULATING GDF15

Ceci Hes[1,2], Abigail Tomlinson[3], Hunter Murdoch[1], Lieke Michielsen[4], Maia Kokoeva[1], Paul Sabatini[1,2]

[1]RI, MUHC, Montreal, Canada, [2]Clinical and Translational Research, McGill U., Montreal, Canada, [3]Dept. Internal Medicine, U. of Michigan, Michigan, MI, [4]LUMC, Leiden University, Leiden, Netherlands

The dorsal vagal complex (DVC) is located in the brainstem and is relevant for food intake regulation in health and disease. In cancer, GDF15 released by tumors reaches the DVC through the circulation and signals through its receptor, GFRAL. This signaling pathway mediates the nausea and anorexia experienced by people undergoing cancer cachexia and other conditions. In cancer GDF15 released by tumors reaches the DVC through the circulation. The resulting aversive behavior is responsible for the involuntary the appetite and weight loss in those patients and compromise their outcomes. The DVC cell identities that express *Gfral* and how they respond to GDF15 in acute and chronic exposures are unknown.
Aim: To establish the cellular activity changes mediated by GDF15 in all *Gfral*-expressing cells in the DVC.
Results: To define all DVC cell identities, we generated the first multi-species DVC cell atlas. Here, we isolated DVC nuclei for sn-RNAseq from adult mice (n=30) and rats (n=2) and established 123 cell identities at 5 granularities of cellular resolution from >180,000 high quality cells analyzed in R and python. This atlas was made available to other researchers in multiple formats as a computational toolbox. The main cell identities found to express *Gfral* were monoaminergic and *Vlugt1* neurons. Next, we injected GDF15 or vehicle to GFRAL:GFP mice (GFRAL-expressing neurons are marked with green fluorescent protein), which allowed us to define GDF15-mediated transcriptional changes in *Gfral*-labelled cells. In this experiment, cell sorting allowed isolation of GFP+ (*Gfral*-expressing) cells from DVC tissue and after performing snRNA-seq, we successfully labeled 48,147 cells using our atlas. We confirmed that some glial cells (e.g. oligodendrocytes and astrocytes) also express *Gfral* and that the main responders to GDF15 are M0 monoaminergic neurons. We observed dramatic changes in transcriptomic profiles of M0 neurons after 4 days of exposure to GDF15 compared to one hour, including collagen and keratin associated genes and phospholipid transfer (i.e. *Pltp*). In opposition, changes in Vglut1-neurons were mild and do not involve these genes. We describe for the first time keratin associated genes expressed by *Gfral*-expressing neurons after a stressor (injection) is applied to mice.
Conclusion: Neuronal and glial cells express *Gfral* in the AP/NTS. The monoaminergic neurons are the main responders to GDF15 signaling in the AP/NTS and their chronic activation involves keratin and collagen associated genes.

INDUCTION OF MENSTRUATION IN MICE REVEALS MECHANISMS UNDERLYING ENDOMETRIAL COMPARTMENTALIZATION

Çağrı Çevrim[1], Nicholas Hilgert[2], Sean R Eddy[2,3], Kara L McKinley[2,3]

[1]Harvard University, Department of Stem Cell and Regenerative Biology, Cambridge, MA, [2]Harvard University, Department of Molecular and Cellular Biology, Cambridge, MA, [3]Harvard University, Howard Hughes Medical Institute, Cambridge, MA

Menstruation occurs in fewer than 2% of mammalian species, limiting mechanistic studies of this fundamental reproductive process. We developed X-Mens, a transgenic mouse model that enables menstruation to be induced physiologically and repeatably. X-Mens mice recapitulate hallmark features of human menstruation, including cyclical decidualization, shedding with vaginal bleeding, and complete functional regeneration, permitting successive pregnancies. Cross-species transcriptomic analyses revealed conserved genetic programs underlying decidualization and menstrual shedding. Using spatially resolved single-nucleus RNA-seq data of the X-Mens and human endometrium, we found that menstruation initiates through spatially patterned waves of decidual cell differentiation, generating a maturation gradient that resolves into distinct tissue layers demarcating the boundary of shedding. The X-Mens model provides a powerful platform for studying endometrial regeneration and offers translational opportunities for therapeutic discovery in menstrual disorders affecting billions worldwide.

DECODING CELL-CELL COMMUNICATION DURING DIFFERENTIATION WITH SINGLE-CELL GENOMICS

Yin Huang[1,2,3], Manu Setty[1,2]

[1]Fred Hutchinson Cancer Center, Computational Biology Division, Seattle, WA, [2]Fred Hutchinson Cancer Center, Basic Science Division, Seattle, WA, [3]University of Washington, Molecular and Cellular Biology Program, Seattle, WA

Cells use a range of mechanisms to communicate, but a major class of signaling involves cell-surface receptors that detect extracellular cues and initiate intracellular signaling cascades. These signals converge on transcription factors (TFs), which serve as key executors by regulating expression of target genes through gene regulatory networks (GRNs). As these relationships vary across cell states, the same receptor input can produce different outcomes depending on the regulatory context of the receiver cell. While single-cell technologies allow us to study signal pathways at unprecedent resolution, most computational methods rely on a single, static GRN to approximate cellular context, which assumes that all cells interpret signals through the same regulatory program. Current methods also tend to infer cellular interactions between discrete cell groups, overlooking the fact that transcriptional responses often change continuously across dynamic processes like differentiation. Capturing these rapid transitions requires higher-resolution approaches that go beyond discretized comparisons.

To address these limitations, we have developed **a novel computational algorithm that quantifies receptor activity based on its transcriptional consequences at single-cell resolution**. Using single-cell multiome data, which jointly profiles gene expression and chromatin accessibility in the same cell, our framework identifies which TFs are activated by each receptor and quantify the changes of their downstream genes. By computing receptor activity at the level of single cell, **this approach enables modeling of cell responses that are both context-specific and continuous**.

Applying our framework to human hematopoiesis, **we recovered known signaling patterns**, including the stage-specific activation of IL7R in B cell development and EPOR in erythropoiesis. Notably, while IL7R expression peaks in pro-B cells, our analysis indicates that its functional activity is delayed, reaching a maximum at the transition from pro-B to pre-B cells. This highlights a measurable discordance between receptor expression and functional output. Similarly, although EPOR is broadly expressed across the erythroid lineage, its signaling activity is restricted to early progenitor stages, indicating a confined window of engagement. **By resolving the dynamics of receptor activity across cell-state transitions in characterized pathways, our approach is now well positioned to uncover novel regulators and interactions in development and disease.**

SINGLE-CELL MULTIOMICS IDENTIFIES MUTATION-DRIVEN CANCER CELL SUBTYPES INVOLVED IN TUMOR MICROENVIRONMENT–MEDIATED PROGRESSION

Tadashi Imafuku[1], Takahiro Sasaki[2], Kyohei Matsumoto[1], Naoyuki Nakao[2], Shinichi Hashimoto[1]

[1]Wakayama Medical University, Department of Molecular Pathophysiology, Wakayama, Japan, [2]Wakayama Medical University, Department of Neurological Surgery, Wakayama, Japan

Recent advancements in the analysis of genome mutations in single cells have revealed the presence of cancer cell clones with diverse mutation patterns in the tumor microenvironment (TME). However, the relationship between these diverse mutation patterns and their role in cancer progression is unclear. To address this issue, we have developed a micro-well device that can perform single-cell genomic mutation analysis using microbeads embedded with oligonucleotide sequences that target driver mutations from a cancer gene panel. These microbeads also enable simultaneous gene expression analysis, allowing for comprehensive single-cell multiomics profiling. In this study, we demonstrated that our single-cell multi-omics analysis enables evaluation of mutation-driven cancer cell subtypes associated with cancer progression in the TME. First, we established a protocol for single-cell multiomics analysis using pancreatic cancer cell lines. As a result, we successfully classified cancer cells based on distinct KRAS G12 mutation patterns, along with their corresponding gene expression profiles. Furthermore, we applied single-cell multiomics analysis to surgical brain tissue specimens from glioblastoma patients. As a result, we identified three distinct cancer cell subtypes: Cancer_0 (S100A9^{high} / ALOX5^{high}, immunosuppressive type), Cancer_1 (SOX4^{high} / CD24^{high}, stem-like type), and Cancer_2 (IGFBP2^{high} / TGM2^{high}, invasive type). In addition, Cancer2 harbored the CDH1 F338Sfs*18 mutation, suggesting that loss of CDH1 function may enhance the invasive capacity of these cells. Cell–cell interaction analysis showed that Cancer2 received strong signals from fibroblasts, implying that fibroblast–cancer cell interactions may promote its invasiveness. Our technology contributes to the understanding of the role of genomic mutations in cancer cells and the identification of novel targets for prediction, diagnosis, and treatment of cancer progression.

MACROPHAGE TM4SF19 CONTRIBUTES TO OBESITY-INDUCED ADIPOSE TISSUE INFLAMMATION

Yujin L Jeong*[1], Cheoljun Choi*[2], Koung-Min Park*[3], Young Suk Jung[4], James G Granneman[5], Young-Min Hyun[3], Jong Kyoung Kim[1], Yun-Hee Lee[2]

[1]Pohang University of Science and Technology (POSTECH), Department of Life Sciences, Pohang, South Korea, [2]Seoul National University, College of Pharmacy and Research Institute of Pharmaceutical Sciences, Seoul, South Korea, [3]Yonsei University College of Medicine, Department of Anatomy and Brain Korea 21 PLUS Project for Medical Science, Seoul, South Korea, [4]Pusan National University, College of Pharmacy, Busan, South Korea, [5]Wayne State University, Center for Molecular Medicine and Genetics, Detroit, MI

Adipose tissue macrophages are a critical player in obesity-induced inflammation and metabolic dysfunction. Among these, TREM2+ lipid-associated macrophages have emerged as key players in diseases characterized by lipid stress and inflammation, including obesity. Here, we identify TM4SF19 as a lysosomal membrane protein that is selectively expressed in TREM2+ macrophages and represses lysosomal acidification by interacting with V-ATPase subunit. In vitro co-culture indicates TM4SF19 knockout accelerates the clearance of dying/dead adipocytes. Single-nucleus RNA sequencing analysis revealed TM4SF19 knockout reduces high-fat diet-induced accumulation of TREM2+ macrophages and increases the proportion of restorative LYVE1+ macrophages. Importantly, TM4SF19 knockout mice adapt to high-fat feeding through adipocyte hyperplasia, rather than hypertrophy. This adaptation significantly improves local and systemic insulin sensitivity, offering TM4SF19-mediated control of lysosomal activity as a potential therapeutic target to resolve obesity-induced adipose tissue inflammation.

A RANKING SYSTEM FOR CELL-CELL INTERACTIONS AND ITS APPLICATION TO EPITHELIAL BARRIER TISSUES TO INVESTIGATE HEALTHY, PRE-MENOPAUSAL BIOLOGY

Olha Kholod[1], Hien M Bui[1], Hildreth R Frost[2], Brittany A Goods[1,3,4]

[1]Dartmouth College, Thayer School of Engineering, Hanover, NH, [2]Dartmouth College, Department of Biomedical Data Science, Hanover, NH, [3]Dartmouth College, Department of Molecular and Systems Biology, Hanover, NH, [4]Dartmouth College, Program in Quantitative Biomedical Sciences, Hanover, NH

Cell-cell interactions (CCIs) are fundamental to tissue homeostasis, immune regulation, and barrier defense. Advances in single-cell RNA-sequencing (scRNA-seq) have transformed our ability to study CCIs in situ, enabling inference of interactions across complex multicellular environments. However, there is no robust framework for comparing CCI patterns across multiple organs and cellular compartments. This gap is particularly critical for heterogeneous tissues, where coordinated signaling between epithelial, immune, and stromal compartments underpins tissue function.

To address this, we introduce the Composite Interaction Score (CIS), a ranking metric designed to prioritize CCIs from scRNA-seq datasets. CIS is based on ranked-biased precision: it emphasizes top-ranked interactions inferred by multiple methods implemented in the LIANA package and calculates a weighted composite score that reflects both cross-tool agreement and rank position. We validated CIS on benchmark datasets used for prior CCI tool development, demonstrating its ability to highlight biologically meaningful interactions.

We then applied CIS to scRNA-seq data from epithelial barrier tissues as a representative use case. Across intestine, skin, and uterus, MIF–CD74 and APP–CD74 consistently ranked among the top interactions in epithelial, immune, and stromal compartments. Notably, APP–CD74 signaling, previously characterized in non-mucosal systems, was enriched in communication between epithelial, endothelial, and fibroblast populations with myeloid cells in the uterus, suggesting a novel, tissue-specific immunoregulatory role.

Next, we investigated epithelial-compartment CCIs unique to each tissue. In the intestine, GUCA2A/GUCA2B–GUCY2C emerged as a mediator of epithelial-endocrine crosstalk, extending its role beyond homeostatic regulation. In the skin, HLA-A/HLA-F–KIR3DL1 ranked highly between keratinocytes and NK cells, with our analysis predicting that keratinocyte-derived HLA-A and inducible HLA-F may engage KIR3DL1 on tissue-resident NK cells to constrain cytotoxicity. This axis likely represents an epithelial-driven mechanism of immune homeostasis at the skin barrier. In the uterine epithelium, we identified several high-ranking interactions between epithelial-ciliated cells and myeloid populations, including SPP1–PTGER4, GAL–HLA-DPA1, and ANXA1–FPR1. Osteopontin (SPP1) signaling through PTGER4 (EP4 receptor) is known to promote epithelial restitution and anti-inflammatory macrophage phenotypes in the reproductive tract and colon, suggesting a role in resolving mucosal inflammation and maintaining uterine homeostasis.

Together, our work establishes CIS as a generalizable framework for prioritizing CCIs from single-cell data, demonstrates its utility in uncovering both conserved and tissue-specific communication networks, and provides a comparative reference resource across epithelial barrier tissues.

INTEGRATIVE SINGLE-CELL MULTIOME AND LONG-READ TRANSCRIPTOMIC PROFILING OF ZEBRAFISH GONADAL DEVELOPMENT

Jisoo Kim

Gwangju Institute of Science and Technology, Department of Life Sciences, Gwangju, South Korea

Embryonic gonadal development is a tightly regulated process beginning with a bipotential gonad that can differentiate into testis or ovary. This transition requires repression of the alternative fate through transcription factor networks, and disruption can cause disorders of sex development. Although non-coding RNAs and epigenetic factors have been implicated, the regulatory framework underlying gonadal differentiation remains incomplete. Zebrafish serve as a powerful model due to conserved cell types and genes with mammals and the ease of obtaining large numbers of staged samples.

To investigate this, we profiled four stages—20 dpf undifferentiated gonad, 30 dpf gonad, adult testis, and adult ovary—using single-cell multi-omics (RNA + ATAC) with PacBio Iso-Seq long-read RNA-seq. This integrative design allows simultaneous analysis of transcriptional programs, chromatin accessibility, and isoform diversity.

Multiome data enabled reconstruction of developmental trajectories. Pseudotime analysis suggested lineage-specific transitions, with some germ and somatic populations showing possible sex-biased branching. Integration of RNA and ATAC profiles supported inference of gene regulatory networks (GRNs), revealing stage-dependent transcription factor activity aligned with developmental progression and pointing to coordination between chromatin remodeling and transcriptional control.

Long-read analysis expanded the view of transcriptomic complexity. Full-length isoforms revealed cell type–specific differentially expressed genes (DEGs), differentially used transcripts (DTUs), and nonDEG-DTU events where gene expression remained stable but isoform usage shifted. Such events highlight regulatory layers often missed by short-read approaches, suggesting that isoform switching may fine-tune gonadal differentiation. These isoform-level changes and novel transcript variants offer insights into regulatory processes in sex determination.

In summary, this study presents the first integrative multiome–long-read resource for zebrafish gonadal development. By combining transcriptional programs, chromatin states, and isoform diversity in a single framework, the dataset enables exploration of stage- and sex-specific regulation at high resolution. Beyond serving as a reference map, it points to developmental transition points, candidate regulators, and isoform-level events for further study. As a community resource, it supports hypothesis generation, mechanistic exploration, and comparative analyses across vertebrates, advancing our understanding of sex determination and reproductive biology.

UNRAVELING THE ROLE OF α-1,2-MANNOSIDASE MAN1C1 IN GLIOMA PROGRESSION AND IMMUNOMODULATION

Don Carlo R Batara, Minseo Kim, Chan-Ho Lee, Sung-Hak Kim

Animal Molecular Biochemistry Laboratory, Animal Science, Gwanju, South Korea

Glioblastoma multiforme (GBM) is the most aggressive form of primary brain tumor, with glioma stem cells (GSCs) contributing to its therapeutic resistance and recurrence. Aberrant glycosylation has been increasingly recognized as a driver of cancer aggressiveness; however, the role and mechanisms of N-glycosylation in regulating the GSC phenotype and GBM malignancy remain poorly understood. In this study, we applied an integrated transcriptomic analysis of publicly available datasets, complemented by qRT-PCR validation, to investigate the impact of N-glycosylation–related genes in GSCs and gliomas. Our results revealed that elevated expression of α-1,2-mannosidase MAN1C1 is associated with immune-related processes and poorer survival in glioma patients. Differential gene expression analyses and qRT-PCR confirmed that MAN1C1 is significantly upregulated in GSCs. Moreover, high MAN1C1 expression correlated with adverse clinical outcomes, enrichment in the perinecrotic regions of GBM, and functional links to immune and inflammatory pathways—hallmarks of the mesenchymal GBM subtype. Additional analyses demonstrated a strong association between MAN1C1 expression, immune cell infiltration, and impaired immune responses within the GBM microenvironment. Collectively, these findings highlight MAN1C1 as a potential biomarker for gliomas and a promising candidate for immunotherapeutic targeting in GBM.

SINGLE-CELL PROFILING OF THE HUMAN CHOROID PLEXUS IDENTIFIES EPITHELIAL SUBTYPES ASSOCIATED WITH ALZHEIMER'S DISEASE

Seon S Kinrot[1,2,3], Milos Pjanic[1,2,3,4], Xuan Cao[1,2,3,4], Christian Porras[1,2,3,4], Xinyi Wang[1,2,3,4], Prashant N.M.[1,2,3,4], Sarah Murphy[1,2,3,4], Clara Casey[1,2,3,4], Zhiping Shao[1,2,3,4], Aram Hong[1,2,3,4], Chenyu Liu[1,2,3,4], Georgios Voloudakis[1,2,3,4], Vahram Haroutanian[2,3,4], Donghoon Lee[1,2,3,4], Jaroslav Bendl[1,2,3,4], Panagiotis Roussos[1,2,3,4]

[1]Icahn School of Medicine at Mount Sinai, Center for Disease Neurogenomics, New York, NY, [2]Icahn School of Medicine at Mount Sinai, Friedman Brain Institute, New York, NY, [3]Icahn School of Medicine at Mount Sinai, Department of Psychiatry, New York, NY, [4]Icahn School of Medicine at Mount Sinai, Department of Genetics and Genomic Sciences, New York, NY, [5]James J. Peters VA Medical Center, Center for Precision Medicine and Translational Therapeutics, Bronx, NY, [6]James J. Peters VA Medical Center, Mental Illness Research, Education and Clinical Center VISN2, Bronx, NY, [7]Icahn School of Medicine at Mount Sinai, Department of Artificial Intelligence and Human Health, New York, NY

The choroid plexus (ChP) is a vascularized extension of one of the innermost brain membranes - the pia mater - located in each of the four ventricles of the brain. The ChP contains highly specialized cellular structures, controlling the flow of material between the brain parenchyma, cerebro-spinal fluid (CSF) and circulating blood. However, a detailed taxonomy of cell types in the ChP is lacking. To fill this gap, we performed single nucleus RNA sequencing (snRNA-seq) on ~1M cells derived from 160 human ChP specimens from the lateral ventricle. Our cohort is composed of donors with 3 neurodegenerative disorders: Alzheimer's Disease (AD), diffuse Lewy body disease (DLBD), and Parkinson's disease (PD), as well as cardiovascular disease (CVD) and controls. We performed iterative taxonomy analysis and defined cell types in 3 hierarchical layers: 18 major cell type classes, 36 subclasses and 64 subtypes. Cell type abundance analysis identified cell types whose proportions significantly change with AD status, including a subset of epithelial cells that are more abundant in disease. These cells also displayed increased DNA damage signatures and evidence of compromised gap junction formation. To investigate these findings further, we used the Xenium in situ platform to collect spatial transcriptomics data from 8 donors (4 AD and 4 controls). Using a label transfer protocol, this data allowed us to characterize the detailed cellular structure of the ChP. Differential abundance analysis on the spatial data revealed a striking difference in the prevalence of vulnerable epithelial cell types between AD and controls, further highlighting the potential role of damaged ChP epithelium in the etiology of AD.

