Ann. N.Y. Acad. Sci. ISSN 0077-8923

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Special Issue: *Keystone Symposia Reports*

Concise Original Report

Single cell biology—a Keystone Symposia report

Jennifer Cable,[1] Michael B. Elowitz,[2,3] Ana I. Domingos,[4,5] Naomi Habib,[6,7] Shalev Itzkovitz,[8] Homaira Hamidzada,[9] Michael S. Balzer,[10] Itai Yanai,[11] Prisca Liberali,[12] Jessica Whited,[13] Aaron Streets,[14,15] Long Cai,[2] Andrew B. Stergachis,[16] Clarice Kit Yee Hong,[17,18] Leeat Keren,[8,19] Martin Guilliams,[20] Uri Alon,[21] Alex K. Shalek,[22] Regan Hamel,[23] Sarah J. Pfau,[24] Arjun Raj,[25,26] Stephen R. Quake,[15,27,28] Nancy R. Zhang,[29] Jean Fan,[30] Cole Trapnell,[31] Bo Wang,[27,32] Noah F. Greenwald,[19] Roser Vento-Tormo,[33] Silvia D.M. Santos,[34] Sabrina L. Spencer,[35] Hernan G. Garcia,[36] Geethika Arekatla,[37] Federico Gaiti,[38] Rinat Arbel-Goren,[39] Steffen Rulands,[40] Jan Philipp Junker,[41] Allon M. Klein,[42] Samantha A. Morris,[18,43] John I. Murray,[25] Kate E. Galloway,[44] Michael Ratz,[45] and Merrit Romeike[46]

[1]PhD Science Writer, New York, New York. [2]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California. [3]Howard Hughes Medical Institute, California Institute of Technology, Pasadena, California. [4]Department of Physiology, Anatomy & Genetics, Oxford University, Oxford, United Kingdom. [5]The Howard Hughes Medical Institute, New York, New York. [6]Cell Circuits Program, Broad Institute, Cambridge, Massachusetts. [7]Edmond & Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel. [8]Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel. [9]Toronto General Hospital Research Institute, University Health Network; Translational Biology and Engineering Program, Ted Rogers Centre for Heart Research and Department of Immunology, University of Toronto, Toronto, Ontario, Canada. [10]Renal, Electrolyte, and Hypertension Division, Department of Medicine and Institute for Diabetes, Obesity, and Metabolism, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania. [11]Institute for Computational Medicine, NYU Langone Health, New York, New York. [12]Friedrich Miescher Institute for Biomedical Research (FMI), Basel, Switzerland. [13]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, Massachusetts. [14]Department of Bioengineering and Center for Computational Biology, University of California, Berkeley, Berkeley, California. [15]Chan Zuckerberg Biohub, San Francisco, California. [16]Division of Medical Genetics, Department of Medicine, University of Washington, Seattle, Washington; and Brotman Baty Institute for Precision Medicine, Seattle, Washington. [17]Edison Center for Genome Sciences and Systems Biology, Washington University in St. Louis, St. Louis, Missouri. [18]Department of Genetics, Washington University in St. Louis, St. Louis, Missouri. [19]Department of Pathology, School of Medicine, Stanford University, Stanford, California. [20]Laboratory of Myeloid Cell Biology in Tissue Homeostasis and Regeneration, VIB-UGent Center for Inflammation Research, and Unit of Immunoregulation and Mucosal Immunology, VIB Inflammation Research Center, and Department of Biomedical Molecular Biology, Ghent University, Ghent, Belgium. [21]Faculty of Sciences, Department of Human Biology, University of Haifa, Haifa, Israel. [22]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts. [23]Department of Clinical Neurosciences and NIHR Biomedical Research Centre, University of Cambridge, Cambridge, United Kingdom. [24]Department of Neurobiology, Harvard Medical School, Boston, Massachusetts. [25]Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania. [26]Department of Bioengineering, School of Engineering and Applied Sciences, University of Pennsylvania, Philadelphia, Pennsylvania. [27]Department of Bioengineering, Stanford University, Stanford, California. [28]Department of Applied Physics, Stanford University, Stanford, California. [29]Graduate Group in Genomics and Computational Biology and Department of Statistics, University of Pennsylvania, Philadelphia, Pennsylvania. [30]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland. [31]Department of Genome Sciences, University of Washington School of Medicine; Brotman Baty Institute for Precision Medicine; and Allen Discovery Center for Cell Lineage Tracing, Seattle, Washington. [32]Department of Developmental Biology, Stanford University School of Medicine, Stanford, California. [33]Wellcome Sanger Institute, Cambridgeshire, United Kingdom. [34]The Francis Crick Institute, London, United Kingdom. [35]Department of Biochemistry and BioFrontiers Institute, University of Colorado Boulder, Boulder, Colorado. [36]Department of Physics; Biophysics Graduate Group; Department of Molecular and Cell Biology; and Institute for Quantitative Biosciences-QB3, University of California at Berkeley, Berkeley, California. [37]ETH Zurich, Zurich, Switzerland. [38]New York Genome Center and Meyer Cancer Center, Weill Cornell Medicine, New York, New York. [39]Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot, Israel. [40]Max Planck Institute for the Physics of Complex Systems, and Center for Systems Biology Dresden, Dresden, Germany. [41]Berlin Institute for Medical Systems Biology, Max

doi: 10.1111/nyas.14692

Delbrück Center for Molecular Medicine, Berlin, Germany. [42]Department of Systems Biology, Blavatnik Institute, Harvard Medical School, Boston, Massachusetts. [43]Department of Developmental Biology and Center of Regenerative Medicine, Washington University School of Medicine, St. Louis, Missouri. [44]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts. [45]Department of Cell and Molecular Biology, Karolinska Institute, Solna, Sweden. [46]Max Perutz Laboratories Vienna, University of Vienna, Vienna, Austria

Address for correspondence: Annals Author. annals@nyas.org

Single cell biology has the potential to elucidate many critical biological processes and diseases, from development and regeneration to cancer. Single cell analyses are uncovering the molecular diversity of cells, revealing a clearer picture of the variation among and between different cell types. New techniques are beginning to unravel how differences in cell state—transcriptional, epigenetic, and other characteristics—can lead to different cell fates among genetically identical cells, which underlies complex processes such as embryonic development, drug resistance, response to injury, and cellular reprogramming. Single cell technologies also pose significant challenges relating to processing and analyzing vast amounts of data collected. To realize the potential of single cell technologies, new computational approaches are needed. On March 17–19, 2021, experts in single cell biology met virtually for the Keystone eSymposium "Single Cell Biology" to discuss advances both in single cell applications and technologies.

Keywords: development; differentiation; lineage tracing; reprogramming; single cell sequencing; spatial transcriptomics

Introduction and keynote address

The ability to isolate and understand the behavior of single cells within a larger system has changed the way biological systems are approached. Single cell biology has the potential to transform our understanding of many critical biological processes and diseases, from development and regeneration to cancer. Despite this potential, single cell biology also faces many challenges. The wealth of data accrued by analyzing transcriptomic, epigenomic, and other data obtained from thousands of cells requires new computational tools and data analysis approaches.

On March 17–19, 2021, experts in single cell biology met virtually for the Keystone eSymposium "Single Cell Biology." Bringing together researchers developing single cell biological tools with those utilizing these technologies, the symposium's goals included identifying new opportunities for technology development and applications.

The meeting focused on (1) *development*, how individual cells differentiate within the context of a developing embryo to generate all the cell types and structures of an adult organism; (2) *spatial analysis*, identifying the different cell types within a tissue or disease site and determining how they interact with and respond to their neighbors to support tissue function and homeostasis; (3) *disease*, understanding how cell populations change in response to disease or injury and how they localize within a diseased tissue; and (4) *new technologies*, new tools for single cell data analysis to make precision measurements on multiple aspects of cellular identity, such as morphology, composition and spatial distribution of proteins and metabolites, gene expression, chromatin organization, and epigenetics.

Michael B. Elowitz from the California Institute of Technology opened the symposium with the keynote address, beginning with an overview of

the recent revolution in single cell biology. Historically, cells have been identified and classified based on morphology and function. More recent single cell analyses are uncovering the immense molecular diversity of cells, revealing a complex picture of the variation among and between different cell types.