THE STUDY OF MOLECULAR SIGNATURES OF COORDINATED CELLULAR RESPONSES DURING SKELETAL MUSCLE REGENERATION USING SPATIAL TRANSCRIPTOMIC ANALYSIS

Kyung Dae Ko[1], Xuesong Feng[1], Shamima Islam[2], Stefania Dell'Orso[2], Davide Randazzo[3], Vittorio Sartorelli [1]

[1]Laboratory of Muscle Stem Cells and Gene Regulation, NIAMS, NIH, Bethesda, MD, [2]Genomic Technology Section, NIAMS, NIH, Bethesda, MD, [3]Light Imaging Section, NIAMS, NIH, Bethesda, MD

Muscle regeneration is a highly orchestrated spatial and temporal process involving diverse cell types and complex molecular interactions. In this study, we employ high-resolution spatial transcriptomics to construct a comprehensive molecular atlas of skeletal muscle regeneration following acute injury, revealing distinct and spatially organized transcriptional programs that guide tissue repair. Spatial clustering uncovers dynamic, niche-specific molecular signatures that evolve across key regenerative phases—including inflammation, proliferation, and tissue remodeling. Comparative analyses between young and aged muscle reveal striking age-related differences in regenerative capacity. Young muscle exhibits more transcriptionally distinct regions enriched for muscle fiber and regeneration markers, and fewer regions dominated by endothelial cell markers, efficiently progressing through regenerative phases. In contrast, aged muscle displays less regions associated with muscle fibers and regeneration, and an increased prevalence of endothelial cell–dominated areas, particularly during transitions into pro-myogenic states. This dysregulation corresponds with impaired satellite cell activation, aberrant immune responses, dysfunctional fibro-adipogenic progenitor activity, and defective extracellular matrix remodeling. Therefore, our findings underscore the power of spatial transcriptomics to decode the regenerative landscape and identify therapeutic targets to counteract age-related impairments in muscle repair.

A PAN-VERTEBRATE SINGLE-CELL ATLAS OF THE PITUITARY GLAND REVEALS EVOLUTIONARILY CONSERVED DRIVERS OF CELL IDENTITY

Bence Kover, Hui-Chun Lu, Cynthia Andoniadou

King's College London, Centre for Craniofacial and Regenerative Biology, London, United Kingdom

The pituitary gland is a central endocrine regulator, with diseases ranging from hormone deficiencies to tumor growth. Previous work on transcription factors (TFs) has already improved diagnosis and patient stratification, but a detailed understanding of the genetic programmes driving pituitary cell types is still lacking. To address this, we have recently generated a uniformly pre-processed Consensus Pituitary Atlas (CPA) using 1.2 million cells from 265 mouse samples. Specifically focusing on pituitary stem cells, we identified 83 TF markers, however, functional validation of all of these hits is not feasible. In addition, we have identified thousands of reproducible cell type markers (e.g., 1736 associated with stem cells), but the biological relevance of these genes remains elusive.

We reasoned that evolutionarily conserved cell type markers, including TFs, would be more likely to be functionally important and relevant to human health. To this end, we obtained single-cell profiling datasets from pituitary glands of 15 vertebrate species, covering fishes (zebrafish, medaka fish, large yellow croaker, grass carp, stickleback), amphibians (frog), birds (chicken), and several mammals (human, macaque, mouse, rat, pig, cow, water buffalo, Siberian tiger). We then created a workflow to identify marker genes, map these to orthogroups, and align them with the known phylogenies. Beyond identifying conserved markers, we were able to infer drivers of cell identity in the most recent common ancestor of these species, which lived ~400 MYA.

Our analysis revealed that pituitary stem cells are present across vertebrates and share a conserved set of 48 TFs, including SOX2. Curiously, we identified TFs whose specificity to stem cells was lost in either humans (7 TFs) or mice (another 7 TFs), highlighting the benefits of a pan-vertebrate approach compared to only studying mice. From stem cell markers other than TFs, we identified several members of Notch, Wnt and Hippo pathways as evolutionarily conserved.

Following this, we showed successful integration of this phylogenetically diverse dataset by modifying the variational autoencoder scVI. Specifically, we show that this is possible through using summed counts in orthogroups, rather than counts for individual genes (as genes map ambiguously across species). Furthermore, we found that supplementing the model with one-hot encoded phylogenetic information improved the integration results.

In summary, we generated a pan-vertebrate single-cell atlas to identify evolutionarily conserved genes and transcription factors driving pituitary cell type identity, highlighting their potential functional importance for human health. Our work demonstrates the power of comparative evolutionary approaches in single-cell biology and can be applied to any other organ of interest.

A LINEAGE-AWARE CROSS-SPECIES FOUNDATION MODEL FOR COMPARATIVE SINGLE-CELL GENOMICS

Chan Woong Lee*, Jaehoon Baek*, Jong Kyoung Kim

POSTECH, Department of Life Sciences, Pohang, South Korea

Single-cell RNA sequencing has revolutionized our understanding of cellular diversity within individual organisms. Comparative analyses across species hold great promise for uncovering fundamental principles of evolutionary conserved cellular functions and gene regulatory networks. However, existing computational methods are hindered by species representation bias, which obscures functional relationships across taxa. Here, we present a single-cell foundation model that enables balanced cross-species single-cell analysis by learning lineage-aware cell embeddings, leveraging cell ontology-based lineage mapping and systematic species oversampling. This approach ensures equitable representation of evolutionarily distant species, facilitating robust inference of functional lineage relationships. Training on five million cells from human, mouse, zebrafish, frog, and axolotl datasets, the model produces embeddings in which functionally related cells cluster by lineage identity rather than species origin. Importantly, the model generalizes effectively to previously unseen species, generating high-quality cell embeddings without the need for species-specific retraining. This lineage-aware foundation model advances comparative single-cell genomics by enabling systematic analysis of the evolution of cellular functions and transcriptional regulation.

MULTI-OMICS SINGLE-CELL ANALYSIS REVEALS KEY REGULATORS OF HIV-1 PERSISTENCE AND ABERRANT HOST IMMUNE RESPONSES IN EARLY INFECTION

Dayeon Lee, Sin Young Choi, So-I Shin, Hyunsu An

Gwangju Institute of Science and Technology (GIST), Life Sciences, Gwangju, South Korea

The clearance of human immunodeficiency virus-1 (HIV-1) remains a significant public health challenge due to impaired cellular immune responses and HIV-1 maintenance during acute infection. However, the genetic and epigenetic changes influencing the immune response on host infected cells remain unclear. Here, this study analyzes HIV-1 infected CD4+ T cells from peripheral blood mononuclear cells from people living with HIV-1 (PLWH) during early infection (<6 months) using single-cell RNA and ATAC sequencing. It is observed that HIV-1 hinders the antiviral response, particularly by interfering with the interferon signaling pathway. Multimodal analysis identifies KLF2 as a key transcription factor in infected CD4+ T cells. Moreover, cells harboring HIV-1 provirus are predominantly identified as Th17 cells, which exhibit elevated KLF2 activity. This suggests an increased susceptibility to HIV-1 infection and a constrained immune response due to the quiescent characteristics of these cells. The finding provides insights into the immune mechanisms and key regulators of HIV-1 maintenance in CD4+ T cells during the early stages of infection.

UNDERSTANDING THE ROLE OF THE RESOLUTION PARAMETER IN MODULARITY CLUSTERING OF SINGLE-CELL RNA-SEQ DATASETS

Sivan Leviyang

Georgetown University, Department of Mathematics and Statistics, Washington, DC

Single-cell RNA-seq (scRNAseq) datasets are commonly used to identify cell types through modularity clustering using the Louvain or Leiden algorithms. Modularity clustering involves a resolution parameter that controls the number of clusters formed. If the resolution parameter is set relatively high, many clusters form, some of which may not correspond to true cell types. Conversely, if the resolution parameter is set relatively low, few clusters form and distinct cell types may be grouped together. The importance of the resolution parameter in scRNAseq clustering is clear empirically and well known to practitioners. However, beyond the basic observation that resolution modulates the number of clusters, we have no quantitative description of its role in shaping scRNAseq clustering outcomes.

I will present novel results that describe clustering as a function of the resolution parameter. First, I will show that if a dataset contains cell types that are sufficiently similar — so that they are not separated at very low resolution (e.g., active vs. naïve B cells) — then the resolution value at which they separate is inversely proportional to the frequency of these cell types within the dataset. This implies that cell types comprising a small fraction of the dataset are only resolved into subtypes at relatively high resolution values. Second, I will demonstrate that high resolution values can induce clustering even in the absence of internal structure. I provide a quantitative description of such clustering by presenting a formula for the minimal resolution value that clusters multivariate normals. Putting the two results together, I will show that modularity clustering inevitably involves a tradeoff between correctly splitting some cell types into their subtypes and incorrectly splitting some cell types into false subtypes.

SINGLE-NUCLEUS TRANSCRIPTOMIC PROFILING IDENTIFIES CANDIDATE CYTOSKELETAL REGULATORS IN VASOPRESSIN NEURONS UNDER HIGH DIETARY SALT

Banruo Li[1], Olena Makashova[1], Ning Gu[1], Behrang Sharif[2], Andrew Bayne[3], Jonathan Fan[4], Masha Prager-Khoutorsky[1]

[1]McGill University, Department of Physiology, Montreal, Canada, [2]McGill University, Department of Neurology & Neurosurgery, Montreal, Canada, [3]McGill University, Department of Pharmacology & Therapeutics, Montreal, Canada, [4]McGill University, Department of Anesthesia, Alan Edwards Centre for Research on Pain, Montreal, Canada

High dietary salt (HDS) strongly correlates with cardiovascular diseases and is a major factor in hypertension development. While the link between HDS and hypertension is well established, the underlying molecular mechanisms remain unclear. Recent studies suggest HDS triggers neurogenic increases in sympathetic activity, vascular resistance, and fluid retention, implicating central sodium detection in salt-sensitive hypertension. Plasma sodium changes are detected by specialized osmosensory neurons in the supraoptic and paraventricular nuclei (SON and PVN). Under normal conditions, increased plasma sodium activates SON and PVN magnocellular neurons releasing vasopressin (VP), antidiuretic hormone causing renal water retention and vasoconstriction, to achieve fluid homeostasis. Chronic exposure to HDS is associated with excessive activation of VP neurons, leading to a VP-mediated blood pressure elevation. The mechanisms underlying excessive VP secretion in HDS are not fully understood. Our previous studies revealed that magnocellular VP neurons harbor unique cytoskeletal networks comprised of a subcortical actin layer, an array of actin comet-like structures, and a somatic scaffold of interweaved microtubules. VP neuron excitability is proportional to the density of these cytoskeletons. Moreover, chronic exposure to HDS increases actin and microtubule density in magnocellular VP neurons. To uncover the molecular mechanisms driving this cytoskeletal reorganization, we performed single-nucleus RNA sequencing of the SON in rats subjected to HDS. Cell-type annotation confirmed the presence of magnocellular VP neurons and revealed differential expression of genes associated with cytoskeletal dynamics. To complement the transcriptomic data, we also conducted bulk-level proteomic profiling of the same brain region. This analysis identified several cytoskeleton-associated proteins whose expression was altered by HDS, providing additional candidates for further investigation. Together, these datasets offer a multi-omic view of cytoskeletal remodeling in VP neurons under high-salt conditions. Future work will integrate the transcriptomic and proteomic findings to identify key differentially expressed genes and pathways, with the goal of selecting leading candidates for experimental validation.

SIMULATION-FREE APPROXIMATE BAYESIAN COMPUTATION FOR STOCHASTIC REACTION NETWORKS

Zekai Li, Mauricio Barahona, Philipp Thomas

Imperial College London, Mathematics, London, United Kingdom

In single-cell biology, stochastic reaction networks (SRNs) are widely used to model the production, degradation, and interaction of molecules. Inferring the underlying network structure and its parameters from observed data is crucial for an accurate understanding of the fundamental mechanism. Among the parameter inference methods proposed for SRNs, approximate Bayesian computation (ABC) is a Bayesian inference framework favored for its flexibility and the rich insights provided by the approximated posterior. However, a common drawback of ABC algorithms is that they typically require a large number of simulations even for relatively small systems, despite the exploration of various advanced sampling schemes aimed at improving efficiency. A range of approaches have been studied to approximate the simulation step, yet they often require strong assumptions such as a large sample size. In this work, we develop a simulation-free ABC (SFABC) framework that leverages theoretical constraints and convex optimization to determine if a sampled parameter set is feasible to produce the observed data for a given tolerance. We show that SFABC can be applied with both rejection and sequential Monte Carlo sampling schemes. Furthermore, we prove that the samples accepted by our algorithms converge to the true posterior distribution as the tolerance approaches zero. We demonstrate that SFABC can effectively handle both steady-state data and time-course data, just like standard ABC algorithms, but more efficiently in many cases. For some applications where ABC struggles to simulate data, we show that SFABC can still be applied and produce quality estimates of the posterior.

THE CURIOUS CASE OF LUNG TISSUE RESIDENT MEMORY T CELLS (T_{RM}): MAPPING CD4 T CELL DIFFERENTIATION

Jean-Christophe J Lone[1,2], Gonçalo Malpica[1,2], Lisa Dratva[3], Catarina Mendes[1], Rui do Amaral Vieira[1,2], Margarida Kirkby[1], Afonso Basto[1], Cristina Ferreira[1], Sarah Teichmann[3,4], Luís Graça[1,2], Marc Veldhoen[1,2]

[1]GIMM-Gulbenkian Institute for Molecular Medicine, Basic Immunology, Lisbon, Portugal, [2]Universidade de Lisboa, Faculdade de Medicina da Universidade de Lisboa, Lisbon, Portugal, [3]Cambridge Stem Cell Institute, University of Cambridge, Cambridge, United Kingdom, [4]University of Cambridge,, Department of Medicine, Cambridge, United Kingdom

Tissue-resident memory T cells (T_{RM}) are a subset of non-circulating T cells that persist in non-lymphoid tissues, where they provide localized and rapid immune responses upon infection. While T_{RM} cells exhibit long-term residency in most epithelial tissues, the lungs present a unique environment, as T_{RM} populations decline over time. This transient residency challenges our understanding of immune memory in the respiratory tract and raises fundamental questions about the mechanisms governing long-term immune protection against pulmonary infections. Our study investigates the molecular and cellular processes that regulate lung T_{RM} cell differentiation, retention, and egress.

To address this question, we designed an adoptive transfer experiment in which CD45.1 T cells were transferred into CD45.2 recipient mice. The recipients were subsequently challenged with three distinct infections: *Influenza X31* virus (Type 1), the helminth *Nippostrongylus brasiliensis* (Type 2), and the fungus *Aspergillus fumigatus* (Type 3). Single-cell RNA sequencing (scRNA-seq) was performed at 1, 2, and 3 weeks post-infection to assess the immune state of the lungs, analysing gene expression profiles and T cell receptor (TCR) VDJ sequences of CD4 T cells. This approach provided a a unique cellular cartography of pulmonary CD4 T cell responses across infection timelines, along with their circulating or lymphoid tissue-restricted phenotype.

Using this cartography of CD4 T cells in space and time, we constructed a high-resolution trajectory tracing the progression from naïve T cells to mature T_{RM} subsets, elucidating the transition from circulating to tissue-resident states. This analysis revealed distinct differentiation pathways leading to Type 1, Type 2, and Type 3 T_{RM} cells, as well as a regulatory subset. Importantly, T_{RM} cell differentiation in the lungs was not stochastic, we identified a strong association between T_{RM} cell fate and TCR specificity, suggesting a selective process shaping local immune adaptation. Using RNA velocity analysis to capture differentiation dynamics, we identified key transcriptional programs that promote tissue residency. These programs not only confer resilience to environmental stressors but may also enhance functional robustness, allowing T_{RM} cells to better adapt to repeated immune challenges.

A DEEP-LEARNING BASED APPROACH FOR REGION OF INTEREST (ROI) OPTIMIZATION IN SPATIAL TRANSCRIPTOMICS

Shupeng Luxu[1,2,3], Bokai Zhu[1,2,4], Alex K Shalek[1,2,4]

[1]Ragon Institute, Shalek Lab, Cambridge, MA, [2]Massachusetts General Hospital, Shalek Lab, Boston, MA, [3]Harvard School of Public Health, Department of Biostatistics, Boston, MA, [4]Massachusetts Institute of Technology, Shalek Lab, Cambridge, MA

After single-cell technology revolutionized our ability to study biology at the cellular level, spatial transcriptomics has further advanced our understanding of cell-to-cell interactions and tissue structure. However, a major challenge remains: compared to single-cell genomics modalities, spatial-omics technologies easily generate millions of cells from one single tissue scan, and up to couple millions with a large sample cohort. This brings in a significant challenge for the field: extensive instrument run time, increase reagent usage, and more importantly, significant burden on the computational resources required to perform general analysis tasks, where most existing tools struggles with multi-million cell inputs.
To overcome this, there is an urgent need for an automated, data-driven strategy to optimize regions of interest (ROIs) selection/covering that can fully cover the tissue's biological complexity, thus enabling more efficient and cost-effective experiments and computational analysis, while maintaining the same biological insight discovery.
To address this critical bottleneck, we propose a novel deep-learning based method for guided ROI selection from standard histological images, specifically H&E stains. Our approach integrates multi-scale embeddings of tissue crops using a self-supervised learning framework. We then apply a geometric sketching algorithm to objectively subsample a small set of non-overlapped tissue crops as representative ROIs. To evaluate the effectiveness of our subsampling strategy, we adapted ecological metrics — species accumulation curve and feature richness estimators—to quantify the biological feature representation of the selected ROIs compared to the entire tissue. Our results demonstrate that our method achieved high coverage of all cell and tissue embedding clusters, indicating a robust representation of the full biological landscape. Furthermore, through power analysis, we found that sampling as few as 10-20 ROIs was sufficient to capture over 90% of the biological features present in the entire tissue. This study introduces an automated, objective, and unbiased approach for ROI selection in spatial transcriptomics, potentially streamlining experimental design and reducing costs without relying on subjective human expertise.

PAIRED SINGLE-CELL MULTIOMICS AND LONG-READ SEQUENCING CAPTURES IMPACT OF SOMATIC MUTATIONS AND HIV ON T CELL CLONAL EXPANSION

Haocong Katherine Ma[1], Tsung-Chih Chen[1], Irini Sereti[2], Ya-Chi Ho[1]

[1]Yale University, Department of Microbial Pathogenesis, New Haven, CT, [2]National Institutes of Health, Division of Intramural Research/NIAID, Bethesda, MD

Multimodal single-cell sequencing enables profiling of heterogeneous populations and rare cells, but detecting sparse somatic mutations remains challenging. Most cancer-associated mutations are distributed across the genome and missed by short-read Illumina platforms, while long-read methods (Nanopore, PacBio) offer broader coverage but lack sufficient throughput for rare events. As cancer often originates from rare mutant clones, early identification of such clones is essential for prevention and intervention. Similarly, HIV-infected cells are extremely rare (<0.01%) but remain the source of viral rebound under ART. Distinguishing intact from defective HIV genomes is therefore essential to understand HIV persistence.