Technologies like single cell RNA sequencing (scRNA-seq) provide an unbiased way to determine the transcriptional state of individual cells; clustering cells based on their transcriptional profile can reveal cell types and provide insight on how cell states change during development and differentiation. Spatial transcriptomics and imaging techniques can provide information on where different cell states localize within a tissue or tumor microenvironment (TME). Dynamic barcoding techniques are able to not only provide information on cell types, but also on how different cell types are related, thus enabling researchers to reconstruct lineage relationships of cells during development or even across evolution.

During his keynote, Elowitz stressed that one of the remaining unmet needs in single cell biology is understanding the underlying molecular circuits that establish a variety of fates that cells can transition to and from in a controlled manner. Although there has been a lot of progress delineating the signals and transcription factors that control cell fate, mapping these interactions as circuits gets complex very quickly. It is difficult to tease out which subsets of interactions are sufficient to result in multistability and controlled transitions.

Elowitz's group is creating synthetic circuits that establish, stabilize, and control cell states to understand how cells can exist in a huge variety of controllable states. There are several features that a synthetic circuit should recapitulate:

- multistability: the ability to produce multiple stable states;
- state switching: the ability to switch states in a controlled way;
- irreversible and hierarchical differentiation programs; and
- expandability: the ability to expand the system to include more states.

Looking at natural circuits, Elowitz's group noticed two common themes among mammalian fate regulators: promiscuous dimerization, for example, many transcription factors function as homo- and heterodimers, and autoregulation.

Ronghui Zhu, a graduate student in Elowitz's group, designed a class of synthetic cell fate control circuits based on these two themes called MultiFate. The simplest MultiFate circuit, MultiFate-2, consists of two transcription factors that can homodimerize to activate expression of their respective genes or form an inactive heterodimer. The system can be simply described mathematically. The phase diagram of MultiFate-2 shows that the system results in three stable conditions that can be destabilized by altering parameters like protein stability. MultiFate-2 can also be expanded to MultiFate-3, which incorporates a third transcription factor to generate seven stable states. Elowitz's group has validated the MultiFate model in cells using zinc finger motifs that are engineered to homodimerize and activate expression of a fluorescent target; heterodimerization with a competing zinc finger blocks gene expression. The system is controllable in that dimerization and protein stability can be controlled via small molecules.[1]

As predicted by the phase diagram, MultiFate-2 produces three stable states in cells that remain stable for at least 18 days. The system allows for state switching—ectopic expression of a transcription factor that switches a cell's state—as well as for irreversible transitions. Transiently destabilizing the heterodimer state by decreasing protein stability forces cells into either a high A or a high B state; cells do not revert back to their original state when protein stability is restored. Elowitz's group has also incorporated a three-component MultiFate system (MultiFate-3) into cells, which similarly behaves as predicted: generating seven states, each stable over extended timescales of weeks.[1]

Elowitz expects that MultiFate will provide a core capability necessary for synthetic multicellular organisms. His lab is also working on other features of organisms, such as intercellular signaling[2] and cell population control, to build such a synthetic system.[3]

Single cell biology of mammalian organs

The first session of the symposium focused on single cell biology of mammalian organs. Speakers showed how single cell techniques can be used to understand how different cell types function

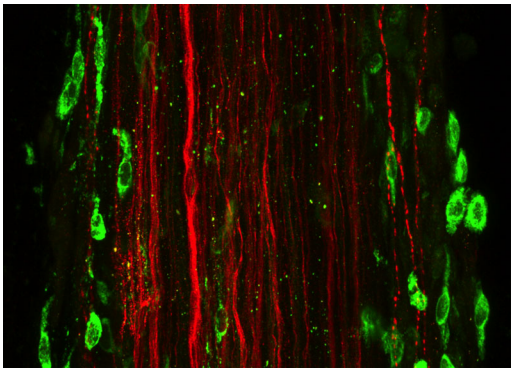

Figure 1. Sympathetic neuron-associated macrophages contribute to obesity by importing and metabolizing norepinephrine. These macrophages integrate a heterogeneous neuroimmune niche within sympathetic nerves (Red: TH, green: LyzM-GFP).

and interact within an organ, both in normal and disease-related processes.

Ana I. Domingos from the University of Oxford showed how sympathetic nerves that enervate white adipose tissue regulate adipose cells (Fig. 1). Domingos's lab was the first to confirm that white adipose tissue is enervated, that is, sympathetic nerves form neuro-adipose junctions. Using optogenetics to locally activate sympathetic neurons in adipose tissue was sufficient to induce leptin-driven lipolysis within adipocytes, leading to depleting of a white adipose mass.[4] Adipose tissue normally produces the hormone leptin in proportion to the amount of fat in the body. When the brain senses leptin, it sends signals to decrease food intake, which reduces adipose tissue. Once leptin levels decrease to a certain point, the brain increases food intake. At homeostasis, leptin signaling keeps weight within a defined, narrow range. During obesity, leptin signaling is not sensed by the brain, and lipolysis is reduced. Domingos is investigating whether activating sympathetic neurons in adipose tissue by targeting sympathetic-associated macrophages can induce lipolysis and promote weight loss.[5] Her lab developed a class of drugs known as sympathofacilitators, which have been shown to increase lipolysis and thermogenesis in mice without altering food intake or locomotor activity.[6] Domingos also presented unpublished single cell sequencing data showing that there are many types of immune cells in the sympathetic ganglia. She hopes that characterizing these immune cells will provide insights into the role of leptin in immunometabolism.[7]

Naomi Habib from the Hebrew University of Jerusalem presented work on dissecting the cellular landscape of Alzheimer's disease. Multiple brain cells have been implicated in disease progression. Instead of taking a neuron-centric view of Alzheimer's disease, Habib's group is interested in the entire cellular environment that may contribute to pathology, hypothesizing that cells such as microglia, oligodendrocytes, astrocytes, and vascular cells may be driving disease. Using single-nucleus RNA-seq (snRNA-seq) to build detailed cellular maps of healthy and diseased brains in mice, Habib's group, in collaboration with the Schwartz and Regev labs, has identified glial and astrocyte populations associated with Alzheimer's disease. Disease-associated astrocytes have a distinct expression profile, are localized near amyloid plaques, appear early in the course of Alzheimer's disease, and increase with natural age.[8] Habib's group is working to understand the functional role of disease-associated astrocytes and on profiling other disease-associated cell types in the mouse and human brains. In collaboration with the De Jager and Menon labs and others, they have used a combination of sn RNA-seq on 24 postmortem human brains and bulk sequencing of 640 brains to identify disease-associated cells. Similar to mouse brains, the data from human brains reveal specific glial cell populations associate with specific Alzheimer's disease-related traits such as cognitive decline or neurofilament tangles load.[9] The group found a network structure among the different cell populations, composed of cellular communities of coordinated cell populations, and linked specific cellular communities with Alzheimer's disease traits. In this way, Habib and her colleagues hope to describe the cellular environment of cognitive decline in aging brains.