To address these challenges, we leveraged an early pre-malignant hematopoiesis disease called clonal hematopoiesis (CH) as a model system in people living with and without HIV. We applied single-cell multi-omic ECCITE-seq coupling with genotyping of transcripts (GoT) to simultaneously capturing DNMT3A genotype, HIV RNA, transcriptome, surface proteins, and T cell receptor sequences. Briefly, we spiked in 5 DNMT3A- and 11 HIV-specific primers before ECCITEseq to enrich for mutations across the DNMT3A coding region and for near full-length HIV RNA. Using the ECCITE-seq cDNA library, we performed targeted DNMT3A and HIV-1 amplification (preserving single-cell barcodes), MAS-ISO-seq concatenation, and PacBio sequencing. This multi-modality platform directly captured somatic mutations and viral genomes in the rare cell populations.

Together, we profiled 307,535 CD4+ T cells from three cohorts: HIV+ CH– patient (55,050 cells), HIV+ CH+ patients (187,999 cells), and HIV– CH– donors (64,486 cells). Short-read scRNA-seq detected DNMT3A RNA in 32,554 cells (10.3%), while long-read sequencing detected 19,898 cells (8.6%) with unbiased coverage across mutational hotspots. Mapping short-read data to autologous HIV genomes identified 137 RNA^+ HIV-infected cells (0.073% of $CD4^+$ T cells). In contrast, long-read sequencing identified 3,297 HIV-infected cells (1.76%), representing a 24-fold increase in sensitivity, and uniquely resolved 9 intact proviruses.

To understand if CH promotes the proliferation of HIV-infected cells, we recapitulated T cells clones based on shared TCR junction sequences. We identified 16,274 unique T cell clones with 54,763 cells in total. Larger clones were enriched in cytotoxic CD4+ T cells expressing high GZMB and GZMH. Notably, we identified 100 HIV-infected cells (58.8%) resided in T cell clones. CD4+ T cells from HIV+ CH+ patients exhibited significantly greater clonal expansion and larger clone sizes than those from either HIV+ CH– or HIV– CH– donors. Overall, this study integrates short- and long-read single-cell multi-omics to overcome technical barriers in DNMT3A and HIV genotyping, revealing that DNMT3A mutation in CH enhances clonal expansion of HIV-infected $CD4^+$ T cells, which may promote HIV persistence.

NOVEL HIGH-THROUGHPUT TECHNOLOGY FOR EVALUATING DNA METHYLATION OF OPEN CHROMATIN IN SINGLE CELLS

Vincenta Mikulenaite, Linas Mazutis

Vilnius University, Department of Single Cell Analytics, Vilnius, Lithuania

Recent advancements in single-cell technologies have paved the way for elucidating the intricate interplay between DNA methylation and chromatin accessibility, revealing their combined impact on gene regulation and cellular heterogeneity. However, the majority of current experimental methods rely on the use of tubes or wells, thus limiting the scalability that could be offered by a microfluidics system.

In this study, we present a novel high-throughput technology for evaluating the connection between DNA methylation and chromatin accessibility in single mammalian cells. High-throughput is ensured by a microfluidics platform that isolates single cells into semi-permeable microcapsules, which exhibit specific properties tailored for capturing the genomic DNA methylation profile of open chromatin regions. Unlike droplets, microcapsules comprise a membrane that acts as a selective barrier, retaining cells and their genomic material while allowing smaller molecules, such as chemical reagents, DNA primers, or enzymes, to pass through via passive diffusion. The selective permeability is crucial not only for maintaining efficient biochemical reactions but also for enabling combinatorial DNA barcoding through a split-and-pool technique to profile millions of single cells. The encapsulated cells are then lysed under mild conditions to preserve the chromatin state. After endonuclease treatment, the fragmented genome (open chromatin) is treated enzymatically to convert cytosines to uracils. Finally, sequencing libraries are prepared by performing multiple rounds of barcoded adaptor ligations using the split-and-pool method to barcode and amplify the genomic DNA fragments for Illumina sequencing.

Sequencing results reveal the potential of this technology to determine the methylation status of open chromatin in single cells, offering future insights into complex cellular mechanisms involving cell differentiation and disease progression.

SINGLE CELL SEQUENCING SHOWS T CELLS ADD COMPLEXITY TO SCLERODERMA 3D SKIN-LIKE TISSUES

Madeline J Morrisson[1,2], Helen C Jarnagin[1,2], Isha Singh[3], Sasha Shenk[3], Jonathan A Garlick[3,4,5,6], Michael L Whitfield[1,2]

[1]Geisel School of Medicine at Dartmouth, Biomedical Data Sciences, Hanover, NH, [2]Geisel School of Medicine at Dartmouth, Molecular & Systems Biology, Hanover, NH, [3]Tufts University School of Dental Medicine, Basic & Clinical Translational Sciences, Boston, MA, [4]Tufts University School of Medicine, Developmental, Molecular, & Chemical Biology, Boston, MA, [5]Tufts University School of Engineering, Biomedical Engineering, Medford, MA, [6]Tufts University School of Dental Medicine, Tufts Initiative in Civic Science, Boston, MA

Scleroderma (systemic sclerosis, SSc) is a rare systemic autoimmune disease characterized by skin fibrosis and internal organ dysfunction. Molecular mechanisms of the disease are poorly understood, but there is a strong immune component. We previously developed and validated an *in vitro* 3D skin tissue model with autologous patient-derived (APD) monocytes and fibroblasts, and keratinocytes forming an epidermis. To better recapitulate SSc, T cells were added to human skin equivalent (HSE) tissues, analyzed functionally and by single cell RNA sequencing (scRNA-seq).

HSE tissues were constructed by growing APD fibroblasts, CD4 T cells, and monocytes in collagen matrix for 7 days. Neonatal keratinocytes were seeded on tissues and cultured for 4 days, before lowering media to allow for keratinization. HSE tissues were digested enzymatically using a Cytivia VIA Extractor or frozen for staining. Cell suspensions were fixed using 10x Genomics' Chromium Fixation, followed by scRNA-seq. Data were analyzed with RStudio.

After Seurat clustering, cells were identified using canonical markers; THY1-fibroblasts, KRT1-keratinocytes, PTPRC-immune, CD3E-T cells, MRC1-macrophages. Staining confirmed the presence of cell types in normal skin architecture.

Fibroblast clustering revealed profibrotic (COL8A1), inflammatory (CXCL12), motility (HOX9), and proliferating (CDC20) groups. We hypothesized that the inflammatory and motility groups could be disease drivers and found these groups make up a greater proportion of SSc cells than healthy control.

T cell clustering revealed naive (CCR7), activated (STAT6), proliferating T cells (MKI67), and Tregs (FOXP3). With CellChat we found signaling interactions between fibroblasts and T cells, namely increased type 1 interferon (IFN-I), a cytokine that guides inflammation and T cell response. IFN-I was sent by proliferating fibroblast population and received by all populations.

New storage and preparation methods supported analysis of complex HSE tissues. Including T cells in HSE tissues allowed us to examine interactions between fibroblast and immune cells, contributing to recapitulation of interactions in SSc patient and healthy human skin.

scFLRNA-SEQ – A NOVEL METHOD FOR SINGLE-CELL FULL-LENGTH RNA QUANTIFICATION USING SHORT-READ SEQUENCING

Simonas Norvaisis, Denis Baronas, Linas Mazutis

Vilnius University, Single Cell Analytics, Vilnius, Lithuania

Full-length RNA sequencing is a crucial advancement in cell biology studies as it enables comprehensive analysis of the complete sequence of RNA molecules, providing insights into transcript diversity, alternative splicing events, and accurate identification of single-nucleotide polymorphisms (SNPs).
Here we present a method for single-cell full-length RNA sequencing by using semi-permeable capsules (SPC). Our method, integrated in semi-permeable capsule technology allows to sequence full-length transcripts of single cells by short-read (Illumina) sequencing technology. We show, how this method can quantify transcript variants and single nucleotide polymorphisms in single human K562 and mouse NIH/3T3 cells.
This approach effectively combines the advantages of single-cell analysis using high throughput and low handling complexity of SPC-based experiments together with the high accuracy and price-efficiency of short-read sequencing technology, enabling scalable and detailed transcriptomic profiling.
scFLRNA-seq promises to enhance the ability to study single-cell cellular processes in greater detail and affordability and offers significant potential for applications in diagnostics and therapeutics, where understanding the diversity and complexity of the full-length transcriptome at a single-cell resolution is paramount.

A SINGLE-CELL VIEW OF SQUID–VIBRIO SYMBIOSIS

Conrad Oakes, Vera Beilinson, Lior Pachter

California Institute of Technology, Biological Engineering, Pasadena, CA

Single-cell RNA-seq has been used to uncover host-microbe interactions - typically by exposing the eukaryotic host to a bacteria and seeing how individual host cells respond. It is also known that bacterial populations themselves are also heterogeneous, though single-cell analysis of this heterogeneity has been far less explored. Here, we investigate how a bacterial symbiont and its host interact at the single-cell level using the squid–vibrio system by applying snRNA-seq to the squid host and scRNA-seq to its bacterial symbiont. Addressing these novel questions required new computational approaches, and we highlight the tools developed here to advance the study of host–symbiont relationships at the single-cell level.

BONE-DEEP FINE-TUNED SPATIAL TRANSCRIPTOMICS MAPS THE SKELETAL CELLULAR LABYRINTH

Hiroyuki Okada[1,2,3], Fumiko Yano[4], Asuka Terashima[5], Sanshiro Kanazawa[6], Yasunori Omata[2,5], Francesca Gori[3], Taku Saito[2], Shigeto Seno[7], Junichi Kikuta[8], Roland Baron[3], Ung-il Chung[1,9], Sakae Tanaka[2], Hironori Hojo[1,9]

[1]the University of Tokyo, Center for Disease Biology and Integrative Medicine, Tokyo, Japan, [2]the University of Tokyo, Department of Orthopaedic Surgery, Tokyo, Japan, [3]Harvard School of Dental Medicine, Department of Oral Medicine, Infection, and Immunity, Boston, MA, [4]Showa Medical University , Research Administration Center, Tokyo, Japan, [5]the University of Tokyo Hospital, Bone and Cartilage Regenerative Medicine, Tokyo, Japan, [6]the University of Tokyo, Department of Oral and Maxillofacial Surgery, Tokyo, Japan, [7]the University of Osaka , Department of Bioinformatic Engineering, Graduate School of Information Science and Technology, Osaka, Japan, [8]Kobe University, Division of Immunology, Department of Future Medical Sciences, Graduate School of Medicine, Kobe, Japan, [9]the University of Tokyo, Department of Bioengineering, Graduate School of Engineering, Tokyo, Japan

Application of spatial transcriptomics (ST) to bone is difficult because decalcification and dendritic osteocytes confound segmentation. We developed a fine-tuned ST pipeline integrating histology to overcome these barriers.

We built a reference from 12 mouse bone and marrow datasets (72k cells, 29k genes). FFPE femoral sections from wild type and ovariectomized (OVX) mice were decalcified for 72 h and profiled on the 10x Genomics Xenium In Situ v2 platform with a custom 100-gene panel. H&E images were co-registered to DAPI and fused with Xenium coordinates, enabling segmentation.

Our atlas found two sclerostin-positive osteocyte clusters with remodeling roles. With our method, ~100k cells per femur were delineated, assigning 93% of transcripts to cells (vs Xenium default: 17k, 31%). Label transfer distinguished cortical from trabecular subsets, with shifts in the OVX model.

Benchmarking tools like Cellpose showed general-purpose tools produced fewer cells or unrealistic morphology. Our approach couples H&E morphology with DAPI anchoring and transcript evidence, yielding better segmentation and fidelity.

This is a breakthrough for skeletal biology. Limited to morphology, bone research lacked a spatial framework. Our pipeline converts bone into digital maps uniting histology, geometry, and gene expression. Through fine-tuned ST, Xenium signals reveal osteocyte heterogeneity and remodeling niches.

The segmentation strategy is transferable beyond bone. Unlocking spatial omics in calcified tissues provides a tool comparable to scRNA-seq, enabling atlases and actionable insights for osteoporosis and skeletal disease.

VISUALIZING METABOLIC FLUX OF DEUTERATED GLUCOSE, GLUCOSE, AND XYLOSE IN S. CEREVISIAE STRAINS ENGINEERED FOR XYLOSE METABOLISM USING STIMULATED RAMAN SCATTERING MICROSCOPY

Erik Oliai[1], Leon Han[1], Bryce Manifold[1], John Dueber[2], Aaron Streets[1,3]

[1]University of California, Berkeley, Bioengineering, Berkeley, CA, [2]University of California, Berkeley, Innovative Genomics Institute, Berkeley, CA, [3]Chan Zuckerberg Biohub SF, University of California, Berkeley, San Francisco, CA

Efficient utilization of xylose, the second most abundant sugar in lignocellulosic biomass, remains a central challenge for engineering *Saccharomyces cerevisiae* strains toward biofuel and biochemical production. Although heterologous expression of *Pichia stipitis* genes has enabled xylose assimilation, flux partitioning between glucose and xylose remains inefficient due to redox imbalance and xylitol accumulation bottlenecks. A clear strategy for designing strains that co-utilize these sugars optimally has yet to be established.

Here, we apply stimulated Raman scattering (SRS) microscopy as a non-destructive, single-cell imaging method to visualize carbon flux from deuterium-labeled sugars in the engineered strain yJD239. The strain was generated through a combinatorial promoter library, tuned for xylose metabolism. We cultured cells under six carbon conditions: glucose, deuterated glucose (d7), xylose, and mixtures of d7-glucose and xylose at varying ratios. SRS imaging enabled direct visualization of isotope incorporation into biomass and lipids.

Our findings demonstrate strong C–D incorporation in cultures supplied with d7-glucose, with signals localized to cytoplasmic regions and lipid droplets. In mixed cultures containing 95% xylose and 5% deuterated glucose (d7), deuterium incorporation was restricted to a subset of cells, indicating heterogeneity in substrate preference. Pure xylose cultures showed altered morphology, with cells adopting more rod-like shapes and reduced growth compared to glucose-fed populations, consistent with reduced metabolic efficiency. Collectively, these results suggest that even engineered strains optimized for xylose assimilation retain a preferential bias for glucose metabolism.

These preliminary findings validate SRS microscopy as a sensitive platform for mapping deuterium flux and uncovering cell-to-cell heterogeneity in engineered yeast cells. Future work will expand to testing additional engineered yeast strains, LL121A and LLTDH3C. Our approach highlights how SRS imaging can serve as a rapid, real-time screening tool for evaluating the functional impact of gene or promoter modifications during strain optimization, allowing researchers to directly visualize the effectiveness of genetic edits on cellular metabolism.

COMPARING PHENOTYPIC MANIFOLDS WITH KOMPOT: DETECTING DIFFERENTIAL ABUNDANCE AND GENE EXPRESSION AT SINGLE-CELL RESOLUTION

Dominik J Otto[1,2,3], Erica Arriaga-Gomez[4], Elana Thieme[1,2,3,4], Ruijin Yang[1,2,3], Stanley C Lee[4], Manu Setty[1,2,3]

[1]Fred Hutchinson Cancer Center, Basic Science Division, Seattle, WA, [2]Computational Biology Program, Public Health Division, Seattle, WA, [3]Fred Hutchinson Cancer Center, Translational Data Science IRC, Seattle, WA, [4]Fred Hutchinson Cancer Center, Translational Science and Therapeutics Division, Seattle, WA

Single-cell data provides snapshots of the diverse cellular states present within biological systems. We have previously demonstrated that the distribution of these cell states can be represented with a continuous density function, providing a comprehensive representation of the phenotypic landscape.
In the Kompot framework, we first co-embed cells from multiple conditions into a shared lower-dimensional cell-state space. We then construct distinct "density maps" - cell-state density functions for each condition. By comparing these maps, we can identify regions of differential abundance. Areas where one condition shows significantly higher or lower cell density than the other, revealing subtle shifts in cellular composition between conditions without discretization. Perturbations often induce condition-specific alterations in gene expression within equivalent cell states. To capture these expression changes, Kompot models gene expression as condition-specific continuous functions over the shared cell-state space. This approach enables direct comparison of expression levels at corresponding positions in the phenotypic landscape, identifying genes that are differentially expressed between conditions. Importantly, our framework derives uncertainty measures for both density and gene expression functions, ensuring that differential abundance and differential expression tests are statistically rigorous for single-cell resolution analysis.

BIOPHYSICAL MODELING OF SPATIAL TRANSCRIPTOMICS DATA

Kayla Jackson[1], Gennady Gorin[2], Lior Pachter[1]

[1]California Institute of Technology, Division of Biology and Biological Engineering, Pasadena, CA, [2]Fauna Bio, Emeryville, CA

Analysis of spatially-resolved transcriptomics (SRT) data has the potential to enhance understanding of cell-cell interactions, developmental processes, and disease mechanisms at an increasingly resolved spatial scale. However, naive application of analysis methods developed for scRNA-seq is fundamentally incompatible with spatial analysis because of implicit and explicit assumptions of independent observations. Further, sequencing-based methods are subject to many of the same technical factors that are present in single-cell RNA-seq (scRNA-seq) that can limit inference of spatial relationships in gene expression and confound the underlying biology. The rapid advancement of sequencing-based SRT technologies offers an opportunity to explore a basis for spatially correlated gene expression

We propose a biophysical model that attributes spatial dependence in gene expression partly to spatially structured technical noise, and we show that this model results in better fits to the count data than the independent model. This perspective underscores the need for physically grounded models that explicitly incorporate spatial structure when analyzing SRT data.

MAPPING LINEAGE-RESOLVED scRNA-SEQ DATA WITH SPATIAL TRANSCRIPTOMICS USING TEMSOMAP

Xinhai Pan[1], Alejandro Danies-Lopez[1], Yiru Chen[2], Xiuwei Zhang[1]

[1]Georgia Institute of Technology, School of CSE, Atlanta, GA, [2]Zhejiang University, College of Life Sciences, Hangzhou, China

Spatial transcriptomics (ST) has become a powerful technique that advances the study of cell spatial organization and cell-cell interactions. While ST can preserve location information of cells or spots, limitations of such technologies include lower number of genes, and lower resolution compared to scRNA-seq datasets. These limitations can be alleviated by integrating scRNA-seq data with the ST data. By mapping the single cells onto the spatial data, we can infer the spatial coordinates of the cells from the scRNA-seq dataset.

We consider leveraging temporal information in this challenging task of spatial location inference. During tissue formation, cells divided from the same ancestor are likely to be located close to each other in the tissue, thus the cell clonal or lineage information can improve cell location inference. Recent lineage tracing technologies such as CRISPR/Cas9-based lineage tracing technologies have enabled paired sequencing of cells' gene expression and lineage barcodes. The lineage barcodes can be used to reconstruct the cell lineage tree, which represents cells' clonal relationships.

We believe that the lineage tree can be a valuable modality to help improve the accuracy of mapping single cells onto space. Therefore, we developed TemSOMap (Temporal dynamics guided Spatial Omics Mapping), which infers the spatial coordinates of cells by mapping a paired gene expression and lineage barcode dataset onto a spatial transcriptomics dataset. TemSOMap takes advantage of the assumption that cells from the same clone are likely to be located closely in space andutilizes a machine learning framework to infer a cell-to-spot mapping matrix by minimizing a loss function based on expression and lineage.

We show that TemSOMap more accurately infers the spatial location of single cells compared to state-of-the-art baseline methods under various scenarios, using both simulated and real datasets. On real datasets, TemSOMap is better at preserving spatial distributions of cell types and maintaining spatial gene expressions. The resulting lineage-resolved ST data can help us better understand the spatio-temporal dynamics of cells in a tissue. TemSOMap is the first method that integrates cell lineage information with ST data, which has the potential to help us better understand the spatio-temporal dynamics of cells.

TemSOMap is publicly available at https://github.com/ZhangLabGT/TemSOMap.