Shalev Itzkovitz from the Weizmann Institute presented work on how enterocytes of the intestinal epithelium change as they migrate from the crypt up the villus. The environment surrounding the villi is not uniform; there are oxygen, nutrient, bacterial, and signaling gradients from the crypt to the tip that could impact cell fate as progenitor cells in the crypt differentiate and migrate up the villus to the tip

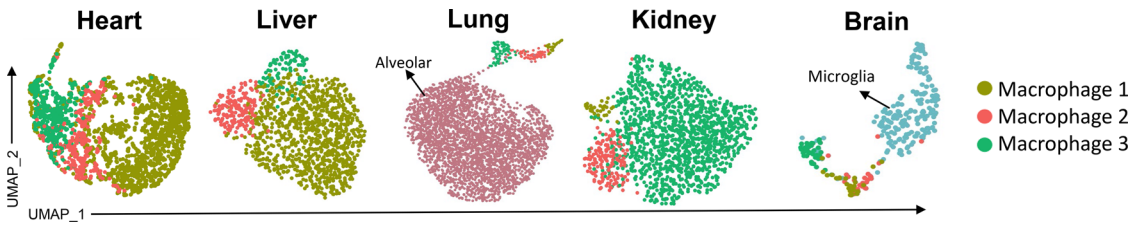

Figure 2. scRNA-seq to identify macrophage signatures and subpopulations that are shared across organs.

before they are finally shed.[10] Itzkovitz's group characterized the transcriptional activity of five regions in the villus by bulk RNA-seq and used these transcriptional phenotypes to localize single-sequenced cells along the villus. The results show that cells along the length of the villus are transcriptionally diverse, indicating that cells have different functions as they migrate. At the bottom of the villus, enterocytes express antimicrobial genes. As they travel up, they sequentially express the machinery for absorption of amino acids, carbohydrates, peptides, and lipids. At the very tip of the villus, enterocytes express genes involved in immune modulation.[11] Itzkovitz's work shows that enterocytes are not terminally differentiated but constantly differentiating, changing their transcriptional profile and function as they move up the villus. Itzkovitz also described unpublished work evaluating the changes in the proteome along the villus and showed that transcriptional profiles do not always correlate with protein profiles, primarily for proteins with long half-lives. These differences between mRNA and protein levels have enabled Itzkovitz's group to refine their map of the function of enterocytes along the villus and gain further insights into how the cells optimize resource utilization for absorption-related functions in the more resource-limited areas of the villus.

Elucidating cell heterogeneity within organs

Homaira Hamidzada from Slava Epelman's lab at the University of Toronto presented unpublished work to understand macrophage heterogeneity. Previous work on macrophages has suffered from a lack of consistency, making it difficult to find commonalities and differences between studies. Hamidzada has characterized macrophages from five different organs in mice by scRNA-seq, identifying macrophage signatures and subpopulations that are shared across organs (Fig. 2).

Michael S. Balzer from Katalin Susztak's lab at the University of Pennsylvania showed how single cell profiling can be used to understand an organ's response to injury. Following acute kidney injury, the kidney can either recover or the injury can progress and become chronic. Working in mice, Balzer *et al.* developed a model of adaptive and maladaptive kidney regeneration by titrating ischemic injury dose. In addition to performing detailed biochemical and histological analysis, they profiled transcriptomic changes at bulk and single cell levels in >110,000 cells over time; these currently unpublished data show that varying degrees of ischemic injury can result in trajectories in renal cells with different expression programs. Further analysis indicated kidney proximal tubule (PT) cells as cells particularly susceptible to injury. Adaptive PT repair correlated with fatty acid oxidation and oxidative phosphorylation. They also identified a specific maladaptive profibrotic PT cluster after long ischemia; these cells expressed proinflammatory and profibrotic cytokines and attracted myeloid cells. Druggability analysis highlighted pyroptosis and ferroptosis as vulnerable pathways in profibrotic cells.

Novel technologies in single cell analysis

Several speakers presented work on developing new single cell technologies. Many of them focused on spatially resolving either single cells within a tissue or subcellular components within a single cell.

Spatial transcriptomics of cell populations within tissues

With regard to localizing cell populations within a tissue, **Itai Yanai** from New York University presented work on spatial transcriptomics in tumors. ScRNA-seq has revealed heterogeneity within malignant cells that may have functional consequences in diverse cancer types. Yanai's group

has detected three cancer cell states in melanoma tumors via scRNA-seq.[12] To confirm that these states exist in the tumor itself, and not merely artifacts of single cell sequencing, Yanai used array-based spatial transcriptomics.[13] In this method, a tissue cryosection is placed on a barcoded array that contains oligonucleotides that can capture an entire transcriptome and correlate with position on the array; the technique does not provide single cell resolution, as each spot on the array represents approximately ten cells. Array-based spatial transcriptomics recapitulated the three transcriptional programs identified by scRNA-seq in melanoma tumors.[12] Array-based spatial transcriptomics can be combined with scRNA-seq to reveal associations among cell states and cell types within the TME. Using these methods together revealed that in pancreatic ductal adenocarcinoma inflammatory fibroblasts are consistently associated with cancer cells that express stress-like gene programs.[14] Yanai's lab is currently working to identify the intrinsic and extrinsic factors that lead to these cancer cell states, as well as their adaptive consequences within the tumor.[15]

Spatial resolution of subnuclear features

Several speakers described new technologies to investigate the spatial arrangement of the genome and nuclear features. Genomes are organized at several levels, for example, DNA wrapped around histones forming nucleosomes and higher order chromatin arrangements. Techniques like Hi-C,[16] GAM,[17] and SPRITE[18] can provide information on what regions of the chromosome are near each other in space, while methods like DNA fluorescence *in situ* hybridization (FISH), chromosome paint,[19] and *in situ* sequencing[20] enable visualization of the three-dimensional (3D) organization of chromosomes in the nucleus. In addition to a genome, the nucleus also contains other features, such as chromatin compartments, nuclear pore complexes, the nucleolus, and lamina, which may affect the spatial organization of chromatin and gene regulation.

Aaron Streets from the University of California, Berkeley described a microfluidic platform dubbed microDamID or "mu"DamID (microfluidic DNA adenine methyltransferase identification) that merges imaging and sequencing measurements within the same cell to map the association of the genome with nuclear features. Previously developed by van Steensel and Henikoff,[21] DamID measures protein–genome interactions by genetically fusing a protein to DNA adenine methyltransferase (DAM). If the protein interacts with DNA, DAM will methylate nearby sequences. DamID has proven to be a powerful sequencing tool for single cell measurements, where it has been used to measure interactions between the genome and the nuclear lamina.[22]

Streets's group has leveraged DamID in a microfluidic platform that sequentially images single cells and then performs DamID through a series of sequential chambers. This technique has been used to generate single cell laminar interaction maps.[23] By comparing data across cells, lamina-associated domains within the genome can be divided into constitutive and variable domains on the basis of the number of cells in which a locus is associated with the lamina. Lamina association can be used to probe cell type, as constitutive lamina-associated domains are associated with low gene expression. Streets's group is further exploring the lamina maps to see if there are associations between nuclear abnormalities and lamina-associated domains. Although the original microfluidic DamID technique was relatively low throughput, microfluidic cell barcoding and sequencing, which link barcodes to cell images, can increase the throughput.[24]

Long Cai from the California Institute of Technology presented work on a separate technology to image subnuclear structures. Cai's group has combined DNA sequential FISH (seqFISH+) with RNA seqFISH and multiplexed immunofluorescence to profile mRNA, DNA, and protein loci *in situ* in single cells.[25] The method expands upon single-molecule FISH, which enables individual mRNA transcripts to be visualized within a single cell.[26,27] SeqFISH uses multiple rounds of hybridization to distinguish different molecular species.[28–30] RNA seqFISH can scale up to the genomic level to localize tens of thousands of mRNA transcripts within a single cell.[31] DNA seqFISH+ can resolve chromosome structures on the order of ∼1 MB across the genome to localize the chromosomes within the nucleus. Multiplexed immunofluorescence is used to localize proteins such as lamins, heterochromatic markers, and nucleolar markers. These three techniques are integrated into a single experiment to

visualize mRNA, DNA, and proteins within a single nucleus.[25]

Cai's group has used this technique to generate the global architecture of the nucleus, reconstructing chromosomes with up to 25-kb resolution and systematically characterizing DNA–chromatin interactions. Cai showed that there are certain DNA loci that are consistently associated with nuclear features across cells: while the arrangement of the nuclear bodies and chromosomes can appear random, when taken together, there are anchoring points with respect to each other. Different chromosomes have different arrangements of these anchoring points, which results in a deterministic scaffold for chromosomal arrangement in the nucleus.

Andrew B. Stergachis from the University of Washington showed how Fiber-seq can be used to determine the chromatin architecture of individual chromatin fibers and, ultimately, provide insights into how neighboring features along the same chromatin fiber work together to regulate gene expression. Fiber-seq improves upon the current resolution limit of single cell epigenomic methods. It can achieve near base-pair resolution for fibers greater than 10 kb in length. In short, nuclei are treated with N6-methyladenine methyltransferase, which selectively marks sites of chromatin accessibility by incorporating a nonnative methyl group onto adenine bases. After DNA extraction and fragmentation, the methylated sequences are determined via single-molecule long-read DNA sequencing. Fiber-seq can identify regions of actuated regulatory DNA as well as protein-bound regions. Stergachis showed that comparing patterns across fibers from multiple cells to construct the chromatin architecture of individual fibers reveals widespread heterogeneity across cells. Because the mapped chromatin fibers are so long, Fiber-seq can provide information on coactuation of neighboring regulatory elements. Fiber-seq mapping showed that neighboring regulatory elements prefer to be coactuated on the same chromatin fiber, suggesting that regulatory architectures can impact neighboring elements.[32]

Clarice Kit Yee Hong from Barak Cohen's lab at Washington University in St. Louis presented unpublished work on a technique to identify chromatin features that influence gene expression noise. Gene expression noise can be thought of as the variance in gene expression within a cell population. Studies of position effect variegation have shown that chromatin environments at different chromosomal locations can affect a gene's expression. Genes in euchromatin are expressed ubiquitously from cell to cell, whereas there is more variation between cells for genes located near heterochromatin.

Hong and coworkers have developed single-cell Thousands of Reporters Integrated in Parallel (scTRIP) to measure the effects of different chromatin environment on gene expression noise; scTRIP, an extension of TRIP,[33] measures the expression noise of the same reporter gene integrated at many genomic locations. In short, reporter genes are integrated into random locations with the genome, and the DNA and mRNA from the reporter genes are counted to determine gene expression; scTRIP can reliably measure noise across the genome and provide insight into how diverse chromatin environments affect expression noise.

Single cell approaches to bridge local and systemic decision making

Single cell technologies have also proven useful for understanding different levels of decision making. One of the key questions for understanding multicellular organisms is how distinct activities at a single cell level affect multicellular processes like development—and vice versa, how signals from an organ, tissue, or embryo have effects at the single cell level.

Prisca Liberali from Friedrich Miescher Institute for Biomedical Research presented work on single cell approaches to understanding collective cell behavior. Liberali's lab is interested in understanding regulation at the individual cell level and how individual decisions are coordinated at the multicellular level to achieve the higher order functions necessary for organs and tissue. Ultimately, Liberali hopes to understand the intrinsic and extrinsic factors that enable genetically identical cells within an organism to coordinate across both spatial and time scales. This requires methods that can bridge large ranges of spatial and temporal scales, for example, relating events that occur at the molecular level within minutes to events that occur at the tissue

level within days. Liberali's group combines imaging technologies with scRNAseq to determine cellular state during cell-type transitions and infer trajectories of development. They have shown that there is a lot of variability between cells at branch points during differentiation; yet, defining the trajectory of a given cell within this population by scRNA-seq is difficult because it is often determined by extrinsic cues. Liberali's group is developing various methods to address these problems. They have characterized hundreds of thousands of individual organoids with high-dimensional measurements and perturbations.[34,35] Liberali also discussed unpublished data focusing on symmetry breaking within embryos.

Jessica Whited from Harvard University presented work on local and systemic responses to injury in the salamander axolotl. Axolotl, like all salamander species, can regenerate limbs after amputation. The process involves formation of a blastema made of undifferentiated progenitor cells that proliferate, differentiate, and eventually form a new limb. Whited's group has performed bulk and single-cell sequencing of blastema cells to understand the genes and trajectories involved in blastema formation.[36–38] During her talk, Whited focused on what occurs before blastema formation. She proposed a two-step model for blastema formation—an activation step in which some cells reenter the cell cycle and a conversion event, in which a portion of activated cells convert to blastema cells. Although many animals undergo the activation step in response to injury, the conversion step is unique to regenerative animals. Whited's group has shown that amputation elicits a systemic response, activating cell proliferation in both injured and uninjured limbs. Many of the activated cells are muscle satellite cells that replenish the muscle after injury.[39] The group is working to understand the signals involved that activate cells both locally and systemically and why only cells at the injured limb undergo conversion to blastema cells. Toward this goal, they are developing technologies to visualize cells at various stages of the cell cycle to learn more about the activated cells and leveraging bulk sequencing data to identify blastema reporters and visualize blastema development in live animals to monitor systemic blastema reporter activation. Additional work is focusing on identifying the full repertoire of cell types in responding tissues and how the activation signal is extinguished in distant sites.

Single cell analysis in pathology

As the roles of, and interactions between, different cell types and populations become more apparent in diseases and injury response, single cell technologies can provide unique insights into how cell populations change in response to injury or diseases such as cancer. Single cell technologies allow researchers to understand not only broad changes in cell populations, but also to pinpoint the spatial and temporal changes in cell populations. Below are presentations focused on using technologies that provide spatial insights on specific cell populations, theoretical modeling to predict how cell populations change in response to injury, and scRNA-seq to characterize how cell populations are affected by SARS-CoV-2 infection and how that relates to disease severity.

Leeat Keren from the Weizmann Institute of Science described work using multiplexed ion beam imaging by time of flight (MIBI-TOF) to better understand the TME. MIBI-TOF provides single-cell as well as subcellular resolution of expression of up to 40 proteins, for example, tumor, immune cell, and immune regulation markers.[40] Keren's group has used MIBI-TOF to characterize immune cells in tumors of patients with triple-negative breast cancer (TNBC). They found that the immune composition was highly heterogeneous between patients. Patients with more leukocytes had a higher proportion of T regulatory cells while those with fewer had a higher proportion of macrophages, suggesting organization in the immune response to TNBC; indeed, spatial enrichment analysis revealed subtypes of immune–tumor organization. Keren showed that immunologically hot tumors, with high tumor infiltrating lymphocytes, can be classified into compartmentalized and mixed tumors. In compartmentalized tumors, large areas of immune cells are spatially separated from tumor cells; the tumor–immune boundary contains suppressive PD-L1–expressing immune cells. By contrast, in mixed tumors, tumor and immune cells are more interspersed, and PD-L1 is primarily expressed on tumor cells. Compartmentalized tumors were associated with higher survival.[40,41] This has also been seen in head and neck squamous cell carcinoma,[42] suggesting that

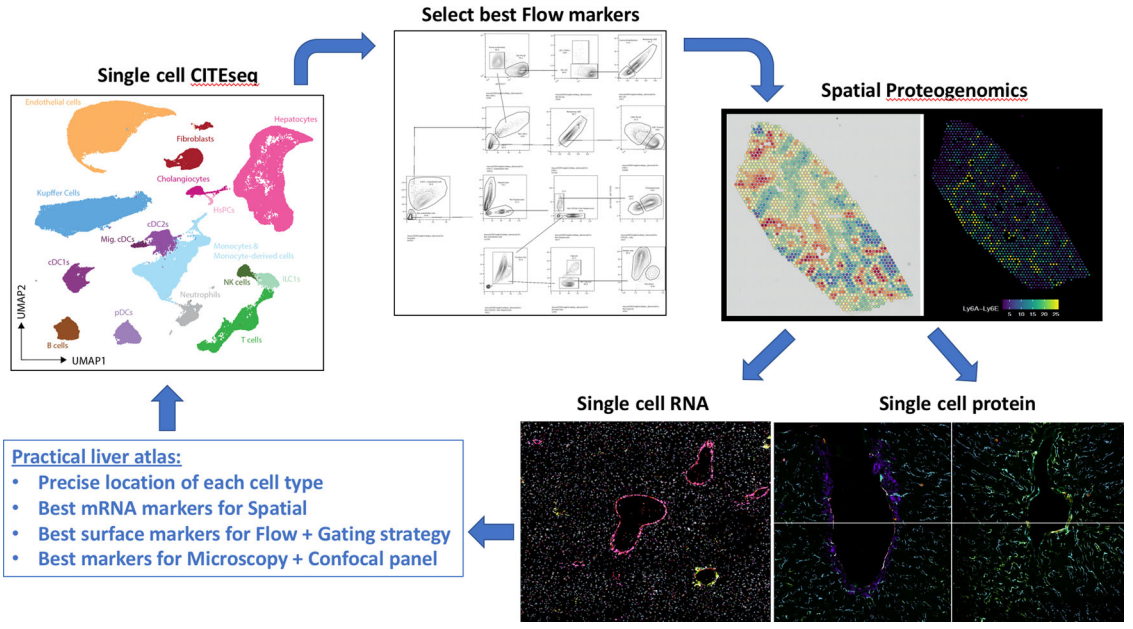

Figure 3. **A liver proteogenomics atlas is being developed using several single cell technologies, including CITE-seq and RNA expression to quantify both RNA and protein expression.**

this tumor–immune architecture may be relevant in several tumor types. Although the TME is generally thought of as more disorganized than normal tissue, Keren suggested that organization of cell types drives function within the TME. Keren's group is also working on developing methods to automatically identify microenvironments, as well as a neural network–based approach to embed multiplexed images, and determine microenvironments without the need for cell segmentation.