TRANSCRIPTS CONTAINING RETROTRANSPOSABLE ELEMENTS PLAY A REGULATORY ROLE IN MODULATING THE FUNCTIONS OF TUMOR INFILTRATING LYMPHOCYTES

Francesco Panariello[1], Rebecca Vadalà[1,2], Valeria Ranzani[1], Alen Stambolliu[1,3], Benedetto Polimeni[1,2], Michele Panepuccia[1,3], Federica Marasca[1] , Salvatore Siena[4,5], Andrea Sartore-Bianchi[4,5], Riccardo Rosati[6,7], Sergio Abrignani[1,2], Beatrice Bodega[1,2]

[1]INGM, Istituto Nazionale di Genetica Molecolare, Genome Biology Unit, Milan, Italy, [2]University of Milan, Department of Clinical Science and Community, Milan, Italy, [3]University of Milan, SEMM, European School of Molecular Medicine, Milan, Italy, [4]University of Milan, Department of Oncology and Hemato-Oncology, Milan, Italy, [5]Grande Ospedale Metropolitano Niguarda, Department of Hematology, Oncology, and Molecular Medicine, Milan, Italy, [6]IRCCS San Raffaele Scientific Institute, Department of Gastrointestinal Surgery, Milan, Italy, [7]Vita-Salute San Raffaele University, School of Medicine, Milan, Italy

Unravelling Tumor Infiltrating Lymphocytes (TIL) heterogeneity is essential to develop more effective immunotherapies. We have recently highlighted the role of LINE1, the predominant class of Retrotransposable Elements (RE), in regulating TIL exhaustion by integrating in transcripts related to T cell activation and downregulating their canonical isoforms. Nevertheless, the role of RE expression as a determinant of TIL diversity and function remains poorly explored. To elucidate this relationship, we have isolated CD3+ and CD45+ T cells from tumor resections of Non-Small Cell Lung Cancer (NSCLC) and Colorectal Cancer (CRC) patients. Due to the challenging task of studying transcripts containing RE with current single-cell technologies, we performed both long- and short-read scRNA-seq. We have developed a custom pipeline to analyze long-read data to generate a cancer-specific Atlas of RE- transcripts, consisting of approximately 40,000 novel transcripts that are specific to NSCLC and CRC, with some shared instances. Among these transcripts, about 46% included REs, resembling their genomic distribution. We next quantified them at isoform level using full-length short-read data. We have introduced a novel strategy of integrating dimensionality reduction performed on novel RE-transcripts with trajectory inference tools. We thus confirmed the positive correlation between LINE1 and exhausted TILs, but it also revealed an enrichment of Alu-containing transcripts in active T cells, while regulatory T cells (Treg) mostly expressed transcripts containing the LTR class. By the integration of cutting-edge sequencing technologies and computational workflows, this project unveils the biological relevance of RE-transcripts in TILs identity. It also introduces the innovative concept that RE-transcripts may serve as novel immunological checkpoints and a promising avenue for cancer therapy.

SINGLE-CELL LANDSCAPE REVEALS FIBRO-INFLAMMATORY DRIVERS OF HUMAN VISCERAL ADIPOSE DYSFUNCTION

Eun Seo Park[1], Joon Ho Moon[2], Jisu Jeong[2], Sung Hee Choi[2], Jong Kyoung Kim[1]

[1]POSTECH, Department of Life Sciences, Pohang, South Korea, [2] Seoul National University Bundang Hospital, Division of Endocrinology and Metabolism, Department of Internal Medicine, SeongNam, South Korea

Adipose tissue dysfunction drives metabolic disease pathogenesis through poorly understood cellular mechanisms. We performed integrated single-nucleus RNA sequencing and bulk transcriptomics on human visceral adipose tissues from metabolically distinct cohorts: lean controls, obese subjects, and type 2 diabetic patients. Single-cell analysis identified progressive expansion of fibro-inflammatory adipose stem cell and adipocyte subpopulations with disease severity. These pathological cells exhibited enhanced extracellular matrix gene expression and inflammatory signaling signatures. Mechanistically, we discovered a novel matricellular protein pathway orchestrating tissue fibrosis via macrophage-mediated TGF-β activation. This cascade disrupts adipogenesis, impairs lipid metabolism, and sustains chronic inflammation. Pathway inhibition reversed fibrotic remodeling and restored metabolic function in human tissue explants. Circulating and tissue-specific pathway biomarkers correlated strongly with BMI, insulin resistance, and glycemic parameters. Transcriptional network analysis revealed cell-type-specific regulatory circuits governing adipose dysfunction. Our findings establish a comprehensive cellular atlas of human visceral adipose pathology and identify therapeutically targetable mechanisms underlying metabolic disease progression. This work provides mechanistic insights into adipose tissue dysfunction and highlights promising therapeutic interventions for metabolic disorders.

BIOLOGICAL INSIGHTS FROM INTEGRATING POPULATION-SCALE, CROSS-TISSUE, CROSS-DISEASE HUMAN CELL ATLASES

Nikolaos Patikas[1,4], Hongcheng Yao[2,3,5], Ilya Korsunsky[2,3,5], Martin Hemberg[1,4]

[1]Brigham and Womens Hospital, Neurology, Boston, MA, [2]Brigham and Womens Hospital, Genetics, Boston, MA, [3]Brigham and Womens Hospital, Center for Data Sciences, Boston, MA, [4]Harvard Medical School, Immunology, Boston, MA, [5]Harvard Medical School, Medicine, Boston, MA

During recent years, the generation of single-cell genomics datasets has become more accessible to researchers. Consequently, the data volumes of available single-cell datasets grow exponentially. Individually, each dataset aims to address specific biological questions. When combined, these datasets can reveal new patterns and associations in human tissues and diseases, and translate to novel biological insights and hypotheses. However, the noisy nature and size of such datasets present significant technical challenges to achieving these goals. Here, we demonstrate an efficient and flexible workflow for integrating such datasets by refining Harmony, one of the most popular batch correction algorithms. We demonstrate the ability to analyze single-cell data from human populations in a single computer using out-of-core approaches. Our methodology involves generating cell embeddings, integrating data across multiple conditions, and utilizing a flexible cell type annotation framework. Our approach enables us to compile and reuse existing single-cell genomics datasets at a population scale. To investigate the biological utility of our approach, we integrated the Chan Zuckerberg Initiative's CellXGene census, utilizing 42 million human cells from 7,000 human donors across 50 tissues and 100 diseases. Using this large cohort, we show how to decompose and extract molecular signatures and biological insights from individual donors. Additionally, we extend this framework to identify novel associations that investigate cross-tissue and disease heterogeneity of human cells, enabling holistic views of disease and tissue-resident cells with a particular focus on the immune system and the brain.

MULTI-OMICS ANALYSIS OF INTERSTASIS IN AGEING AND NEURODEGENERATION

Martina Pedna[1,2], Silvia Hnatova[2], Owen Gwydion James[2], Charlotte Capitanchik[2], Jernej Ule[2], Alessandra Vigilante[1]

[1]Centre for Gene Therapy and Regenerative Medicine, King's College London, London, United Kingdom, [2]UK Dementia Research Institute, King's College London, London, United Kingdom

Interstasis is a mechanism that promotes dosage homeostasis for a class of genes encoding condensation-prone proteins, characterised by mixed-charge, low-complexity domains (R-MCDs). Condensation prone proteins are often altered in abundance and solubility in the context of ageing and neurodegenerative diseases.
Inducing interstasis, either through overexpression of various R-MCDs or inhibition of Cdc2-like kinases (CLKs), leads to altered nuclear/cytoplasmic localisation of mRNA groups with distinct nucleotide compositions. Specifically, GA-rich mRNAs exhibit enriched nuclear localisation due to retention in nuclear speckles (pro-retention mRNAs), whereas GU-rich mRNAs preferentially locate in the cytoplasm (anti-retention mRNAs).

In addition, we observe characteristic alternative polyadenylation (APA) of mRNAs in interstatic cells. We leveraged these APA signatures, together with nuclear retention profiles from bulk sequencing and nucleotide composition features, to classify single nuclei and investigate interstasis in a cell-type specific context and in relation to demographic and disease metrics. We focused on the Seattle Alzheimer's Disease Brain Cell Atlas (SEA-AD) cohort, which comprises single-nuclei from over 80 individuals across two brain regions: the Middle Temporal Gyrus and the Dorso Lateral Prefrontal Cortex. Our results indicate that neurons coming from both brain areas populate a range of states, which shifts in ageing and dementia. Ultimately, our work presents an additional layer of understanding to the widespread dysregulation of RNA and protein homeostasis.

GCAT, IT'S A SCAM: AN OPEN SOURCE PIPELINE FOR SINGLE CELL DATA PROCESSING AND QUALITY CONTROL

Rishvanth K Prabakar, Hannah V Meyer

Cold Spring Harbor Laboratory, The Simons Center for Quantitative Biology, Cold Spring Harbor, NY

The era of designing bioinformatic pipelines using open source tools is giving way to monolithic, and often proprietary, tools designed for end-to-end analysis. While these tools can be convenient for going from sequencing data to an almost complete data analysis "in a few clicks", they are severely limiting to perform any custom analysis. Here, we developed Gene and Chromosome Analysis Toolkit (gcat) and It's a scam To Sell Another Single Cell Analysis Method (its_a_scam) for processing and quality control of single cell genomic data, freeing users from the reliance on proprietary tools for single cell data analysis. gcat is a C++ library (dependent only on HTSlib) for most common tasks associated with genomics data such as processing and handling reads, alignments, and annotations. its_a_scam is built on top of gcat, and provides a suite of tools for a wide range of single cell data processing and quality control. its_a_scam can be used in conjugation with commonly used open source tools for scRNA-seq count matrix generation, scATAC-seq peak matrix generation, and for performing QC. We show that the count matrices generated with the aid of its_a_scam for Parse Biosciences scRNA-seq and 10x scATAC-seq data are highly concordant with those generated by the respective proprietary pipelines. gcat and its_a_scam are well documented and we provide detailed descriptions for a user to generate count matrices and QC single cell datasets starting from FASTQ files. Of course, they are open source (licensed under GNU GPL version 2) enabling users to adapt and build on our tool as desired.

UNTANGLING THE TRANSCRIPTOMIC SIGNATURES OF TISSUE AND CELL TYPE IDENTITY

Dominic Pruss, Sarah Chapin, Rishvanth Prabakar, Erin Groce, Hannah Meyer

Cold Spring Harbor Laboratory, Simons Center for Quantitative Biology, Cold Spring Harbor, NY

Single-cell analyses typically classify cells into cell types defined by structural and functional commonalities across different groups of cells. However, cells of the same type often specialize across tissues: for example, gene expression in epithelial cells present in digestive organs differs significantly from the gene expression of epithelial cells present in the skin. A more nuanced classification of cell phenotypes that considers both cell type and tissue contributions will thus better characterize cellular function in particular tissue contexts. We hypothesize that cellular phenotypes can be decomposed into gene regulatory programs shared by cells of similar function situated in different tissues. In order to identify these programs, we curated a human scRNA-seq dataset downsampled from the CELLxGENE Census database. The complete CELLxGENE database includes data from over 500 scRNA-seq studies and 60 million unique cells and spans over 800 human cell types and 370 human tissues; however, cell tissue and type classifications are inconsistent across component studies and the dataset is impractical to use in its entirety due to its size. We therefore harmonized the cell tissue and type annotations into 53 distinct tissues and 58 distinct cell types, and developed a protocol to downsample the full dataset while yielding reproducible results in single-cell analyses. We determined the minimum sample size required per cell-tissue type combination to obtain robust results across three analytic contexts: correlation analysis of pseudobulked gene-expression matrices, MetaNeighbor analysis of cell tissue-type replicability and Weighted Gene Co-expression Network Analysis of gene modules across cell tissue-types. Our final downsampled dataset consists of 1.6 million cells and can be used to identify cell tissue-type defining gene regulatory programs and in single-cell analyses more broadly.

VOLUMETRIC DNA MICROSCOPY FOR THREE-DIMENSIONAL SPATIAL TRANSCRIPTOMICS IN INTACT TISSUES

Nianchao Qian[1], Joshua Weinstein[1,2]

[1]University of Chicago, Medicine, Chicago, IL, [2]University of Chicago, Molecular Engineering, Chicago, IL

Complex tissues such as lymphatic, nervous, and tumoral systems are inherently three-dimensional (3D), where cells interact to form microenvironments that shape both physiological and pathological processes. To capture these interactions directly within intact tissues, we developed volumetric DNA microscopy, a scalable, optics-free approach that encodes molecular proximities into DNA sequences through *in situ* chemical reactions, followed by standard sequencing. This generates a dense network of molecules that preserve both gene identity and spatial proximity information, from which 3D coordinates and cellular relationships can be reconstructed using geodesic spectral embedding, a manifold-learning framework optimized for large-scale spatial-genomic datasets. Here, we apply volumetric DNA microscopy to whole zebrafish embryos, achieving high-resolution spatial mapping of gene expression within intact developmental contexts. Thus, volumetric DNA microscopy provides a broadly applicable method for large-scale analysis of cell-cell interactions and tissue organization within complex biological systems.

SINGLE-CELL RNA SEQUENCING OF PERIPHERAL IMMUNE CELLS FROM PATIENTS WITH RECURRENT OR INCESSANT PERICARDITIS

Subbaramireddy Remala [1], Tamara Salloum[2], Kevin S Wei[2], Michael Garshick[3], Antonio Abbate[4], Leo F Buckley[1], Brittany N Weber[5]

[1]Brigham and Women's Hospital, Department of Pharmacy, Boston, MA, [2]Brigham and Women's Hospital, Center for Cellular Profiling Core, Boston, MA, [3] NYU Grossman School of Medicine , Department of Medicine, New York, NY, [4]University of Virginia – School of Medicine, Department of Medicine, Charlottesville, VA, [5]UT Southwestern Medical Center , Department of Internal Medicine and Dermatology, Dallas, TX

Introduction: Pericardial inflammation (pericarditis) causes debilitating chest pain and 30-50% of patients experience one or more recurrent episodes despite current standards of care. The specific immune cell responses driving pericardial inflammation remain incompletely understood. Single cell RNA sequencing of peripheral blood mononuclear cells may identify cell type-specific transcriptomic programs responsible for initiating and propagating pericardial inflammation.
Objective: To compare the transcriptomic profile of immune cell subtypes between patients with recurrent or incessant pericarditis and controls.
Methods: We performed single-cell RNA sequencing of peripheral blood mononuclear cells (PBMCs) from patients with (n = 7) and without recurrent or incessant pericarditis (n = 4) using the 10x Genomics platform. After standard preprocessing, doublet removal and integration, we annotated major cell types using standard cellular marker databases. We conducted differential gene expression analysis to compare the transcriptomic profile of cell types between the pericarditis and control groups (log2-fold difference > 1.7 and FDR < 0.05). We performed pathway enrichment analysis of differentially expressed genes in each cell type.
Results: We included 7 patients with recurrent or incessant pericarditis (mean age 51 years, 3 women, 5 on colchicine) and identified 25 distinct cell clusters within the 99,963 cells across all 11 samples. Compared to controls, PBMCs from patients exhibited significantly different cellular variation without changes in composition, except for CD14+ monocytes, which showed significance in both. We identified 33 genes differentially expressed between pericarditis and control across all cell types. CXCR4 was the gene that was differentially expressed in the most cell types (conventional dendritic cells, CD14+ monocytes and NK cells). Pathway enrichment analysis suggested overexpression of genes from the VEGF-VEGFR signaling pathway in conventional dendritic cells, from the TGF-beta receptor signaling pathway in CD14+ monocytes, from the epoetin receptor pathway in NK cells and from the IL-6 pathway in naïve B cells in patients with pericarditis compared to controls.
Conclusion: This study identified CXCR4 as a top differentially expressed gene and overexpression of inflammatory, fibrotic and hypoxic pathways in conventional dendritic cells, CD14+ monocytes and NK cells among patients with pericarditis. Further investigation of these novel, hypothesis-generating findings may identify potential therapeutic targets and biomarkers for the management of pericarditis.

REFERENCE-BASED VARIANT DETECTION WITH VARSEEK

Joseph M Rich[1,2], Laura Luebbert[1,3,4], Delaney K Sullivan[1,5], Reginaldo Rosa[1,2], Lior Pachter[1,6]

[1]California Institute of Technology, Biology and Biological Engineering, Pasadena, CA, [2]Keck School of Medicine of the University of Southern California, Medicine, Los Angeles, CA, [3]Massachusetts Institute of Technology and Harvard University, Broad Institute, Cambridge, MA, [4]Harvard University, Organismic and Evolutionary Biology, Cambridge, MA, [5]David Geffen School of Medicine of the University of California, Los Angeles, Medicine, Los Angeles, CA, [6]California Institute of Technology, Computing and Mathematical Sciences, Pasadena, CA

Variant detection from sequencing data is a cornerstone of genomics, enabling applications from genome-wide association studies to disease diagnosis. Standard de novo variant calling methods, which combine read alignment with statistical modeling, perform well for many tasks but often struggle with indels, low coverage, and single-cell data. We present varseek to overcome these challenges. varseek screens sequencing reads from DNA-, bulk RNA-, or single-cell RNA-Seq against a user-provided database of variants (e.g., COSMIC, ClinVar) by creating and pseudoaligning to an index of Variant-Containing Reference Sequences (VCRSs)—short sequences spanning each variant, flanked so that every k-mer overlaps the variant. To control false positives, varseek applies optimizations including shortening flanks to avoid overlaps, removing low-complexity or genome-matching VCRSs, trimming/adaptor removal for low-quality reads, enforcing gene-consistent mapping, and requiring a minimum count threshold for variant calls.

In benchmarks on synthetic data, varseek maintained >99% sensitivity for both substitutions and indels at ≥3 variant-supporting reads, even at low coverage, outperforming all alternative tools tested. It is also highly efficient, completing variant calling for 256 million reads against 5.4 million variants in just 18 minutes. Validation on real-world bulk RNA-Seq and whole-exome sequencing data from the Geuvadis project further confirmed its high sensitivity and specificity.

We applied varseek to glioblastoma multiforme (GBM) single-cell RNA-Seq data, identifying thousands of tumor-specific COSMIC variants enriched in neoplastic versus healthy astrocytes. Among these, *EGFR*, *C3*, *THBS1*, and *SPP1* harbored cancer-specific variants and showed differential expression between cancer and non-cancer cells, revealing mutation- and cell type-specific patterns within the GBM tumor microenvironment. We characterized individual cells by total COSMIC mutational burden, identifying highly mutated genes in these cells. By combining speed, sensitivity, and technology versatility, varseek provides a versatile framework for high-throughput variant screening in research and clinical genomics.

ALLUVIAL PLOT OPTIMIZATION WITH WOMPWOMP

Joseph M Rich, Conrad G Oakes, Lior Pachter

California Institute of Technology, Division of Biology and biological Engineering, Pasadena, CA

Alluvial plots are widely used for visualizing multivariate categorical data, particularly to reveal relationships across stratified variables. Each observation is mapped to a sequence of category memberships across variables, represented by a flow (or alluvium) connecting vertical strata in adjacent columns. However, two challenges hinder interpretability: (1) the ordering of blocks and flows can obscure structure due to excessive crossings, and (2) inconsistent coloring can make correspondence across variables unclear.

We formalize these as two distinct but related optimization problems. The first, W_POMP (Weighted Permutation Optimization of Multiple Partitions), seeks the variable order and within-layer block orderings that minimize weighted edge crossings, subject to contiguity constraints on partition blocks. The second, W_LOMP (Weighted Label Optimization of Multiple Partitions), seeks color assignments that maximize label agreement across adjacent partitions. Together, these optimizations yield cleaner and more interpretable visualizations.

W_POMP generalizes the NP-hard Weighted One-Layer Free (WOLF) problem to multi-layer graphs without a fixed side, and remains computationally challenging. To address this, we adapt the NeighborNet agglomerative clustering algorithm—originally developed in phylogenetics—to produce high-quality orderings efficiently in typical use cases. We use the sum of products of overlapping edge weights as the objective function to minimize, with a Fenwick tree implementation for reducing runtime from $O(a^2)$ to O(aloga) where a is the number of alluvia. W_LOMP is addressed via hierarchical clustering of blocks by element overlap, enabling consistent color propagation across layers.

We implement both algorithms in a new R package, wompwomp, available via Bioconductor and GitHub. wompwomp achieves substantial improvements over randomized or naive visualizations, both in edge crossing reduction and color coherence. In a visualization challenge involving tissue and clustering labels in single-cell RNA-seq data, our method reduces the edge crossing objective from 182 to just 1. On a complex dataset tracking political affiliations of Game of Thrones characters across 9 timepoints (488 observations, 183 alluvia, and over 12 billion possible block permutations), wompwomp reduces the loss function by more than a factor of 5.