Martin Guilliams from Ghent University described a collaboration between his lab and Charlotte Scott's group to develop a proteogenomics liver atlas (Fig. 3), which is still a work in progress. Guilliams described some of the goals, methods, and expected offerings of the project. The two groups have begun developing a mouse liver atlas to aid in preclinical and basic research; they will also develop a human liver atlas to correlate clinical and mechanistic research. The proteogenomics atlas is being developed using several single cell technologies, including Cellular Indexing of Transcriptomes and Epitopes by Sequencing (CITE-seq)[43] and RNA expression to quantify both RNA and protein expression. Guilliams showed several examples where CITE-seq can detect proteins that are poorly detectable by mRNA. Using the algorithm TotalVI,[44] they also showed that clustering using mRNA and protein achieves better resolution than clustering using mRNA levels only. Guilliams stressed that the liver atlas will have practical uses for biological researchers. Not only will it provide information on the precise location of each cell type within the liver, it will also include the best mRNA markers for spatial analyses, the surface markers for flow cytometry–gating strategies, and the surface markers for microscopy.

Uri Alon from the Weizmann Institute of Science presented work of former graduate student Miri Adler together with Ruslan Medzhitov from Yale that yielded a theoretical model for understanding inflammation and fibrosis.[45] Alon's group uses a cell circuit approach to model fibrosis that includes fibroblasts and macrophages that support each other's growth via growth factors. Using this simplified circuit, the phase diagram reveals three steady states. If the system contains just a few macrophages and fibroblasts, they are unable to support long-term growth, resulting in a healing

phenotype. If the system contains an excess of fibroblasts that can support their own growth via an autocrine loop, it leads to a fibroblast-only fibrosis, which Alon has dubbed cold fibrosis. If the system contains enough macrophages and fibroblasts to support long-term growth, it leads to fibrosis containing both cell types, which Alon has dubbed hot fibrosis. The model captures the phase portrait of the *in vitro* coculture of fibroblasts and macrophages, and provides several insights on experimental observations on fibrosis. For example, the model predicts that a transient injury will lead to healing, whereas a prolonged or repeated injury will lead to fibrosis because of the continued flow of macrophages.[45] It also explains experimental observations that removing macrophages when there are few fibroblasts avoids fibrosis, while removing macrophages when there are many fibroblasts leads to fibrosis.[46–50] The model also predicts that fibroblasts can be prevented or reversed by weakening the fibroblast autocrine loop or cell proliferation, or by depleting macrophages.[45] Alon's group is working to experimentally validate some of these predictions in mouse models of heart fibrosis as well as in the TME. Alon concluded with a musical summary from the sound of the Beatles (https://www.youtube.com/watch?v=8CU1c-b4QHI).

Alex K. Shalek from MIT discussed a collaboration between their group, Boston Children's Hospital the University of Mississippi Medical School, and others using single cell genomics to better understand the cellular targets of SARS-CoV-2 and link clinical and molecular features. scRNA-seq data from a number of data atlases has revealed that cells that coexpress two membrane proteins key for viral entry, the ACE2 receptor and TMPRSS2, are quite rare within a set of epithelial barrier tissues.[51–53] To understand the *in vivo* viral targets in the tissue first encountered by the virus, the group performed scRNA-seq on cells isolated from nasopharyngeal swabs used for viral testing. Protocols to recover rare viable cells from frozen nasopharyngeal swabs are accessible to researchers.[a] The results revealed major differences in nasopharyngeal cell populations and trajectories in individuals with COVID-19. Infected individuals had an increase in secretory, deuterosomal, and developing ciliated cells and a decrease in mature ciliated cells. Shalek also showed that severe COVID-19 is characterized by a muted interferon (IFN) signature in ciliated epithelial cells,[54] a finding supported by other studies.[55–57] Codetection of host and viral RNA via scRNA-seq revealed that infected cells have increased expression of IFN-response genes and receptors used for viral entry, while bystander cells express genes associated with major histocompatibility complex class II presentation.[54] scRNA-seq data are available to researchers (at www.covid19cellatlas.org). Shalek's group is also profiling lung tissue in patients who died of COVID-19 to understand how severe infection affects lung cell populations and gene expression.[58]

Regan Hamel from Stefano Pluchino's lab at the University of Cambridge presented work using scRNA-seq to understand cell dynamics during spinal cord injury. Spinal cord injury induces persistent inflammation that does not resolve, unlike normal wound healing. Both microglia from the central nervous system and macrophages from the periphery infiltrate the site of injury. Infiltrating macrophages, which are generally short lived, can adopt the morphology and transcriptional program of long-lived microglia and drive persistent inflammation. Hamel used fate mapping in mice to isolate infiltrating macrophages and microglia at the site of injury at different time points and characterized them by scRNA-seq. Hamel showed that microglia can follow two trajectories after injury, a cycling trajectory and a trajectory in which cells transition through three states, first expressing proinflammatory genes, including FSBP5, then converting to a cytotoxic phenotype that negatively regulates canonical wound healing, and finally achieving a neuroprotective, anti-inflammatory phenotype. Microglia and infiltrating macrophages that express FSBP5 may disrupt the progression of normal wound healing after spinal cord injury.[59] Hamel is confirming the role of FSBP5 with other methods like single-molecule FISH and immunohistochemistry as well as functional studies.

Sarah J. Pfau from Chenghua Gu's lab at Harvard Medical School presented work on understanding how different cell types coordinate to

[a]www.protocols.io/view/human-nasopharyngealswab-processing-for-viable-si-bjhkkj4w.

regulate the blood–brain barrier (BBB).[60] The permeability of the BBB is heterogenous throughout the brain; this heterogeneity is important for the brain's function and its ability to communicate with the periphery.[61,62] Although the BBB is made up of endothelial cells, it is known that the brain environment contains factors that affect BBB properties and that brain cells such as pericytes and astrocytes are important for the formation and maintenance of the BBB.[63] Pfau is using scRNA-seq and 3D imaging to assess morphological and transcriptional differences between various cells types in regions of the brain with different BBB properties to see how local differences modulate barrier function. The hope is that understanding how regional BBB heterogeneity is achieved can provide insights on regional brain function and the development of more effective, potentially region-specific targeted therapies.

Computational approaches

The wealth of data generated by single cell technologies requires new computational approaches to integrate multichannel measurements that may provide information on transcriptomic, epigenetic, and protein expression across thousands of cells. Speakers presented new computational methods that can integrate single cell technologies to provide deeper insights on the relationship between molecular state and fate, increase the throughput of single cell technologies, achieve more accurate cell segmentation in imaging studies, and predict future cell states on the basis of static scRNA-seq data, as well as large-scale projects to develop whole-organism atlases.

Arjun Raj from the University of Pennsylvania described work to understand how nongenetic variability within cell populations can lead to drug resistance. Raj's group has developed a method called Rewind, which combines genetic barcoding and RNA FISH, to look at a cell's fate, for example, drug resistance and to determine the state that led to resistance. In short, cells are barcoded and allowed to divide. Twin barcoded cells are then separated; half are subjected to the drug to determine which are drug resistant. Drug-resistant cells are identified by RNA FISH probes, and the twin cells that have not been subjected to the drug can be characterized by a variety of single cell techniques to determine what it was about that cell that allowed it to be resistant. Using Rewind, Raj's group showed that the initial molecular state of a cell predetermines its phenotype. Raj showed that cells contain cryptic variability that only becomes apparent after a perturbation, like drug exposure.[64] Raj also presented unpublished data characterizing morphology, transcriptome, and phenotype across resistant cell colonies to classify resistant cells into different subtypes. Raj hopes that understanding how cells are primed for different fates will be applicable across a range of biological processes, including stem cell differentiation and induced pluripotent stem cell reprogramming.

Stephen R. Quake from Stanford University and the Chan Zuckerberg Biohub is working to create organism-wide single cell transcriptome atlases. Quake argued that while transcriptomics data essentially simplify the cell into a "bag" of mRNAs, providing little information on protein expression, activity, or localization, it has proven useful to provide insight on complex processes. For example, Quake's group has used single cell transcriptomics to understand long-term memory formation in mice[65] and the evolution of anatomical structures in the brain.[66] Quake is involved in several large-scale collaborative projects within the Biohub to establish organism-wide cell atlases. The Mouse Aging Cell Atlas characterizes eighteen tissues at various ages, providing insights on aging-related effects in shared cell types across tissues.[67,68] The group has also developed an atlas using heterochronic parabiosis, in which the circulatory systems of young and old mice are joined, revealing gene expression and molecular-level changes consistent with accelerated aging and rejuvenation.[68] The Biohub is also working to develop a fly cell atlas of approximately 400,000 cells and a few hundred cell types, and a mouse lemur atlas of approximately 250,000 cells and 30 tissues, which will be the first nonhuman primate atlas. Finally, they have recently begun work on a human atlas that will include all human cell types from 25 tissues from a single donor. Quake stressed that these projects are large-scale, collaborative efforts that require a range of expertise. The single cell transcriptome atlases are useful references for researchers to provide insight on how genes are used in various cell types.

Nancy R. Zhang from the University of Pennsylvania described her group's development of a new method, Alleloscope, to estimate the allele-specific copy number from scDNA-seq and the single cell assay for transposase-accessible chromatin sequencing (scATAC-seq) data. Copy number variation represents gains and losses of large portions of the genome and is a hallmark of many cancers. Most current single cell methods are limited to assessing total copy number and do not provide information on which allele is affected during large-scale deletions or duplications. Zhang argued that these methods therefore miss a lot of intratumor heterogeneity. For example, it has been observed that different haplotypes of a genome segment can be amplified across different regions of a tumor.[66] This intratumor heterogeneity would not be detected by looking at total copy number alone or by averaged allelic ratios across the tumor. Alleloscope estimates the allele-specific copy number for scDNA-seq and scATAC-seq data. Unlike other methods, it does not require external phasing information and can achieve high accuracy under low coverage. Alleloscope allows one to dissect the contribution of chromosomal instability (copy number variations) and epigenetic plasticity (chromatin accessibility changes) to intra-tumor heterogeneity. Zhang's group has applied Alleloscope to primary breast tumor and primary and metastasized colorectal and gastric tumors, revealing pervasive intratumor heterogeneity, including highly complex multiallelic copy number aberrations differentiated by haplotype ratios that have previously been underappreciated or ignored.[69]

Jean Fan from Johns Hopkins University presented work using computational modeling to infer changes in cellular state from spatially resolved transcriptomic imaging data. Fan's group applies multiplexed error-robust FISH (MERFISH), which uses combinatorial labeling, barcoding, and sequential imaging, to profile spatially resolved genome-wide transcriptomes in single cells within fixed cultures and tissues.[70] MERFISH provides information at both the intracellular level, revealing the organization of mRNAs within cells, as well as at the tissue level, revealing the organization of transcriptionally distinct cell types and states within tissues.[71] To explore the wealth of data produced by MERFISH, Fan's group developed an interactive web tool, MERmaid (available at https://jef.works/MERmaid/). Although such spatially resolved transcriptomic imaging data provide a fixed snapshot of the transcriptome, not all temporal information is lost. Fan's group uses RNA velocity *in situ* analysis to infer future cellular transcriptional states.[71,72] RNA velocity *in situ* analysis models the rate of change of cytoplasmic mRNA levels as a function of transcripts exported from the nucleus and the rate of degradation. For example, if a cell has more nuclear expression than expected at steady state, the model infers that the gene is being upregulated. RNA velocity analysis has been used to demonstrate changes in gene expression throughout the cell cycle, as well as during neurodevelopment.[72] However, approaches to visualize RNA velocities generally rely on projecting the observed and future cellular transcriptional states onto a two-dimensional (2D) embedding. Depending on which 2D embedding is chosen, different aspects of cellular dynamics may be featured. Fan's group has developed VeloViz to directly integrate cellular dynamics predicted from RNA velocity analysis in constructing 2D embeddings to visualize cellular trajectories. VeloViz has been used to visualize the differentiation of pancreatic ductal cells (Fig. 4) and is robust enough to predict global trajectories even if intermediate cell states are missing.[73] The software package and tutorials are available (at https://jef.works/veloviz/).

Bo Wang from Stanford University presented work using single cell sequencing to map cell-type evolution. Molecular evolution can blur cell-type relationships across species: genes are lost while others expand or duplicate, and regulatory networks expand and change throughout evolution. To address these challenges, Wang's group has developed self-assembling manifold mapping (SAMap), which mutually maps cell type and gene homology across species. Unlike other algorithms, which require one-to-one gene orthology, SAMap allows one-to-many homologs to allow for gene duplication events. SAMap is an extension of SAM, a robust manifold reconstruction algorithm.[74] SAMap accounts for molecular evolution during cell-type mapping and can identify gene substitution events, where genes exhibit more similar expression to their paralogs than their orthologs. SAMap was able to correctly recapitulate gene