Our formulation provides a principled approach to visualizing structured categorical data. By combining partition-aware sorting with meaningful color alignment, wompwomp enables generation of clear alluvial plots that scale to large datasets with numerous applications in single-cell genomics.

scMTR-SEQ: PROFILING MULTIPLE HISTONE MODIFICATIONS AND RNA IN THE SAME SINGLE CELLS

Jack J Riley, Yang Wang, Peter J Rugg-Gunn

Babraham Institute, Epigenetics Programme, Cambridge, United Kingdom

The combination of histone modifications defines the local chromatin state to regulate gene expression. Understanding how these modifications differ between different cell types and developmental stages provides insight into the mechanisms governing cell identity and cell fate decisions. As such, the ability to simultaneously profile multiple histone modifications alongside the transcriptome is highly desirable, enabling a more direct comparison within the same single cells. Here we apply single-cell multitargets and RNA sequencing (scMTR-seq) to jointly profile the transcriptome alongside up to six histone modifications simultaneously in the same single cells. scMTR-seq utilizes barcoded antibody-proteinA-Tn5 complexes to mediate the tagmentation of genomic regions associated with histone modifications of interest within isolated nuclei. Meanwhile, RNA is captured simultaneously via barcoded poly-dT reverse transcription primers. Subsequently, scMTR-seq utilizes split-pool combinatorial barcoding to confer unique IDs to each cell, and as such, scMTR-seq is highly scalable. scMTR-seq has been successfully applied to complex and heterogeneous cellular systems, including human pluripotent stem cell differentiation, models of early development, and early-stage mouse embryos. By leveraging the paired (transcriptome-epigenome) nature of our data, analysis of the transcriptome is used to define cell states, sub-populations, and transcriptional features such as pseudotime. This, in turn, allows us to compare the chromatin landscapes between these groups using epigenome data from the same cells. Through this approach, we reveal dynamic changes in chromatin states, identify poised (co-marked) regions, and track the direction of their switching. We identify asymmetries in chromatin state between cell lineages and link this to differences in gene expression. Additionally, we identify enhancer-driven gene regulatory networks and compare these between lineages, revealing lineage-specific enhancer usage. To address data sparsity, scMTR-seq data can be aggregated into pseudobulks or used to generate metacells, enabling more robust downstream analyses. Together our results highlight scMTR-seq as a powerful tool to interrogate the chromatin regulatory landscape in a broad range of contexts, allowing us to capture the transcriptome and up to six histone modifications simultaneously in the same single cells. Therefore, scMTR-seq data is highly multimodal and represents a novel source of training data to model epigenome-linked gene regulatory networks with single cell data.

SINGLE-CELL LANDSCAPE OF NORMAL BREAST ORGANOIDS REVEALS LINEAGE-SPECIFIC AND ORGANOID-ENRICHED CELL STATES

Yunus Sahin, James Rouse, Raditya Utama, Sonam Bhatia, David L Spector

Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Understanding the cellular hierarchy and plasticity of the human mammary epithelium is critical for modeling tumorigenesis and lineage disruption. Here, we performed single cell transcriptomic profiling of six patient derived breast organoid lines generated from reduction mammoplasty samples to characterize epithelial diversity in a three dimensional context. Unsupervised clustering identified three major epithelial lineages: basal, luminal hormone responsive (lumHR), and luminal secretory (lumSEC), consistent with the canonical in vivo architecture. Using both established and new markers from our data, we assembled gene panels to define and validate these lineages. Further analysis revealed heterogeneous subpopulations within each lineage, including actively cycling progenitors, differentiated states, and previously undescribed intermediates. We also observed a transcriptional continuum from basal to lumHR, with a “tail” of cells occupying transitional positions in the diffusion graph. These cells resemble the lineage primed intermediates reported by Pal et al. (EMBO J, 2021). This transitional population is consistent with the Human Breast Atlas by Gray et al. (Dev Cell, 2022), which identified basal luminal hybrid cells in normal breast tissue co-expressing KRT5 or KRT14 with luminal markers such as MUC1 and AGR2. All three lineages also contained proliferative subsets. Yet, some states appeared enriched in organoid culture. Our lineage framework is further consistent with the Human Breast Cell Atlas by Kumar et al. (Nature, 2023), which profiled more than 700,000 cells using integrated single cell and spatial transcriptomics. Taken together, these findings show that normal breast organoids recapitulate the in vivo epithelial hierarchy while also revealing transitional and proliferative states that are difficult to detect in tissue based analyses. We are now extending this framework to triple negative and metastatic breast organoids to map tumor associated changes in cell state.

THE IMPACT OF IMPAIRED CHROMATIN 3D STRUCTURE ON EMBRYONIC DEVELOPMENT USING SINGLE-CELL ANALYSES.

Manuel Durán-Acosta[1], Martin Franke[2], José M Santos-Pereira[1,3]

[1]Departamento de Biología Celular, Facultad de Biología, Universidad de Sevilla, Seville, Spain, [2]Centro Andaluz de Biología del Desarrollo (CABD), Consejo Superior de Investigaciones Científicas/Universidad Pablo de Olavide, Seville, Spain, [3]Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville, Spain

During animal embryonic development, precise spatiotemporal regulation of gene expression is required to ensure proper organogenesis and the formation of the body plan. This regulation largely occurs at the transcriptional level through cis-regulatory elements (CREs), such as promoters and enhancers, which physically interact to form loops. These interactions require the genome to be organized in 3D, giving rise to topologically associating domains (TADs), which favor enhancer-promoter contacts while preventing them from forming with incorrect genes (Tena & Santos-Pereira, 2021). The CTCF protein mediates the formation of these contacts by establishing boundaries between TADs and is essential for embryonic development. Its absence results in altered enhancer-promoter interactions and the dysregulation of developmental gene expression (Franke et al., 2021). However, it remains unclear whether all cell types require chromatin 3D structure for differentiation and what the specific impact of CTCF absence is in each embryonic tissue. To address these questions, we have leveraged our ctcf knockout mutant in zebrafish and generated single-cell data during embryonic development. This provides new insights into the link between chromosome organization and gene regulation.

LINKING TRANSCRIPTOMIC PROFILES TO PERTURBATION PHENOTYPES AT SINGLE-CELL RESOLUTION

Leon Schwartz*[1], Ben Kuznets-Speck*[1], Gillian Primavera[1], Ekta Prashnani[3], Carsten Marr[4], Yogesh Goyal[1,2]

[1]Northwestern University, Department of Cell and Developmental Biology, Chicago, IL, [2]Chan-Zuckerberg Biohub Chicago, LLC, Chicago, IL, [3]NVIDIA, NVIDIA Research, Santa Clara, CA, [4]Helmholtz Munich, Institute of AI for Health, Munich, Germany

Mapping genetic features to specific phenotypes has been a central aspect of biology since Mendel's landmark pea plant experiments, through to modern CRISPR genome engineering technology. Recent methods such as Perturb-seq enable massively parallel genotype-to-phenotype screening by introducing thousands of CRISPR-based perturbations in a single experiment, while simultaneously capturing molecular readouts at single-cell resolution. Dozens of published datasets from across cell types and disease states offer a unique opportunity to invert the conventional problem: can a genotype—or in this context, a specific genetic perturbation—be causally inferred solely from the resulting phenotypic response? Mapping phenotype-to-genotype would significantly accelerate mechanistic discovery of basic and applied biology. A central challenge is that experimental context drives a significant degree of transcriptome differences between samples, obscuring true perturbation effects.

To address this challenge, we propose ExPert, a novel contrastive Variational Autoencoder framework that jointly models diverse Perturb-seq datasets across multiple cell types and predicts gene-level effects of CRISPR-based perturbations. Classification is based on latent space similarity to pre-computed gene embeddings, capturing gene-gene relationships. To disentangle context-specific variation from true perturbation signatures, we combine a context-aware contrastive loss with a conditional decoder, yielding a harmonized latent space that factors out confounders while reintroducing cell line-specific context during reconstruction. By integrating 16 datasets and over 17 million cells, we demonstrate that our model accurately predicts target perturbations within trained cell types across more than 3,500 essential genes while preserving biologically meaningful structure in the latent space. In a major breakthrough, our model exhibits strong zero-shot capabilities, generalizing to cell types not included during training and outperforming existing approaches both within and outside the training data. We further observed zero-shot predictions for gene perturbations absent during training, extending inference beyond the limited perturbation coverage of current Perturb-seq experiments. By capturing phenotype-gene relationships independently of cell type context, our approach renders insights from perturbation data transferable across biological systems. Our work opens new opportunities for the discovery of causal genes in fields like targeted therapy, cellular differentiation, or cancer resistance.

INTEGRATED SINGLE-CELL AND SPATIAL PROFILING REVEALS MYELOID-STROMAL NICHES DRIVING STAGE-SPECIFIC IMMUNE REMODELING IN MYCOSIS FUNGOIDES

Shahin Shahsavari[1], Shirin Shahsavari[1], Marissa Mojena[1], Vrinda Madan[1], Daniela Barata[1], Soroosh Solhjoo[2], Courtney Johnson[1]

[1]Johns Hopkins School of Medicine, Department of Dermatology, Baltimore, MD, [2]F. Edward Hébert School of Medicine, Department of Medicine, Bethesda, MD

Introduction:
Mycosis fungoides (MF) is a rare cutaneous T-cell lymphoma characterized by profound immune and stromal heterogeneity. Mechanisms driving disease progression remain poorly defined. Here, we present the first integrated single-cell and spatial atlas of MF across disease stages, resolving cellular states and their spatial niches at unprecedented resolution.
Methods:
We performed single-cell RNA sequencing (scRNA-seq, 23,700 cells) and spatial transcriptomics (10x Visium, 5,600 spots) on six skin samples (healthy control, early-stage MF n=2, late-stage MF n=3).
Results:
We identified diverse malignant and non-malignant T-cell subpopulations, including proliferative, cytotoxic, and exhausted clones. Early lesions displayed robust cytotoxic surveillance ($IFNG^+$, $GZMB^+$, $PRF1^+$ T cells) supported by IFN-activated myeloid cells ($CXCL9^+$, $CXCL10^+$, $STAT1^+$). In contrast, late-stage lesions accumulated TOX^+ $PDCD1^+$ neoplastic T cells enriched for immune checkpoint and exhaustion signatures (PDCD1, CCR4, CD7 loss). Stromal and myeloid compartments were extensively remodeled: $COL11A1^+$ myofibroblastic, $CXCL13^+$ inflammatory, and $HLA\text{-}DRB1^+$ antigen-presenting CAFs transitioned toward fibrotic, immunosuppressive niches enriched in $CD163^+$ M2-like macrophages and $CCR7^+$ $CCL22^+$ migratory DCs. Spatial mapping revealed exhausted T cells co-localized with M2 macrophages and CAFs ($P<0.01$). CellChat uncovered late-stage-specific ligand-receptor interactions (CXCL12/13-CXCR4, MIF-CD74-CD44, TGFB1-TGFBR2, CD40LG-CD40), underscoring emergent stromal-myeloid-tumor crosstalk that suppresses anti-tumor immunity.
Conclusion:
MF progression reflects a shift from diffuse cytotoxic activity in early disease to highly organized stromal–myeloid suppressive niches in advanced lesions. Our findings nominate CXCR4, CD74, and TGFB1 signaling as actionable therapeutic pathways and highlight spatially organized immunosuppressive hubs as candidate biomarkers of disease stage and prognosis. This integrated approach provides a conceptual and analytic framework for precision immunotherapy in cutaneous T-cell lymphoma.

MITOCHONTROL: ADAPTIVE mtRNA FILTERING OF SINGLE CELL RNA SEQUENCING DATA

Caitlin Strassburgh[1], Danielle Pitlor[1], Rachel Gottschalk[2], Shikhar Uttam[1,3]

[1]University of Pittsburgh, Computational and Systems Biology, Pittsburgh, PA, [2]University of Pittsburgh, Immunology, Pittsburgh, PA, [3]University of Pittsburgh, Hillman Cancer Center, Pittsburgh, PA

We propose Mitochontrol, a novel preprocessing method for single cell RNA-sequencing data, which performs adaptive quality control at the cell-type resolution to identify and remove compromised cells.
Single cell RNA-sequencing (scRNA-seq) quantitatively measures the gene expression profiles of individual cells by counting mRNA transcripts. This data is acutely vulnerable to noise resulting from artifact reads and unhealthy cells. Extensive preprocessing is performed to ensure data quality. One critical quality control step evaluates cell quality based on the fraction of mitochondrial transcripts (mtRNA). High mtRNA levels often indicate forms of cellular distress such as apoptosis, but mtRNA levels also fluctuate due to metabolic demands, which are closely tied to cell type and tissue of origin. The current standard for mtRNA filtering is arbitrarily chosen fixed thresholds, applied uniformly to the full sample, which disregards the biological diversity of mtRNA levels and results in the unavoidable inclusion of compromised cells and/or exclusion of intact cells. This presents an urgent need for robust, adaptive mtRNA filtering methods which ensure the observed expression profiles accurately reflect the sample's biological state without bias from compromised cells. Existing adaptive methods have failed to appropriately address these concerns in an unbiased and efficient manner.
Mitochontrol addresses these concerns by selecting adaptive thresholds for each independent cell type using data-driven confidence bounds; these bounds are derived by applying online expectation maximization to assign cells to components of a Gaussian Mixture Model. The removed fraction of cells is validated by comparative pathway enrichment analysis, to confirm the removed cells reflect a 'compromised' functional profile.
We demonstrate that Mitochontrol results in higher retention of intact vs compromised cells and significantly reduced bias in cell type retention rates to create a biologically informed mtRNA quality control protocol which removes compromised cells while respecting the biological diversity of mtRNA levels.

SINGLE-CELL MULTIOMIC PROFILING REVEALS IMMUNE ACTIVATION AND NEURAL DEVELOPMENT PATHWAYS ASSOCIATED WITH ORGANOPHOSPHATE ESTER FLAME RETARDANTS EXPOSURE

Qing Tan[1], Agnieszka D'Antonio-Chronowska[2], Dana Barr[3], Nicholas Mancuso[1,4], Kelly Street[4], Theresa Bastain[5], Carrie Breton[5], Chang Shu[1,6]

[1]USC, Center for Genetic Epidemiology, PPHS, KSOM, LA, CA, [2]UCSD, Center for Epigenomics, La Jolla, CA, [3]Emory University, Department of Environmental Health, RSPH, Atlanta, GA, [4]USC, Division of Biostatistics, PPHS, KSOM, LA, CA, [5]USC, Division of Environmental Health, PPHS, KSOM, LA, CA, [6]USC, Department of Pediatrics, KSOM, LA, CA

Background:
Organophosphate esters (OPEs) are widely used flame retardants and plasticizers prevalent in indoor air, dust and water. While studies suggest exposures to OPEs lead to adverse health effects on maternal and child health, the underlying biological mechanisms remain unclear.
Methods:
We profiled single-cell(sc) multiome of five blood samples from five Hispanic women aged 26 – 40 with varying levels of OPE exposure in the Maternal and Developmental Risks from Environmental and Social Stressors cohort. Concurrent measurement of OPE metabolites, i.e., BDCPP and DPHP, was conducted in urine samples. 29,055 high-quality cells were retained. Integration of scRNA-seq and scATAC-seq was performed by *harmony* and weighted nearest neighbors in *Seurat*. Eleven blood cell types were annotated by transferring reference cell types from OneK1K. Associations between OPE metabolites and scRNA/ATAC-seq were assessed by negative binomial models with adjustment of sample-level random effect and library size. Gene ontology pathway enrichment was performed.
Results:
The expression of four genes were consistently associated with both BDCPP and DPHP metabolites across four cell types (CD4 T cells, effector CD8 T cells, naive CD8 T cells and monocytes) after Bonferroni correction: *ACTB* ($P<4.36\times10^{-}10$; cytoskeleton/neural development), *TMSB4X* ($P<2.21\times10^{-}5$; calcium regulation) and *FTL* ($P<3.44\times10^{-}5$; immune regulation). The ATAC-seq peak in the enhancer region of *CAMKK2* ($P<2.87\times10^{-}6$; calcium signaling/neural development) was also associated with both BDCPP and DPHP metabolites across four blood cell types (CD4 T cells, effector CD8 T cells, memory B cells and monocytes). At the pathway level, cell-type-specific differentially expressed genes significantly associated with BDCPP and DPHP metabolites were enriched in immune activation and cellular trafficking pathways, including platelet activation ($P=5.0\times10^{-3}$) and leukocyte transendothelial migration ($P=2.3\times10^{-2}$) among CD4 T cell.
Conclusions:
Maternal exposure to OPEs is associated with immune activation, cellular trafficking, and neural development across diverse blood cell types, suggesting potential mechanisms through which OPEs perturb maternal immune regulation and neural development pathways.

COMPARATIVE ANALYSIS OF SINGLE-CELL RNA SEQUENCING TECHNOLOGIES FOR COMPLEX TISSUES: INSIGHTS FROM COLORECTAL CANCER & NORMAL ADJACENT TISSUE

Tze Guan Tan[1], Yvonne Teng[2], Lu Zhang[1], Karen Teng[1], Samantha Ho[1], Iain Tan[3], Si-Lin Koo[3], Shuting Han[3]

[1]Merck & Co, Quantitative Biosciences, Singapore, Singapore, [2]Merck & Co, Translational Research, Beijing, China, [3]National Cancer Centre Singapore, Division of Medical Oncology, Singapore

The rapid evolution of single-cell RNA sequencing (scRNA-seq) methodologies has significantly advanced our understanding of cellular diversity across tissue types and disease states. However, systematic assessments of scRNA-seq technologies for complex tissues remain lacking. Here we provide a comprehensive evaluation of three scRNA-seq platforms, including emulsion-based and microwell-based methods, using human colorectal cancer and normal adjacent colon tissues. Assessing key metrics at library, gene, and cell levels, we found that the Chromium GEM-X technology by 10x Genomics exhibited the best overall performance, particularly in transcript recovery, neutrophil detection, and ambient RNA levels. Our results provide critical insights that will assist researchers in selecting the most appropriate scRNA-seq method, in particular for studies involving clinical cancer specimens.

MAPPING THE ISR LANDSCAPE IN COGNITIVE DISORDERS VIA SINGLE-CELL MULTI-OMICS

Kristof Torkenczy[1], Lucas C Reineke[1], Sean W Dooling[1], Benjamin Henderson[2], Daniel N Itzhak[1], Benjamin Yang[1], Dongze He[1], Richard M Myers[2], Peter Walter[1], Stefka Tyanova[1], Mauro Costa-Mattioli[1,3]

[1]Altos Labs, Inc., Bay Area Institute, Redwood City, CA, [2]HudsonAlpha , Institute for Biotechnology, Huntsville, AL, [3]Baylor College of Medicine, Department of Neuroscience, Houston, TX

Persistent activation of the integrated stress response (ISR) is a major driver of cognitive decline in both neurodevelopmental and neurodegenerative disorders. Using a new mouse model (Ppp1r15b-R658C mice) that mimics the persistent ISR activation and cognitive decline observed in humans, we generated the first single-cell ISR atlas of the brain. By integrating single-cell RNA-seq and single-cell ATAC-seq with proteomics, we discovered that distinct brain cell types respond differently to persistent ISR activation and elicit cell-type-specific ISR programs. Interestingly, chromatin accessibility analyses revealed that the ISR downstream factor ATF4 is a key ISR effector in GABAergic neurons, while AP-1 (JUNB) is implicated in glutamatergic neurons. More importantly, selective deletion of ATF4 in GABAergic neurons—but not in glutamatergic neurons—impacts ISR-mediated cognitive decline in Ppp1r15b-R658C mice, demonstrating that different neuronal subtypes rely on unique ISR downstream effectors to regulate mnemonic processes. Furthermore, we defined a comprehensive molecular signature of persistent ISR activation, which we showed could serve as a biomarker for cognitive dysfunction across neurodevelopmental, neurodegenerative disorders and normal aging. This multi-omic framework provides a key platform for exploring and validating new scientific hypotheses, significantly advancing our understanding of ISR-related brain disorders.

SINGLE-CELL TRANSCRIPTION FACTOR PERTURBATIONS REVEAL GENE MODULE ARCHITECTURE DURING hPSC DIFFERENTIATION INTO DEFINITIVE ENDODERM

Nan Zhang*[1], Daniel Cizin*[2], Denis Torre*[1], Tingfeng Guo[1], Thomas Norman#[2], Danwei Huangfu#[1]

[1]Memorial Sloan Kettering Cancer Center, Developmental Biology Program, Sloan Kettering Institute, New York, NY, [2]Memorial Sloan Kettering Cancer Center, Computational and Systems Biology Program, Sloan Kettering Institute, New York, NY

* Contributed equally
# Corresponding authors

Transcription factors (TFs) are master regulators of cell identity. However, their role in shaping cell state during human embryogenesis has not been systematically interrogated. Here, we developed a scalable pipeline that integrates Perturb-seq with sgRNA multiplexing and module-level regression analysis to interrogate the function of 1,983 TFs and epigenetic regulators during differentiation of human pluripotent stem cells (hPSCs) into definitive endoderm (DE), a critical early step in the formation of the respiratory, endocrine and digestive systems. Through analysis of the resulting single-cell data, we define gene modules that recapitulate key dimensions of cell state, capturing regulatory networks underlying differentiation, cell cycle control, signaling and metabolism across both cell types. Integrating perturbation data in the context of such modules reveals how TF silencing reshapes transcriptional programs, providing a quantitative framework to understand how TFs contribute to cell state transitions. Mechanistically, we confirm the known role of SOX17 and FOXH1 in regulating endodermal development, and nominate candidate genes including SOX11, SOX4 and GATAD1 for further experimental investigation. Leveraging this system, we further illustrate how combinations of gene modules can be used to define phenotypic landscapes and map perturbation-induced shifts in cell state, demonstrating a generalizable strategy to dissect regulatory programs across cellular contexts.

SINGLE-CELL RNA SEQUENCING OF MURINE LIVER REVEALS AN ALIGNED CIRCADIAN CLOCK AND CELL-POPULATION SPECIFIC CIRCADIAN REGULATED PATHWAYS

Anthony J Veltri*[1], Manabu Nukaya*[2], Ksenija Korac[2], Patrick B Schwartz[2], Sean M Ronnekleiv-Kelly[2,3]

[1]Bioinformatics Resource Center, University of Wisconsin Biotechnology Center, Madison, WI, [2]Department of Surgery, Division of Surgical Oncology, University of Wisconsin School of Medicine and Public Health, Madison, WI, [3]University of Wisconsin Carbone Cancer Center, University of Wisconsin School of Medicine and Public Health, Madison, WI

*Authors contributed equally

The circadian clock is tightly connected to metabolism, which is evident in various metabolic processes performed by the liver. Perturbation of these processes due to circadian dysregulation leads to liver specific pathology. The liver is composed of multiple different cell populations each with distinct functions contributing to organ homeostasis, but individual cell population contributions to circadian clock function is not yet known. Single-cell RNA sequencing provides the opportunity to understand clock function and oscillating gene expression within an organ system at the individual cell population level that would allow for better understanding of the crosstalk between the circadian clock and metabolic pathways within the liver. In the past, barriers to achieving this goal included complexity associated with generating single-cell RNA sequencing time series data as well as the complexity of data analysis. Here, we established a protocol that enabled the generation of murine liver cell population time series data, as well as a methodological approach to evaluate the core molecular clock and oscillating gene expression in individual cell populations. Using a combination of normalized coefficient of variation, clock-correlation and aggregate pseudobulk, we found a robust and aligned circadian clock in each of the cell populations. We then employed a pseudoreplicate / pseudobulk strategy to identify oscillating gene expression and benchmarked against bulk RNA sequencing data; we demonstrated that many metabolic genes were oscillating in several of the cell populations, including non-hepatocyte clusters. Finally, we identified oscillating genes unique to specific cell populations that play critical roles in liver function. The findings in this study lay an important foundation for understanding clock function and contributions of oscillating gene function at the individual cell population level in liver.

MAPPING CROSSOVER EVENTS OF MOUSE MEIOTIC RECOMBINATION BY RESTRICTION FRAGMENT LIGATION-BASED REFRESH-SEQ

Yan Wang*, Yijun Chen*, Junpeng Gao, Jun'e Liu, Fuchou Tang

Peking University, School of Life Sciences, Beijing, China

Single-cell whole-genome sequencing methods have undergone great improvements over the past decade. However, allele dropout, which means the inability to detect both alleles simultaneously in an individual diploid cell, largely restricts the application of these methods particularly for medical applications. Here, we develop a new single-cell whole-genome sequencing method based on third-generation sequencing (TGS) platform named Refresh-seq (restriction fragment ligation-based genome amplification and TGS). It is based on restriction endonuclease cutting and ligation strategy in which two alleles in an individual cell can be cut into equal fragments and tend to be amplified simultaneously. As a new single-cell long-read genome sequencing method, Refresh-seq features much lower allele dropout rate compared with SMOOTH-seq. Furthermore, we apply Refresh-seq to 688 sperm cells and 272 female haploid cells (secondary polar bodies and parthenogenetic oocytes) from F1 hybrid mice. We acquire high-resolution genetic map of mouse meiosis recombination at low sequencing depth and reveal the sexual dimorphism in meiotic crossovers. We also phase the structure variations (deletions and insertions) in sperm cells and female haploid cells with high precision. Refresh-seq shows great performance in screening aneuploid sperm cells and oocytes due to the low allele dropout rate and has great potential for medical applications such as preimplantation genetic diagnosis.

TISSUE-SPECIFIC CELL ATLAS OF *DAPHNIA*: TRACING EVOLUTIONARY SHIFTS IN SEX DETERMINATION

Wen Wei, Man Lin, Michael Lynch

Arizona State University, Biodesign Center for Mechanisms of Evolution, Tempe, AZ

A central question in evolutionary biology is how reproductive systems diversify and persist across lineages. In animals, one of the most frequent transitions involves shifts between environmental sex determination (ESD), in which external cues such as temperature or photoperiod bias sex, and genetic sex determination (GSD), in which sex is specified by chromosomal factors. Pancrustaceans span the full spectrum of these modes, including ESD, GSD, and intermediate reproductive strategies such as hermaphroditism and gynodioecy, and thus provide an informative system for dissecting the cellular and molecular foundations of sex-determination evolution.

Using *Daphnia*, a classical ESD model, we constructed a tissue-resolved single-cell atlas by sequencing >1.2 million cells from 54 tissues sampled from males and females across six species spanning ~150 million years of divergence. This resource substantially exceeds existing pancrustacean atlases in tissue coverage, cell number, and phylogenetic depth, and thereby provides a foundation for comparative cell-type analyses.

Cross-species comparisons revealed both conserved and lineage-specific cell types within pancrustaceans. For example, germline cells are universally present, but their transcriptional states diverge: ribosomal proteins are consistently upregulated in the germlines of GSD species relative to *Daphnia*. Even more striking is the absence of follicle cells in *Daphnia*. In GSD taxa, follicle cells surround oocytes, mediating responses to growth and sex hormones. Their absence in *Daphnia* likely reflects a distinct hormonal regime: in ESD species, sex is established during embryogenesis by levels of juvenile hormone and 20-hydroxyecdysone, whereas in GSD species, hormones function primarily in gamete development, not sex specification.

By linking cell-type variation to differences in reproductive mode, this work provides a mechanistic view of how changes in hormonal signaling and cellular composition shape transitions in sex determination. More broadly, the *Daphnia* atlas contributes to the growing repertoire of tissue-specific single-cell resources, offering a comparative framework for investigating how cellular and regulatory changes underlie the long-term evolution of reproductive systems.

EPIGENETIC AND SPATIAL REMODELING SHAPE LYMPHOMA TUMOR MICROENVIRONMENT

Yulong Wei[1], Paige M Bracci[2], Michael S McGrath[3], Ya-Chi Ho[1]

[1]Yale University School of Medicine, Department of Microbial Pathogenesis, New Haven, CT, [2]The University of California at San Francisco, Department of Epidemiology and Biostatistics, San Francisco, CA, [3]The University of California at San Francisco, Department of Medicine, San Francisco, CA

Non Hodgkin lymphoma (NHL) ranks among ten most frequently diagnosed cancers in the United States. Most NHL cases are B cell malignancies that accumulate epigenetic lesions and aberrant transcription factor activities (IRF4, FOXP1, BCL6, FOXO1, and MYC) that rewire B cell gene expression programs. Within the tumor microenvironment (TME), spatial organization of immune subsets further shapes malignant B cell fate. Concomitant viral infections, particularly HIV-1 and EBV, promote malignant B cell survival and immune escape. We hypothesized that epigenetic reprogramming in the TME drives dysfunctional immune states and creates immune sanctuaries permissive to lymphoma survival, and viral infections potentiate TME dysfunction. Using snap-frozen lymphoma samples from 12 participants with B cell lymphoma (6 HIV-1+ and 6 HIV-1−; including 5 ABC- and 3 GCB-DLBCL, 2 follicular lymphoma, and 2 Burkitt lymphoma clinical diagnoses) and non-malignant lymph nodes from 17 donors (9 HIV-1+ and 8 HIV-1−), we performed single-cell 10x Multiome and simultaneously captured chromatin accessibility and transcriptome in 415,548 cells, including 230,828 B cells, 32,417 CD8+ T cells, and 36,810 CD5+ T cells, In addition, we identified 164 HIV-1+ CD4+ T cells from HIV-1+ donors. From the same tissues, we performed 10x Xenium In Situ profiling (380 Immuno-Oncology panel plus 100 custom gene targets, including HIV and EBV probes) to map spatial organization within the TME. Integrating single-cell Hierarchical Poisson Factorization, Weighted Gene Co-expression Network Analysis, and spatial immune profiling revealed lymphoma-subtype-specific immune programs. Notably, high MYC activity in Burkitt lymphoma suppressed VEGFA and HIF1a expression and drove upregulation of glycolysis and stress-adaptation genes (PFKP, ENO1, GPI, LDHA, LDHB, TXNRD1, CALM2, PARGC1B, ACACA, ACSL1, AFMID), consistent with a Warburg-like glycolytic program and oxidative-stress tolerance instead of VEGF-mediated angiogenesis. Single-cell and spatial immune profiling further revealed enrichment of HIV-1+ CD4+ T cells in Type 1 regulatory T cells (Tr1), enrichment of cytotoxic CD8+ T cells, and upregulated antiviral response, proliferation, and exhaustion programs in immune cells within the TME of HIV-1+ donors across lymphoma subtypes. Our multiomic atlas links epigenetic and spatial remodeling to uncover key drivers of immune dysfunction in infection-associated and idiopathic lymphomas.

scIsoSim: SIMULATING SINGLE-CELL RNA-SEQ READS WITH FULL-LENGTH ISOFORM GROUND TRUTHS

Guanao Yan[1,2], Weijian Wang[3], Zhiyin Liu[3], Jingyi Jessica Li[1,3,4]

[1]University of California, Los Angeles, Department of Statistics, Los Angeles, CA, [2]Michigan State University, Department of Computational Mathematics Science and Engineering, East Lansing, MI, [3]University of California, Los Angeles, Bioinformatics Interdepartmental Ph.D. Program, Los Angeles, CA, [4]Fred Hutchinson Cancer Center, Biostatistics Program, Seattle, WA

The advancement of single-cell RNA sequencing (scRNA-seq) has enabled the investigation of isoform expression and alternative splicing (AS) at single-cell resolution. However, benchmarking computational tools for these tasks remains challenging due to the lack of full-length isoform information in experimental data and the inability of existing simulators to generate scRNA-seq reads with ground-truth isoform structures and proportions. To address this gap, we present scIsoSim, the first simulator capable of generating realistic scRNA-seq reads from isoforms with predefined structures and proportions. scIsoSim is a flexible simulator that supports both plate-based and droplet-based technologies, including Smart-seq2 and 10x Genomics Chromium 3' and 5' protocols. It also realistically mimics read sequences and splicing observed in real scRNA-seq data. Using scIsoSim, we performed a comprehensive benchmark of isoform and AS quantification tools on scRNA-seq datasets. First, we assessed the applicability of isoform quantification tools originally developed for bulk RNA-seq to Smart-seq2 data, finding that all tested tools achieved high accuracy, with Salmon demonstrating the best computational efficiency. Second, we evaluated Scasa, the only existing isoform quantification tool for 10x 3' scRNA-seq, and found its accuracy limited due to data sparsity. Finally, we benchmarked AS quantification tools and observed that while brie outperformed outrigger in overall accuracy, it does not always capture cell-specific AS events. Overall, scIsoSim provides a valuable resource for benchmarking isoform and AS quantification methods in scRNA-seq analysis.

A NOVEL METHOD TO IMPROVE SIGNAL-TO-NOISE FOR PEAK-CALLING IN scATAC-SEQ

Shiyi Yang[1], Kevin Y An[2], Jase Gehring[3], Sina Booeshaghi[1], Aaron Streets[3]

[1]UC Berkeley, Center for Computational Biology, Berkeley, CA, [2]UC Berkeley, Quantitative Biosciences, Berkeley, CA, [3]UC Berkeley, Bioengineering, Berkeley, CA

Single-cell Assay for Transposase Accessible Chromatin with sequencing (scATAC-seq) has become a widely used method for investigating chromatin accessibility at single-cell resolution. The ensemble landscape of accessible chromatin for a given experiment is discovered through computational peak-calling algorithms, and these peak coordinates which represent regions of accessible chromatin are then used as genomic features for downstream, single-cell analysis. Due to the intrinsic low read depth property of each single-cell in the assay, peak calling algorithms estimate peaks by pseudo-bulking the DNA fragment reads from all single-cells. However, background noise, corresponding to non-specific reads from empty droplets or overtagmented cell droplets can diminish peak-calling performance by obscuring open regions with spurious DNA fragments or systematically decreasing the signal to noise ratio. Here, we present a method to improve signal to noise in scATAC-seq analysis by filtering low-quality barcodes prior to peak-calling. To achieve this, we measure transposase tagmentation complexity by calculating entropy to remove barcodes corresponding to empty droplets and overtagmented cells. With this method, we were able to discover 39% more peaks with entropy based barcode pre-filtering. We found downstream analysis using peaks after entropy filtering has better cell cluster consistency with their RNA counterpart in multi-omic data.

THE CELL LANDSCAPE OF DENTAL PULP POLYP

Wei Yin

Wuhan University, School & Hospital of Stomatology, Wuhan, China

Objective:
Dental pulp polyp is a granulomatous lesion formed by abnormal proliferation of dental pulp tissue after long-term inflammatory stimulation. It is commonly found in young permanent or deciduous teeth with deep caries that have not been treated in a timely manner. The formation mechanism and cellular composition of dental pulp polyp is still not fully understood. This study aims to draw a cell map of dental pulp polyp through single-cell sequencing.

Methods:
The dental pulp polyp sample was obtained from a right mandibular first molar of an 18-year-old patient. With the patient's consent, polyp sample was collected for single-cell sequencing using 10X Genomics platform.

Results:
A total of 2266 cells were obtained from the sequencing of this dental pulp polyp sample. According to the reported markers for various types of cells, the cells in the sample were mainly immune cells and endothelial cells. Immune cells include macrophages, T cells, and B cells. The obvious cellular heterogeneity was a significant characteristic of macrophages in this dental pulp polyp. Analysis of cellular function revealed that macrophages not only had anti-inflammatory effects, but also promoted angiogenesis.

Conclusion:
This study depicted the cellular composition of dental pulp polyp and provided a preliminary analysis of the functions of the main cells.

PREDICTING TUMOR MICROENVIRONMENT CHANGES INDUCED BY CHEMORADIOTHERAPY IN PANCREATIC DUCTAL ADENOCARCINOMA USING MACHINE LEARNING ON SPATIAL SINGLE-CELL PROTEIN PROFILING AND CELL TYPE COMPOSITION

Sergio Zamora-Erazo[1], Caneta Brown[2], Edna Cukierman[2], Janusz Franco-Barraza[2], Hayan Lee[1,3,4]

[1]Cancer Epigenetics Institute, Nuclear Dynamics and Cancer Program, Fox Chase Cancer Center, Temple University, Philadelphia, PA, [2]M&C Greenberg Pancreatic Cancer Institute, Fox Chase Cancer Center/Lewis Katz School of Medicine, Temple University, Philadelphia, PA, [3]Department of Cancer and Cellular Biology, Lewis Katz School of Medicine, Temple University, Philadelphia, PA, [4]Department of Medical Genetics and Molecular Biochemistry, Lewis Katz School of Medicine, Temple University, Philadelphia, PA

Pancreatic ductal adenocarcinoma (PDAC), the most common and aggressive form of pancreatic cancer, is characterized by a poor prognosis and therapeutic resistance. Tumor microenvironment (TME) in PDAC plays a critical role in promoting tumor progression, immune evasion, and resistance to therapy, making it a key target for improving treatment outcomes. The latest technological advances have allowed protein abundance to be measured at single-cell resolution. Further, Lunaphore COMET™ technology enables spatial mapping of the proteome at single-cell resolution, providing crucial insights into the cellular interactions within the TME and functional states of immune and cancer-associated fibroblasts (CAFs) in PDAC.

We collected 32 biospecimens from a surgical cohort of PDAC patients, including 8 treatment naïve (TN) patients and 8 treated with total-neoadjuvant therapy including chemotherapy and conventional radiotherapy (TNT-CRT). We developed a computational workflow to cluster TMEs and CAFs, allowing us to identify sub-TME populations and sub-CAF characteristics, exploiting Logistic Regression (LR) and Random Forest (RF) algorithms.

Our ensemble approach showed that LR trained with the proportion of 29 protein markers in cells was the best option for sample classification, with a contribution of RF trained with the proportion of sub-CAFs grouped by 15 protein markers. The models' performance suggests that CD56, a classic marker for NK cells, predicts TNT-CRT-treated samples, while CD68 and FOXP3, markers for tumor-associated macrophages and T regulatory cells, may be associated with TN TMEs.

Our study identifies spatial proteomic biomarkers capable of predicting TMEs in TN and TNT-CRT–treated PDAC patients, as well as associated treatment outcomes. Predicting a patient's response to therapy is a critical component of clinical decision-making, and our findings may help guide treatment strategies to improve success rates. In the future, we will characterize tumor-resistant and tumor-supportive TMEs to further understand mechanisms of treatment response and resistance.

BARTsc: MOTIF-INDEPENDENT IDENTIFICATION OF FUNCTIONAL TRANSCRIPTION FACTORS FROM SINGLE-CELL MULTIOMICS DATA

Hongpan Zhang[1,2], Jingyi Wang[1], Zhenjia Wang[1], Chongzhi Zang[1,2,3,4]

[1]University of Virginia, Department of Genome Sciences, Charlottesville, VA, [2]University of Virginia, Department of Biochemistry and Molecular Genetics, Charlottesville, VA, [3]University of Virginia, Department of Biomedical Engineering, Charlottesville, VA, [4]University of Virginia, UVA Comprehensive Cancer Center, Charlottesville, VA

Transcription factors (TRs), including transcription factors (TFs) and chromatin regulators, play crucial roles in transcriptional regulation and are implicated in numerous cellular processes, including cancer progression. Advances in single-cell genomics techniques, particularly single-cell multiome (RNA+ATAC), provide new opportunities for inferring functional TRs at a cell type level in silico. However, most existing methods employ conventional approaches that are based on gene-proximal motif enrichment and co-expression between regulator and target gene, which hold inherent limitations. Here, we present BARTsc, a computational method that accurately predicts functional TRs by inferring cis-regulatory profiles from differential genomic features and accessing their association with a large collection of TR ChIP-seq profiles. BARTsc introduces a novel cell-group-based framework that identifies TRs shaping cell group signatures, analyzes relative TR activity across cell groups and predicts key regulators for each cell group within a given dataset. BARTsc implements a flexible approach that handles scRNA-seq and scATAC-seq data, while accommodating single-cell multiome data by inferring consensus cis-regulatory profiles through the integration of both modalities. We validated BARTsc on diverse datasets, including mouse cortex, human peripheral blood mononuclear cells (PBMCs), and pancreatic ductal adenocarcinoma (PDAC). In each case, BARTsc successfully identified known lineage-defining TRs and key regulators of cell identity. We demonstrate that integrating paired transcriptomic and chromatin accessibility data significantly enhances predictive accuracy. Systematic benchmarking reveals that BARTsc consistently outperforms existing state-of-the-art methods in various aspects. In the PDAC dataset, BARTsc pinpointed critical regulators and pathways driving tumor development, highlighting the potential of BARTsc for uncovering candidate therapeutic targets. As a robust and versatile tool, BARTsc enables deeper insights into cell-type-specific regulatory programs and facilitates the discovery of key regulators in development, disease, and cellular reprogramming.