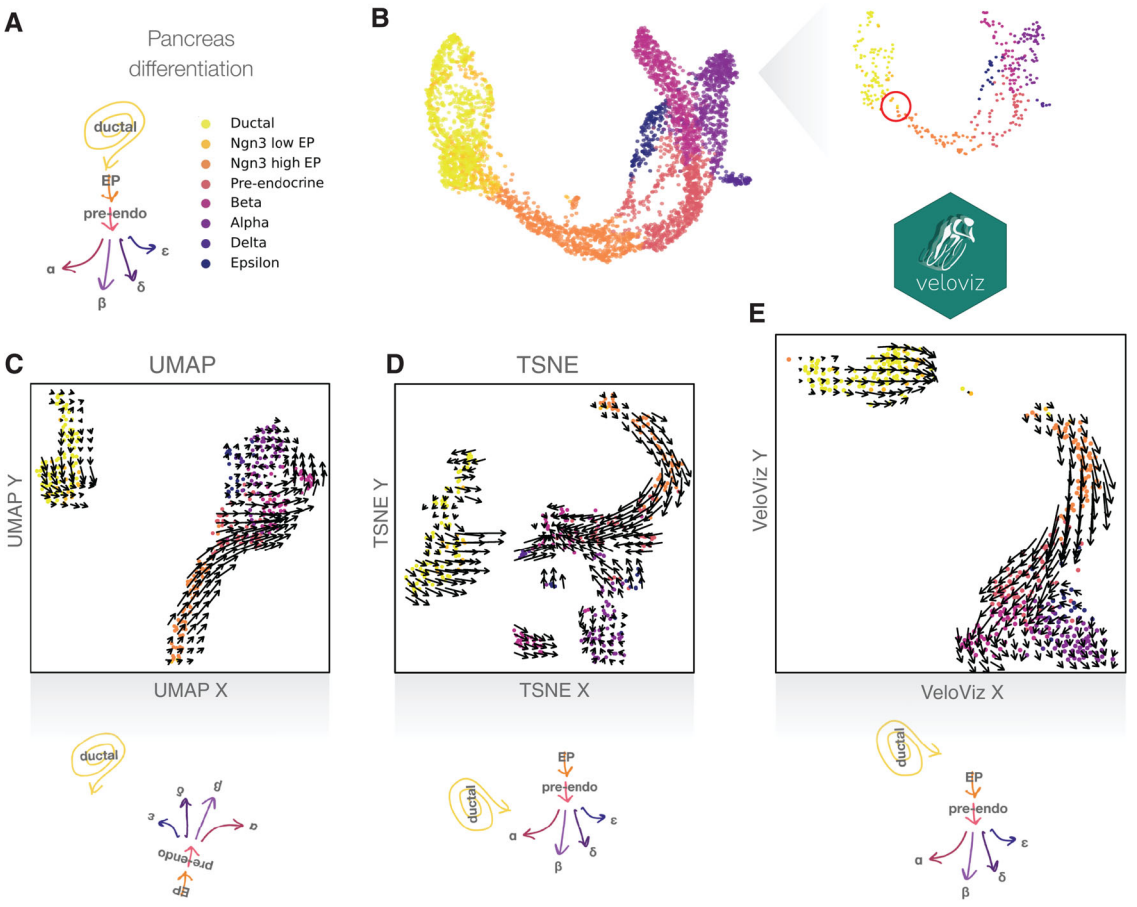

Figure 4. **VeloViz integrates RNA velocity information to visualize cellular trajectories. (A) Cartoon of the expected cellular trajectory for pancreatic endocrinogenesis (see https://doi.org/10.1242/dev.173849). (B) Single cell RNA-seq of the developing mouse pancreas visualized using UMAP and colored by cell type. Inset shows a subsample of the data where the number of differentiating endocrine progenitor cells is sparse (highlighted in the red circle). Visualization of the subsampled data using UMAP (C) and tSNE (D) results in two distinct clusters containing cells before and after the differentiation gap. Arrows show the projection of velocities onto the embeddings, suggesting that these clusters of cells are part of biologically distinct subpopulations rather than the same biological trajectory. A cartoon of the potential biological misinterpretation from such visualizations is shown below. (E) Visualization of the subsampled data using VeloViz with projected velocity arrows. By incorporating information about each cell's predicted future transcriptional state from RNA velocity, VeloViz can visualize relationships among cells that are more consistent with the expected cellular trajectory.**

ontogeny between zebrafish and frog, suggesting broad concordance between cell-type transcriptomic signatures and development origins. It also detected several homologous cell types arising from distinct developmental lineages, and even different germ layers (Fig. 5). Wang's group has subsequently successfully mapped 20 species using SAMap. Mapping cell atlases across long evolutionary distances reveals two distinct cell-type conservation patterns: (1) a one-to-one cell-type mapping, particularly between close animals, and (2) conserved families of related cell types sharing common gene expression programs that originate from a common ancestral state across distantly related species.[75]

Cole Trapnell from the University of Washington showed how Sci-Plex can be used to profile millions of cells from thousands of specimens. Trapnell's group has used Sci-Plex to follow embryonic

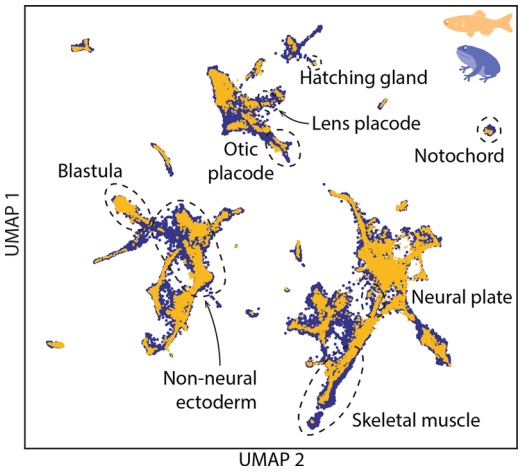

Figure 5. **Self-assembling manifold mapping (SAMap) analysis used to detect several homologous cell types arising from distinct developmental lineages and even different germ layers. Adapted from Ref. 75.**

development in zebrafish to better understand how genomes encode robust developmental programs. Looking at every cell type and gene in embryos across different stages of development is technically difficult to do with traditional scRNA-seq technologies. Sci-Plex can profile millions of cells from thousands of specimens. In short, embryos are deposited into wells and dissociated into a single-cell suspension; each is then labeled with an oligonucleotide and then data are combined and analyzed. The results reveal whole-genome transcriptome data. Trapnell's group has used Sci-Plex to understand how the abundance of each cell type varies across embryos and whether perturbations affect that variability.[76] Sci-Plex can accurately phenotype and stage individual embryos on the basis of cell-type frequencies and detect shifts in abundance and variance in cell types throughout the embryo. Using Sci-Plex, Trapnell's group has created a time-resolved atlas of zebrafish embryonic development that contains approximately 1 million cells from 859 embryos at 1 of 15 time points. They have also created atlases of embryos that were perturbed either genetically or environmentally to understand how these perturbations affect cell differentiation and the distribution of cell types within the embryo.

Noah F. Greenwald from Stanford University presented work on a new algorithm for cell segmentation of cell imaging data, Mesmer (Fig. 6). Single cell analyses of tissue data require single cell segmentation, which has been challenging owing to the density and heterogeneity of cells throughout tissues as well as the differences between imaging platforms, tissue types, and experimental conditions. Mesmer is a novel deep learning algorithm for cell segmentation that was specifically designed for tissue data. It uses a two-channel image and a deep learning model to predict the center and shape of each cell and nucleus. To train the deep learning model, Greenwald created TissueNet, the largest training dataset for cell segmentation. TissueNet contains 1 million paired annotations for nuclear and whole-cell data from a range of tissue types. Mesmer works across imaging platforms and tissue types and is more accurate than previous models, achieving human-level performance for cell segmentation.[77]

Tracking dynamics of single cells

Although some single cell technologies offer a snapshot of the cell populations present at a given time, many enable researchers to track the fate and differentiation of a specific cell or population over time during key events like differentiation and development. Speakers showed how cell states change with respect to the extracellular environment, how tracking individual cell response to drugs can provide insights for drug resistance, and how synchronicity can be achieved among cells.

Roser Vento-Tormo from the Wellcome Sanger Institute developed CellPhoneDB, a computational tool to identify cell–cell interactions from single cell data.[78,79] Updates to CellPhoneDB integrate differential expression of interacting ligands and receptors from scRNA-seq data with spatial transcriptomics to improve cell–cell interaction predictions.[79] CellPhone DB can inform how cellular microenvironments influence cellular identify and responses. Vento-Tormo's group is using the technology to study regulation of regeneration and differentiation in the human endometrium. Since not much is known about the human endometrium, Vento-Tormo's group first used scRNA-seq to profile the cell types in endometrial biopsies and full thickness uterine wall samples. They identified two novel SOX9^{+} epithelial populations (SOX9^{+}LGR5$^{+/-}$) enriched in the

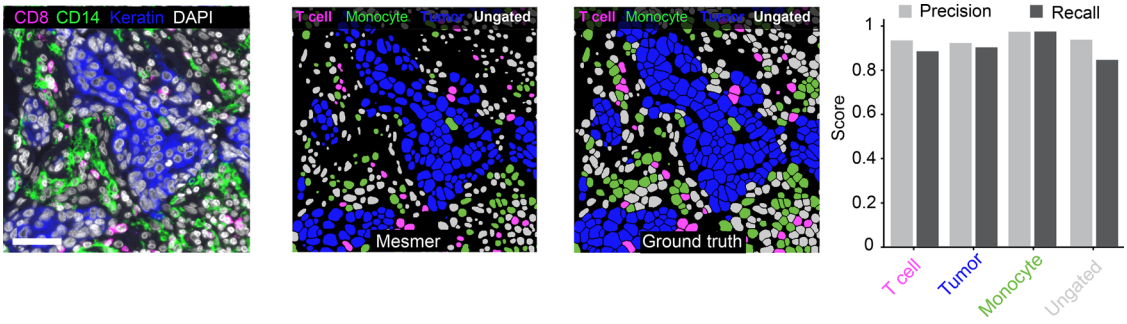

Figure 6. **Building a spatial proteogenomic atlas of the liver by combining single cell CITE-seq, single-nuclei sequencing, spatial transcriptomics, and spatial proteomics. By integrating these multiomic datasets, we aim to provide validated strategies to reliably discriminate and localize all hepatic cells.**

proliferative phase of the menstrual cycle. Spatial transcriptomics showed that the two SOX9$^+$ populations have distinct spatial positions: SOX9$^+$LGR5$^+$ populations in the lumen give rise to luminal cells, while SOX9$^+$LGR5$^-$ in the basal layer give rise to secretory cells.[80,81] The group is using CellPhone DB to study the signals from WNT and NOTCH that give rise to ciliary and secretory lineage. Vento-Tormo also showed that endometrial organoids derived from primary tissue accurately recapitulate the expected distribution of cell types by scRNA-seq as well as the expected hormone-induced changes in cell types. The organoids also validated the effects of WNT and NOTCH on cell differentiation: inhibiting NOTCH promoted ciliated cells, whereas inhibiting WNT promoted secretory cells. Using these results, Vento-Tormo's group has built a model of how cell states change across the menstrual cycle, as well as how the lumen and glandular environments differ with regard to WNT and NOTCH signaling to affect cell fate.