DISCOVERING CELL TYPES AND STATES FROM REFERENCE ATLASES WITH HETEROGENEOUS SINGLE-CELL ATAC-SEQ FEATURES

Yuqi Cheng, Xiuwei Zhang

Georgia Institute of Technology, School of Computational Science and Engineering, Atlanta, GA

Despite substantial recent advances in query mapping and cell type or cell state discovery tools, their application to single-cell assay for transposase-accessible chromatin using sequencing (scATAC-seq) data remains challenging. The heterogeneous nature of peak feature spaces across samples hinders the effectiveness of existing methods, while the absence of dedicated tools for detecting perturbed cell types and states in scATAC-seq data further limits the depth of downstream analyses. To address these limitations, we present EpiPack, an integrative computational toolkit that leverages heterogeneous transfer learning and graph-based modeling strategies to advance scATAC-seq analysis. At its core, the Peak Embedding Informed Variational Inference (PEIVI) framework within EpiPack enhances mappable reference construction, query mapping, and label transfer, demonstrating that leveraging heterogeneous features in scATAC-seq data outperforms methods relying solely on conventional homogeneous features. In addition, EpiPack's global-local out-of-reference (OOR) detection framework achieves robust and efficient detection of perturbed cell types and states, extending the utility of scATAC-seq to disease and perturbation contexts. With its modular design and transferable pre-trained references, EpiPack can be readily applied to diverse analytical tasks and is available as a Python package at https://github.com/ZhangLabGT/EpiPack.

scREBOUND: AN EFFICIENT DESIGN OF SINGLE-CELL FOUNDATION MODEL WITH BATCH REPRESENTATION

Ziqi Zhang, Chandra Sekhar Reddy Edula, Xiuwei Zhang

Georgia Institute of Technology, School of Computational Science and Engineering, Atlanta, GA

Recent advances in single-cell foundation models (scFMs) have demonstrated the promise of large-scale pretraining on single-cell RNA sequencing (scRNA-seq) data for a wide range of downstream biological tasks. However, existing models such as scGPT, UCE, scFoundation, and scMulan demand substantial computational resources, limiting their accessibility and practical deployment in academic settings. Furthermore, the systematic technical noise within different experimental batches of scRNA-seq datasets (batch effect) cannot be well removed with the masked token prediction tasks that are commonly used by scFMs. This significantly jeopardizes the zero-shot performance of these models.

In this work, we tackle both challenges through a new design of scFM, named scREBOUND. First, we introduce a protein-language-model-informed gene compression network that reduces the total number of gene tokens per cell while preserving key biological information. The compression network groups genes with similar functionality, informed by their protein embeddings from a protein language model, and is fine-tuned during training to better retain the expression information in the original data. Since scFMs are built based on the transformer architecture, reducing the number of gene tokens reduces the computation time of the transformer quadratically. This compression significantly accelerates the speed of scREBOUND and reduces GPU memory usage. Second, we address batch effects by incorporating a batch embedding network and a multi-granular contrastive loss as regularization during model pretraining. The batch embedding network learns representations from batch-specific features summarized from expression data of certain categories of genes and can generalize to unseen batches. The multi-granular contrastive loss aligns cells of common cell types across batches and is designed to accommodate mismatches in label granularity across different experiments.

We evaluate our model on multiple scRNA-seq datasets, under four downstream tasks: (1) Batch effect removal; (2) Cell type annotation; (3) Gene expression imputation; (4) Cross-species knowledge transfer: the model that is trained solely on human dataset can infer the cell embedding of data from other species (mouse, monkey, etc). We compared scREBOUND with state-of-the-art (SOTA) scFMs in a zero-shot manner. Our results demonstrate that scREBOUND achieves comparable or superior accuracy across all four tasks, while achieving up to 17x reduction in inference time and 30x reduction of memory usage compared to the SOTA model scGPT. This work paves the way for enabling wider adoption of scFMs in academic research and improving the robustness of single-cell analyses across heterogeneous datasets.

MANIFOLD LEARNING REVEALS CELL CYCLE PLASTICITY UNDERLYING FRACTIONAL RESISTANCE TO PALBOCICLIB IN ER+/HER2− BREAST TUMOR CELLS

Tarek M Zikry[1,2,5], Samuel C Wolff[3,5], Jolene S Ranek[3,5], Michael R Kosorok[2], Philip M Spanheimer[4], Jeremy E Purvis[3,5]

[1]UNC-Chapel Hill, School of Data Science and Society, Chapel Hill, NC, [2]UNC-Chapel Hill, Department of Biostatistics, Chapel Hill, NC, [3]UNC-Chapel Hill, Department of Genetics, Chapel Hill, NC, [4]UNC-Chapel Hill, Department of Surgery, Chapel Hill, NC, [5]UNC-Chapel Hill, Computational Medicine Program, Chapel Hill, NC

The CDK4/6 inhibitor palbociclib blocks cell cycle progression in Estrogen receptor–positive, human epidermal growth factor 2 receptor–negative (ER+/HER2−) breast tumor cells. Despite the drug's success in improving patient outcomes, a small percentage of tumor cells continues to divide in the presence of palbociclib—a phenomenon we refer to as fractional resistance. It is critical to understand the cellular mechanisms underlying fractional resistance because the precise percentage of resistant cells in patient tissue is a strong predictor of clinical outcomes. Here, we hypothesize that fractional resistance arises from cell-to-cell differences in core cell cycle regulators that allow a subset of cells to escape CDK4/6 inhibitor therapy. We used multiplex, single-cell imaging to identify fractionally resistant cells in both cultured and primary breast tumor samples resected from patients. Resistant cells showed premature accumulation of multiple G1 regulators including E2F1, retinoblastoma protein, and CDK2, as well as enhanced sensitivity to pharmacological inhibition of CDK2 activity. Using trajectory inference approaches, we show how plasticity among cell cycle regulators gives rise to alternate cell cycle "paths" that allow individual tumor cells to escape palbociclib treatment. In additional work, we posit that spherical manifold approximations represent these single-cell populations, suggesting that significant changes in latent lower dimensional manifold structures correspond to distinct cell cycle behaviors. Leveraging an existing manifold approximation method, we fit single-cell data to a hypersphere and establish an empirical hypothesis testing framework to quantify differences in these spheres across conditions. Our model-agnostic approach enables the direct quantification of the effect of single-cell perturbations, treatments, or differences between patient tumors, revealing cellular behaviors in a novel paradigm. Understanding drivers of cell cycle plasticity, and how to eliminate resistant cell cycle paths, could lead to improved cancer therapies targeting fractionally resistant cells to improve patient outcomes.

RECONSTRUCTING CHROMATIN DYNAMICS WITH MPATH: A LABEL-FREE APPROACH FOR TEMPORAL EPIGENOME PROFILING VIA LONG-READ METHYLATION

Annie Trinh[1], Navied Akhtar[2], Kwadwo Bonsu[3], Nandor Laszik[2], Asia Mendelevich[4], Tanye Wen[2], Julien Morival[2], Katelyn E Diune[2], Mitchell Frazeur[2], Justin E Vega[2], Alexander A Gimelbrant[4], Elizabeth L Read[3], Timothy L Downing[1,2]

[1]University of California, Irvine, Microbiology and Molecular Genetics, Irvine, CA, [2]University of California, Irvine, Biomedical Engineering, Irvine, CA, [3]University of California, Irvine, Chemical and Biomolecular Engineering, Irvine, CA, [4]Altius Institute for Biomedical Sciences, Seattle, WA

Faithful propagation of epigenetic states is essential for maintaining cell identity, yet the temporal coordination of chromatin modifications during DNA replication remains difficult to resolve. Existing strategies to track epigenetic inheritance often rely on exogenous labeling or chemical conversion steps that can perturb native chromatin and limit temporal precision.

We developed Methylation Pseudotime Analysis Through read-level Heterogeneity (MPATH), a label- and conversion-free computational framework that infers post-replication strand maturity directly from single-molecule methylation profiles. By leveraging intrinsic heterogeneity captured by long-read sequencing, MPATH enables reconstruction of the relative timing of methylation and hydroxymethylation events across individual DNA molecules.

Applying MPATH to long-read NOMe-seq data allowed for integration with nucleosome occupancy information, revealing relationships between DNA strand maturity and local chromatin accessibility, while phased analyses hinted at allele-specific pseudotime trends linked to X-chromosome activity.

Together, these results highlight the potential of MPATH to extract temporal information from static methylation data without external perturbation. More broadly, they point toward new opportunities to explore how chromatin structure and DNA modification states are re-established following replication, opening a window into the dynamic interplay between genome architecture and epigenetic memory at single-molecule resolution

NOTES

NOTES

NOTES

NOTES

NOTES

Participant List

Dr. Ayush Aggarwal
University of California San Francisco
ayush.aggarwal@ucsf.edu

Dr. Kadir Akdemir
The University of TX MD Anderson Cancer Center
kcAkdemir@mdanderson.org

Ms. Ekaterina Aladyeva
The Ohio State University
ekaterina.aladyeva@osumc.edu

Dr. Bremy Alburquerque
Icahn School of Medicine at Mount Sinai
bremy.alburquerque@mssm.edu

Mr. Kevin An
University of California, Berkeley
kevinan@berkeley.edu

Rebecca Anderson
Cold Spring Harbor Laboratory
randerso@cshl.edu

Dr. Jianpeng Ao
Boston University
jpao@bu.edu

Kevin Aris
University of California, Berkeley
kevin_aris@berkeley.edu

Dr. Daniel Armendariz
Yale University
Daniel.Armendariz@yale.edu

Kanishk Asthana
University of California San Diego
kasthana@ucsd.edu

Petra Bachanova
Dartmouth College
petra.bachanova.gr@dartmouth.edu

Dr. Akilandeswari Balasubramanian
Chan Zuckerberg Biohub San Francisco
akilandeswari.b@czbiohub.org

Dr. Douglas Barrows
Rockefeller University
dbarrows@rockefeller.edu

Dr. Tobias Beyer
Cytosurge
tobias.beyer@cytosurge.com

Mr. Nischal Bhandari
Cold Spring Harbor Laboratory
nbhandar@cshl.edu

Mr. Timothy Bi
MD Anderson Cancer Center
tmbi@mdanderson.org

Dr. Paul Blainey
The Broad Institute of MIT and Harvard
pblainey@broadinstitute.org

Madison Blanton
University of Kentucky
mbbl227@uky.edu

Elise Boisvert
Cornell University
eab75@cornell.edu

Dr. A. Sina Booeshaghi
UC Berkeley
sinab@berkeley.edu

Dr. Michael Bronson
10x Genomics
michael.bronson@10xgenomics.com

Prof. Liz Brunk
UNC Chapel Hill
elizabeth_brunk@med.unc.edu

Meghana Budankayala
Cleveland Clinic
budankm@ccf.org

Prof. Long Cai
California Institute of Technology
lcai@caltech.edu

Dr. Siu Chiu Chan
Stony Brook University
siuchiu.chan@stonybrookmedicine.edu

Kevin Chang
University of Southern California
kchang42@usc.edu

Dr. Laura Chen
Broad Institute
ldoherty@broadinstitute.org

Ms. Shihui Chen
University of Southern California
shihuich@usc.edu

Dr. Lan Chen
MSD
lan.chen@msd.com

Dr. Jian Chen
CHDI Management, Inc.
jian.chen@chdifoundation.org

Mr. Xiaofeng Chen
Virginia Tech
xiaofengchen@vt.edu

Ms. Gayeon Cho
POSTECH
rkdus9977@postech.ac.kr

Ms. Urshila Choubal
Georgetown University
uvc2@georgetown.edu

Gulam Chuwdhury
Iowa State University
sarwar@iastate.edu

Dr. Paolo Cifani
Cold Spring Harbor Laboratory
cifani@cshl.edu

Mr. Rodrigo Cotrim Chaves
University of California, Berkeley
rchaves@berkeley.edu

Prof. Mehdi Damaghi
Stony Brook University
mehdi.damaghi@stonybrookmedicine.edu

Gargi Damle
Icahn School of Medicine
gargi.damle@mssm.edu

Mr. Ojonugwa David
Georgetown University
od126@georgetown.edu

Dr. Tatiane De Rossi
Brazilian Center for Research in Energy and Materi
tatiane.mazo@lnbio.cnpem.br

Dr. Iwjin De Vlaminck
Cornell University
vlaminck@cornell.edu

Prof. Deniz Demircioglu
Icahn School of Medicine at Mount Sinai
deniz.demircioglu@mssm.edu

Dr. Ahmet Denli
Genome Research, Associate Editor
denli@cshl.edu

Dr. Erica DePasquale
Cincinnati Children's Hospital
erica.depasquale@cchmc.org

Mr. Jairav Desai
RefinedScience
jairav.desai@refinedscience.com

Dr. Shreya Deshmukh
Cellanome, Inc.
shreya@cellanome.com

Mr. Ziteng Diao
Stony Brook University
ziteng.diao@stonybrookmedicine.edu

Ms. Trang Dinh
Case Western Reserve University
ttd28@case.edu

Dr. Jennifer Dionne
Stanford University
jdionne@stanford.edu

Mr. Igor Dolgalev
NYU Grossman School of Medicine
igor.dolgalev@nyulangone.org

Dr. Timothy Downing
University of California, Irvine
tim.downing@uci.edu

Prof. Bianca Dumitrascu
Columbia University
bmd2151@columbia.edu

Boden Eakins
University of California, Berkeley
eakins@berkeley.edu

Mr. Petur Einarsson
University of Iceland
phe1@hi.is

Dr. Sara Elgamal
University of Cincinnati
elgamasa@ucmail.uc.edu

Melanie Evans
IFF
melanie.brennan@iff.com

Mr. Jonathan Fan
McGill University
jonathan.fan@mail.mcgill.ca

Ms. Chih-Wei Fan
Icahn School of Medicine at Mount Sinai
chih-wei.fan@mssm.edu

Dr. Yilin Fan
Massachusetts General Hospital
yilin.fan@mgh.harvard.edu

Mr. Connor Finkbeiner
Fred Hutch
cfinkbei@fredhutch.org

Dr. Scott Fraser
Chan Zuckerberg Imaging Institute
sfraser@chanzuckerberg.com

Dr. Gabriella Frosi
Douglas Research Centre
gabriella.frosi.comtl@ssss.gouv.qc.ca

Rob Frost
Dartmouth College
rob.frost@dartmouth.edu

Dr. Jase Gehring
University of California, Berkeley
jgehring@berkeley.edu

Mr. Ian Gingerich
Dartmouth College
ian.gingerich.gr@dartmouth.edu

Alyx Gray
University of Colorado
alyx.gray@cuanschutz.edu

Ms. Emanuelle Grody
Northwestern University
emanuelle.grody@northwestern.edu

Dr. Abigail Groff
Conception Biosciences
abbie@conception.bio

Dr. Qijie Guan
Yale University
qijie.guan@yale.edu

Dr. Jeremy Gygi
AbbVie
jeremy.gygi@abbvie.com

Dr. Batoulsadat Haerian
Georgia Tech
batoolsadat@yahoo.com

Yeaeun Han
Mayo Clinic
yeaehan@gmail.com

Wuji Han
University of California-Berkeley
leonhan@berkeley.edu

Prof. Shinichi Hashimoto
Wakayama Medical University
hashimot@wakayama-med.ac.jp

Dr. Amy Herr
University of California, Berkeley
aeh@berkeley.edu

Ceci Hes
McGill University
ceci.hermosillasamayoa@mail.mcgill.ca

Mr. Nicholas Hilgert
Harvard University
nhilgert@g.harvard.edu

Dr. Zhiyuan Hu
Wuhan University
zhiyuan.hu@whu.edu.cn

Ms. Sarah Huang
Fred Hutchinson Cancer Center
huangy0204@gmail.com

Dr. Hiroaki Ikushima
University of Tokyo
ikushimah-int@h.u-tokyo.ac.jp

Dr. Tadashi Imafuku
Wakayama Medical University
imafuku@wakayama-med.ac.jp

Ms. Yujin Jeong
Pohang University of Science and Technology
yjjeong@postech.ac.kr

Dr. Harris Jeong
Altos Labs
hjeong@altoslabs.com

Ms. Vrinda Jethalia
Icahn School of Medicine at Mount Sinai
vrinda.jethalia@mssm.edu

Dr. Ping Kao
RIKEN
ping.kao@riken.jp

Dr. Anna Karlsson
Lund University
anna_f.karlsson@med.lu.se

Dr. Shana Kelley
Northwestern University
shana.kelley@northwestern.edu

Dr. Olha Kholod
Dartmouth College
olha.kholod@dartmouth.edu

Mr. Seung Tea Kim
CSHL
sekim@cshl.edu

Prof. Sung-Hak Kim
Chonnam National University
shkpales3@gmail.com

Dr. Junhyong Kim
University of Pennsylvania
junhyong@sas.upenn.edu

Ms. Minseo Kim
Chonnam National University
kimminseo512@jnu.ac.kr

Ms. Jisoo Kim
Gwangju Institute of Science and Technology
kimjisoo18@gm.gist.ac.kr

Dr. Seon Kinrot
Icahn School of Medicine at Mount Sinai
seon.kinrot@mssm.edu

Dr. Kyung Dae Ko
National Institute of Health
kok3@nih.gov

Dr. Timothy Koh
University of Illinois Chicago
tjkoh@uic.edu

Mr. Bence Kover
King's College London
bence.kover@kcl.ac.uk

Dr. Spencer Krieger
Carnegie Mellon University
skrieger@andrew.cmu.edu

Kelly Lagor
Illumina
klagor@illumina.com

Ms. Sofía Lavista Llanos
ArgenTag
sofia@argentag.com

Juliana Lee
University of California, Berkeley
juliana_lee@berkeley.edu

Dr. Linda Lee
Icahn School of Medicine at Mount Sinai
linda.lee2@mssm.edu

Ms. Dayeon Lee
Gwangju Institute of Science Technology
dayon927@gm.gist.ac.kr

Mr. Chan-Ho Lee
Chonnam National University
chanolee@jnu.ac.kr

Mr. Chan Woong Lee
pohang university of science and technology
cwlee@postech.ac.kr

Dr. Sivan Leviyang
Georgetown University
Sivan.Leviyang@georgetown.edu

Dr. Dan Li
CSHL
dali@cshl.edu

Banruo Li
McGill University
banruo.li@mail.mcgill.ca

Dr. Jingyi Jessica Li
Fred Hutchinson Cancer Center/ University of Washington
lijy03@g.ucla.edu

Mr. Zekai Li
Imperial College London
zekai.li18@imperial.ac.uk

Qinglan Ling
UMass Chan Medical School
qinglan.ling@umassmed.edu

Mr. Jean-Christophe Lone
Gulbenkian Institute for Molecular Medicine - GIMM
jean.clone@gimm.pt

Will Lounsbery-Scaife
Mount Sinai Icahn School of Medicine
will.lounsbery-scaife@mssm.edu

Ms. Shupeng Luxu
Ragon Institute
sluxu.Alisa@gmail.com

Wenxiu Ma
University of California Riverside
wenxiu.ma@ucr.edu

Ms. Katherine Ma
Yale University
katherine.ma@yale.edu

Mr. Vaidhyanathan Mahaganapathy
Ellison Medical Institute
vmahaganapathy@emila.org

Bryce Manifold
University of California, Berkeley
manifold@berkeley.edu

Yizi Mao
University of California, Berkeley
y.mao@berkeley.edu

Michael Marciano
Syracuse University
mamarcia@syr.edu

Dr. Luciano Martelotto
Adelaide University

Dr. David McKinnon
Stony Brook University
david.mckinnon@stonybrook.edu

Prof. Pall Melsted
University of Iceland
pmelsted@hi.is

Lieke Michielsen
Weill Cornell Medicine
lim4020@med.cornell.edu

Ms. Vincenta Mikulenaite
Vilnius University
mikulenaitev@gmail.com

Jason Morrison
Stowers Institute for Medical Research
mlc@stowers.org

Madeline Morrisson
Dartmouth College
madeline.j.morrisson.gr@dartmouth.edu

Lambda Moses
Columbia University
dl3764@columbia.edu

Dr. John Murray
Perelman School of Medicine, University of Pennsylvania
jmurr@upenn.edu

Dr. Sandeep Nadella
Cold Spring Harbor Laboratory/ Northwell Health
nadella@cshl.edu

Mr. Simonas Norvaisis
Vilnius University
simonas.norvaisis@gmail.com

Conrad Oakes
California Institute of Technology
coakes@caltech.edu

Dr. Hiroyuki Okada
the University of Tokyo
hokada-tky@g.ecc.u-tokyo.ac.jp

Erik Oliai
Cornell University
ero27@cornell.edu

Dr. Dominik Otto
Fred Hutchinson Cancer Center
dotto@fredhutch.org

Prof. Lior Pachter
California Institute of Technology
lpachter@caltech.edu

Mr. Xinhai Pan
Georgia Institute of Technology
xpan78@gatech.edu

Dr. Francesco Panariello
Fondazione Istituto Nazionale Genetica Molecolare
panariello@ingm.org

Dr. Eun Seo Park
POSTECH
espark0403@postech.ac.kr

Mr. Nikolaos Patikas
Brigham and Women's Hospital
npatikas@bwh.harvard.edu

Ms. Martina Pedna
Kings College London
martina.2.pedna@kcl.ac.uk

Sebastian Pott
University of Chicago
spott@uchicago.edu

Dr. Rishvanth Prabakar
Cold Spring Harbor Laboratory
kaliapp@cshl.edu

Mr. Dominic Pruss
Cold Spring Harbor
dominic_pruss1@baylor.edu

Dr. Nianchao Qian
University of Chicago
nianchaoqian@uchicago.edu

Dr. Jieqiong Qu
Altos Labs Ltd
jqu@altoslabs.com

Dr. FNU Rajan
University of Arizona
rpandey@arizona.edu

Angie Ramirez
Icahn School of Medicine at Mount Sinai
angie.ramirez@mssm.edu

Dr. Ben Raphael
Princeton University
braphael@princeton.edu

Subbaramireddy Remala
Brigham and Women's Hospital
sremala@bwh.harvard.edu

Mr. Joseph Rich
Caltech
jmrich@caltech.edu

Kent Riemondy
RefinedScience
kent.riemondy@refinedscience.com

Dr. Jack Riley
Babraham Institute
jack.riley@babraham.ac.uk

Mr. Scott Robinson
10x Genomics
scott.robinson@10xgenomics.com

Dr. Barbara Rosati
Stony Brook University
barbara.rosati.1@stonybrook.edu

Michal Rozenwald
UC Berkeley
michal_roz@berkeley.edu

Dr. Yunus Sahin
Cold Spring Harbor Laboratory
sahin@cshl.edu

Md. Abul Hassan Samee
Baylor College of Medicine
samee@bcm.edu

Dr. Jose Santos-Pereira
University of Seville
jsantos2@us.es

Dr. Tatjana Sauka-Spengler
Stowers Institute for Medical Research and the University of Oxford
tatjana.sauka-spengler@imm.ox.ac.uk

Mr. Leon Schwartz
Northwestern University
leon.schwartz@northwestern.edu

Dr. Snehal Shah
Texas A&M University
snehal.shah@ag.tamu.edu

Mr. Shahin Shahsavari
Johns Hopkins School of Medicine
shhsvri@gmail.com

Praveer Sharma
Cellanome
praveer.sharma@cellanome.com

Dr. Alaullah Sheikh
Washington University in St. Louis
asheikh@wustl.edu

Dr. Kohei Shimono
Sumitomo Pharma Co. Ltd. / UCLA
k0hei4m0n0@gmail.com

Kalyanee Shirlekar
University of Florida
kshirlekar@ufl.edu

Dr. Chang Shu
University of Southern California
aprilchangshu@gmail.com

Dr. Bogdan Sieriebriennikov
Pacific Biosciences (PacBio)
bsieriebriennikov@pacb.com

Dr. Peter Sims
Columbia University
pas2182@cumc.columbia.edu

Dr. Gunjan Singh
Brown University
gunjan_singh@brown.edu

Mr. Stephen Staklinski
Cold Spring Harbor Laboratory
staklins@cshl.edu

Caitlin Strassburg
University of Pittsburgh
cms496@pitt.edu

Dr. Aaron Streets
University of California, Berkeley
astreets@berkeley.edu

Ms. Pallavi Surana
SUNY-Stony Brook University
pallavi.surana@stonybrook.edu

Dr. Hillary Sussman
Genome Research, Executive Editor
hsussman@cshl.edu

Ms. Nikhila Swarna
California Institute of Technology
nswarna@caltech.edu

Dr. Qing Tan
University of Southern California
qingtan@usc.edu

Dr. Tze Guan Tan
Merck & Co
tze.guan.tan@msd.com

Dr. Kristof Torkenczy
Altos Labs
thoor444@gmail.com

Mr. Denis Torre
Memorial Sloan Kettering Cancer Center
torred1@mskcc.org

Dr. Leslie Torres Ulloa
Harvard University
ltorresulloa@fas.harvard.edu

Dr. Shirngwern Tsaih
Medical College of Wisconsin
stsaih@mcw.edu

Prof. Shikhar Uttam
University of Pittsburgh
shf28@pitt.edu

Dr. Anthony Veltri
University of Wisconsin-Madison
ajveltri@wisc.edu

Ms. Twinkle Bapuji Vhatkar
Mount Sinai Health System
twinklebapuji.vhatkar@mssm.edu

Mr. Chieh Wang
University of California, Santa Barbara
chieh_wang@ucsb.edu

Ms. Yan Wang
Peking University
pkuwangyan@163.com

Wen Wei
Arizona State University
wwei39@asu.edu

Dr. Yulong Wei
Yale University
yulong.wei@yale.edu

Prof. Joshua Weinstein
University of Chicago
jaweinst@uchicago.edu

Dr. Sherman Weissman
Yale University
sherman.weissman@yale.edu

Prof. Michael Whitfield
Dartmouth College
michael.l.whitfield@dartmouth.edu

Dr. Noa Wigoda
Weizmann Institute of Science
noa.wigoda@weizmann.ac.il

Peipei Wu
Cold Spring Harbor Lab
pwu@cshl.edu

Dr. Guanao Yan
Michigan State University
yanguana@msu.edu

Ms. Shiyi Yang
University of California, Berkeley
sy.yang@berkeley.edu

Mr. Rui Ye
MD Anderson Cancer Center
rye@mdanderson.org

Dr. Wei Yin
Wuhan University
kq000933@whu.edu.cn

Dr. Sergio Zamora-Erazo
Fox Chase Cancer Center
sergio.zamora-erazo@fccc.edu

Xinyi Zhang
Memorial Sloan Kettering Cancer Center
xinyizxx@gmail.com

Lin Zhang
University of Pennsylvania
linzhang@pennmedicine.upenn.edu

Hongpan Zhang
University of Virginia
hz9fq@virginia.edu

Dr. Xiuwei Zhang
Georgia Institute of Technology
xiuwei.zhang@gatech.edu

Dr. Anni Zhang
Stanford University
annizh@stanford.edu

Beatrice Zhang
Memorial Sloan Kettering
zhangb2@mskcc.org

Rui Zhang
Chinese Institute for Brain Research
ruizhangdmn@outlook.com

Mr. Ziqi Zhang
Georgia Institute of Technology
zhangziqibuaa@gmail.com

Bokai Zhu
MIT
zhubokai@mit.edu

Ms. Xiaohan Zhu
University of California-Berkeley
xiaohanzhu@berkeley.edu

Dr. Tarek Zikry
UNC-Chapel Hill
tarek@unc.edu

CODE OF CONDUCT FOR ALL PARTICIPANTS IN CSHL MEETINGS

Cold Spring Harbor Laboratory (CSHL or the Laboratory) is dedicated to pursuing its twin missions of research and education in the biological sciences. The Laboratory is committed to fostering a working environment that encourages and supports unfettered scientific inquiry and the free and open exchange of ideas that are the hallmarks of academic freedom. To this end, the Laboratory aims to maintain a safe and respectful environment that is free from harassment and discrimination for all attendees of our meetings and courses as well as associated support staff, in accordance with federal, state and local laws.

Consistent with the Laboratory's missions, commitments and policies, the purpose of this Code is to set forth expectations for the professional conduct of all individuals participating in the Laboratory's meetings program, both in person and virtually, including organizers, session chairs, invited speakers, presenters, attendees and sponsors. This Code's prohibition against discrimination and harassment is consistent with the Laboratory's internal policies governing conduct by its own faculty, trainees, students and employees.

By registering for and attending a CSHL meeting, either in person or virtually, participants agree to:

1. Treat fellow meeting participants and CSHL staff with respect, civility and fairness, without bias based on sex, gender, gender identity or expression, sexual orientation, race, ethnicity, color, religion, nationality or national origin, citizenship status, disability status, veteran status, marital or partnership status, age, genetic information, or any other criteria prohibited under applicable federal, state or local law.

2. Use all CSHL facilities, equipment, computers, supplies and resources responsibly and appropriately if attending in person, as you would at your home institution.

3. Abide by the CSHL Meeting Alcohol Policy (*see below)*.

Similarly, meeting participants agree to refrain from:

1. Harassment and discrimination, either in person or online, in violation of Laboratory policy based on actual or perceived sex, pregnancy status, gender, gender identity or expression, sexual orientation, race, ethnicity, color, religion, creed, nationality or national origin, immigration or citizenship status, mental or physical disability status, veteran status, military status, marital or partnership status, marital or partnership status, familial status, caregiver status, age, genetic information, status as a victim of domestic violence, sexual violence, or stalking, sexual reproductive health decisions, or any other criteria prohibited under applicable federal, state or local law.

2. Sexual harassment or misconduct.

3. Disrespectful, uncivil and/or unprofessional interpersonal behavior, either in person or online, that interferes with the working and learning environment.

4. Misappropriation of Laboratory property or excessive personal use of resources, if attending in person.

BREACHES OR VIOLATIONS OF THE CODE OF CONDUCT

Cold Spring Harbor Laboratory aims to maintain in-person and virtual conference environments that accord with the principles and expectations outlined in this Code of Conduct. Meeting organizers are tasked with providing leadership during each meeting, and may be approached informally about any breach or violation. Breaches or violations should also be reported to program leadership in person or by email:

- Dr. David Stewart, Grace Auditorium Room 204, 516-367-8801 or x8801 from a campus phone, stewart@cshl.edu
- Dr. Charla Lambert, Hershey Laboratory Room 214, 516-367-5058 or x5058 from a campus phone, clambert@cshl.edu

Reports may be submitted by those who experience harassment or discrimination as well as by those who witness violations of the behavior laid out in this Code.

The Laboratory will act as needed to resolve the matter, up to and including immediate expulsion of the offending participant(s) from the meeting, dismissal from the Laboratory, and exclusion from future academic events offered by CSHL.

If you have questions or concerns, you can contact the meeting organizers, CSHL staff.

For meetings and courses funded by NIH awards:

Participants may contact the Health & Human Services Office for Civil Rights (OCR). See this page for information on filing a civil rights complaint with the OCR; filing a complaint with CSHL is not required before filing a complaint with OCR, and seeking assistance from CSHL in no way prohibits filing complaints with OCR. You may also notify NIH directly about sexual harassment, discrimination, and other forms of inappropriate conduct at NIH-supported events.

For meetings and courses funded by NSF awards:

Participants may file a complaint with the NSF. See this page for information on how to file a complaint with the NSF.

Law Enforcement Reporting:

- For on-campus incidents, reports to law enforcement can be made to the Security Department at 516-367-5555 or x5555 from a campus phone.
- For off-campus incidents, report to the local department where the incident occurred.

In an emergency, dial 911.

DEFINITIONS AND EXAMPLES

Uncivil/disrespectful behavior is not limited to but may take the following forms:

- Shouting, personal attacks or insults, throwing objects, and/or sustained disruption of talks or other meeting-related events

Harassment is any unwelcome verbal, visual, written, or physical conduct that occurs with the purpose or effect of creating an intimidating, hostile, degrading, humiliating, or offensive environment or unreasonably interferes with an individual's work performance. Harassment is not limited to but may take the following forms:

- Threatening, stalking, bullying, demeaning, coercive, or hostile acts that may have real or implied threats of physical, professional, or financial harm
- Signs, graphics, photographs, videos, gestures, jokes, pranks, epithets, slurs, or stereotypes that comment on a person's sex, gender, gender identity or expression, sexual orientation, race, ethnicity, color, religion, nationality or national origin, citizenship status, disability status, veteran status, marital or partnership status, age, genetic information, or physical appearance

Sexual Harassment includes harassment on the basis of sex, sexual orientation, self-identified or perceived sex, gender expression, gender identity, and the status of being transgender. Sexual harassment is not limited to sexual contact, touching, or expressions of a sexually suggestive nature. Sexual harassment includes all forms of gender discrimination including gender role stereotyping and treating employees differently because of their gender. *Sexual misconduct* is not limited to but may take the following forms:

- Unwelcome and uninvited attention, physical contact, or inappropriate touching
- Groping or sexual assault
- Use of sexual imagery, objects, gestures, or jokes in public spaces or presentations
- Any other verbal or physical contact of a sexual nature when such conduct creates a hostile environment, prevents an individual from fulfilling their professional responsibilities at the meeting, or is made a condition of employment or compensation either implicitly or explicitly

MEETING ALCOHOL POLICY

Consumption of alcoholic beverages is not permitted in CSHL's public areas other than at designated social events (wine and cheese reception, picnic, banquet, etc.), in the Blackford Bar, or under the supervision of a licensed CSHL bartender.

No provision of alcohol by meeting sponsors is permitted unless arranged through CSHL.

Meeting participants consuming alcohol are expected to drink only in moderation at all times during the meeting.

Excessive promotion of a drinking culture at any meeting is not acceptable or tolerated by the Laboratory. No meeting participant should feel pressured or obliged to consume alcohol at any meeting-related event or activity.

VISITOR INFORMATION

EMERGENCY (to dial outside line, press 3+1+number)	
CSHL Security	**516-367-8870 (x8870 from house phone)**
CSHL Emergency	**516-367-5555 (x5555 from house phone)**
Local Police / Fire	**911**
Poison Control	**(3) 911**

CSHL SightMD Center for Health and Wellness ***(call for appointment)*** Dolan Hall, East Wing, Room 111 **cshlwellness@northwell.edu**	**516-422-4422** **x4422 from house phone**
Emergency Room **Huntington Hospital** 270 Park Avenue, Huntington	**631-351-2000**
Dentists Dr. William Berg Dr. Robert Zeman	 **631-271-2310** **631-271-8090**
Drugs - 24 hours, 7 days Rite-Aid 391 W. Main Street, Huntington	**631-549-9400**

GENERAL INFORMATION

Meetings & Courses Main Office

Hours during meetings: M-F 9am – 9pm, Sat 8:30am – 1pm
After hours – See information on front desk counter
For assistance, call Security at 516-367-8870
(x8870 from house phone)

Dining, Bar

Blackford Dining Hall (main level):
Breakfast 7:30–9:00, Lunch 11:30–1:30, Dinner 5:30–7:00
Blackford Bar (lower level): 5:00 p.m. until late

House Phones

Grace Auditorium, upper / lower level; Cabin Complex; Blackford Hall; Dolan Hall, foyer

Books, Gifts, Snacks, Clothing

CSHL Bookstore and Gift Shop
516-367-8837 (hours posted on door)
Grace Auditorium, lower level.

Computers, E-mail, Internet access

Grace Auditorium
Upper level: E-mail and printing in the business center area
WiFi Access: GUEST (no password)

Announcements, Message Board Mail, ATM, Travel info

Grace Auditorium, lower level

Russell Fitness Center
Dolan Hall, east wing, lower level
PIN#: **(On your registration envelope)**

Laundry Machines
Dolan Hall, lower level

Photocopiers, Journals, Periodicals, Books
CSHL Main Library
Open 24 hours (with PIN# or CSHL ID)
Staff Hours: 9:00 am – 9:00 pm
Use PIN# (On your registration envelope) to enter Library
See Library staff for photocopier code.
Library room reservations (hourly) available on request between 9:00 am – 9:00 pm

Swimming, Tennis, Jogging, Hiking
June–Sept. Lifeguard on duty at the beach. 12:00 noon–6:00 p.m.
Two tennis courts open daily.

Local Interest

Fish Hatchery	631-692-6758
Sagamore Hill	516-922-4788
Whaling Museum	631-367-3418
Heckscher Museum	631-351-3250
CSHL DNA Learning Center	x 5170

New York City
Helpful tip -
Take CSHL Shuttle OR Uber/Lyft/Taxi to Syosset Train Station
Long Island Railroad to Penn Station
Train ride about one hour.

TRANSPORTATION

Limo, Taxi

Syosset Limousine	516-364-9681
Executive Limo Service	516-826-8172
Limos Long Island	516-400-3364
Syosset Taxi	516-921-2141
Orange & White Taxi	631-271-3600
Uber / Lyft	

Trains

Long Island Rail Road	718-217-LIRR (5477)
Amtrak	800-872-7245
MetroNorth	877-690-5114
New Jersey Transit	973-275-5555

CSHL's Green Campus

Cold Spring Harbor Laboratory is pledged to operate in an environmentally responsible fashion wherever possible. In the past, we have removed underground oil tanks, remediated asbestos in historic buildings, and taken substantial measures to ensure the pristine quality of the waters of the harbor. Water used for irrigation comes from natural springs and wells on the property itself. Lawns, trees, and planting beds are managed organically whenever possible. And trees are planted to replace those felled for construction projects.

Two areas in which the Laboratory has focused recent efforts have been those of waste management and energy conservation. The Laboratory currently recycles most waste. Scrap metal, electronics, construction debris, batteries, fluorescent light bulbs, toner cartridges, and waste oil are all recycled. For general waste, the Laboratory uses a "single stream waste management" system, removing recyclable materials and sending the remaining combustible trash to a cogeneration plant where it is burned to provide electricity, an approach considered among the most energy efficient, while providing a high yield of recyclable materials.

Equal attention has been paid to energy conservation. Most lighting fixtures have been replaced with high efficiency fluorescent fixtures, and thousands of incandescent bulbs throughout campus have been replaced with compact fluorescents. The Laboratory has also embarked on a project that will replace all building management systems on campus, reducing heating and cooling costs by as much as twenty-five per cent.

Cold Spring Harbor Laboratory continues to explore new ways in which we can reduce our environmental footprint, including encouraging our visitors and employees to use reusable containers, conserve energy, and suggest areas in which the Laboratory's efforts can be improved. This book, for example, is printed on recycled paper.

Cold Spring Harbor Laboratory Bookstore & Gift Shop

Main campus, lower level of Grace Auditorium

Store Hours

Did you miss your chance to shop at the CSHL Bookstore and Gift Shop during the conference?

No problem! Visit our Online Bookstore and Gift Shop.

It's a great way to bring home a piece of the experience!

Contact Us
bookstore@cshl.edu
x8837

science books
CSHL-branded merchandise
unique gifts
souvenirs
. . . and more!

Visit our website
cshlvirtualstore.com

CSHL Campus Map

CSHL Map