Silvia D.M. Santos from the Francis Crick Institute showed how her group combines experimental and computational methods to understand cell decision making, such as undergoing cell division or differentiation. Santos's group uses human embryonic stem cells (hESCs) as a model for understanding cell differentiation. Human ESCs can divide or differentiate into one of the three germ layers, endoderm, mesoderm, or ectoderm. Differentiation is driven by BMP4, which enables cells to differentiate into all three embryonic germ layers. Santos's group combines scRNA-seq and micropatterns of protein expression within an *in vitro* gastrulation assay to understand the spatial-temporal features of hESC differentiation. Using neural networks, they can track individual cells as they differentiate and undergo dramatic morphologic changes. Santos showed that hESCs commit to differentiation unexpectedly early, within minutes of BMP4 exposure, and well before morphological or expression changes are apparent. This early commitment is driven by BMP4-mediated SMAD signaling. After BMP4 stimulation, SMAD is irreversibly activated via a positive feedback loop, with cells at the periphery activated first. scRNA-seq revealed the genes affected by SMAD activations, including *GATA3*. *GATA3* was previously unknown to be a SMAD-responsive gene, but Santos showed that its expression dynamics mimic SMAD activation both temporally and spatially.[82] She presented unpublished work investigating whether *GATA3* is an early commitment gene for hESC differentiation.

Sabrina L. Spencer from the University of Colorado Boulder presented their work using real-time, long-term time-lapse microscopy to study cell proliferation at the single cell level. Her talk focused on work with melanoma cells and the rapid adaptation subsets of cells used to escape from the drug and to reenter the cell cycle. Although there has been a lot of research on sequencing late-stage resistant tumors in which cells acquire drug-resistant mutations that lead to tumor relapse, less is known about the quick, nongenetic adaptations that cells undergo within the first few days of treatment, which may represent the inception of drug resistance. Using a live-cell CDK2 activity sensor

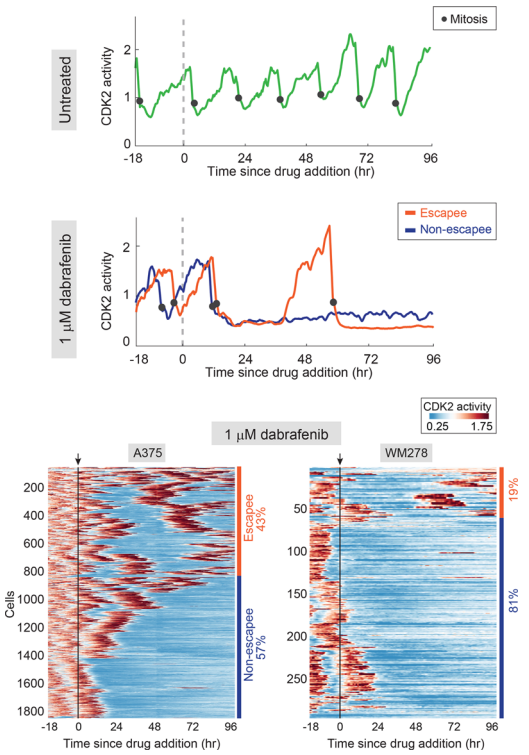

Figure 7. **Long-term live-cell microscopy of thousands of tumor cells shows that exposure to the BRAF inhibitor dabrafenib causes cells to enter a prolonged quiescence (blue). Subsets of cells escape drug-induced quiescence and resume proliferation after a couple of days (red). Increasing the drug dosage reduces the proportion of cells that can escape but does not eliminate it. From Ref. 85.**

that reports on cell proliferation and quiescence[83] along with EllipTrack, a new machine learning–based cell-tracking pipeline for long-term live-cell microscopy,[84] Spencer's group can visualize and track thousands of melanoma cells over many days. Exposure to the BRAF inhibitor dabrafenib causes melanoma cells to enter a prolonged quiescence, but subsets of cells are able to rapidly rewire to escape drug-induced quiescence and resume proliferation within 3 days. Although the proportion of cells that escape depends on the drug dose, even at high doses, Spencer's group was never able to eliminate this cycling subpopulation (Fig. 7). They further showed that the ability to escape is reversible upon drug withdrawal, suggesting that escape is caused by nongenetic mechanisms. Similar results were seen with other BRAF inhibitors across several melanoma cell lines as well as in patient biopsies. ScRNA-seq revealed that ATF4, which is involved in the integrated stress response, and p53, which is involved in the DNA damage response, are upregulated in escapee cells. Spencer's group then independently confirmed that escapees exhibit heightened DNA replication stress and DNA damage, yet are able to out-proliferate nonescapees, suggesting that they could be the seed population driving eventual acquisition of drug-resistance mutations.[85]

Hernan G. Garcia from the University of California Berkeley showed his lab's work to understand the spatiotemporal control of gene expression during development. Garcia's ultimate goal is to augment these gene networks with molecular and quantitative information to be able to predict phenotypes from these gene networks. Although a challenging task, Garcia and collaborators have developed theoretical models that successfully predict transcriptional regulation in bacteria.[86–88] Garcia's group is now moving to more complicated systems like *Drosophila*. Garcia focused on understanding expression of the gene *even-skipped* in the fruit fly embryo to demonstrate how single cells regulate their transcriptional dynamics to dictate macroscopic patterns in gene expression. *In vivo* imaging of *even-skipped* transcripts in living embryos showed that the promoters are turned on for short periods of time, resulting in transcriptional bursts rather than continuous transcription.[89–91]

In collaboration with Chris Wiggins's group at Columbia University, Garcia's group devised a computational model that infers promoter status from single cell fluorescence data showing the number of RNA polymerase molecules at the gene locus over time. The model has revealed that the transcriptional burst frequency and amplitude are regulated to control gene expression across the embryo.[92,93] The group is now working to understand how transcriptional bursting is controlled at the molecular level. Garcia's group is developing new data analysis methods to visualize activator binding events at a gene locus to eventually relate activator binding with transcriptional activity to understand what transcription factors are doing before, during, and after transcriptional bursts.

Geethika Arekatla from Timm Schroeder's lab at ETH Zurich presented unpublished work using optogenetics and time-lapse microscopy[94] to

understand how the dynamics of ERK activation affects embryonic stem cell (ESC) differentiation. ERK inhibition allows ESCs to maintain pluripotency. It has previously been shown that different dynamics in ERK signaling can result in different cell fates.[95] However, how ERK signaling dynamics affect differentiation is not well understood. Arekatla's system uses optoFGFR1 to quickly and reversible trigger ERK activation,[96] as well as a fluorescent ERK sensor to detect ERK activity. They showed how modifying the amplitude, duration, and frequency of ERK activation affects ESC differentiation.

Federico Gaiti from Dan Landau's lab at Weill Cornell Medicine presented unpublished work on a multiomics approach to understanding cancer evolution. The complexity of cancer cell states revealed by scRNA-seq is often independent of genetic diversity. Gaiti uses a single cell multiomics platform that captures genetic, epigenetic, and transcriptomic information within the same cell.[97] Using this platform, Gaiti has shown that stochastic, heritable DNA methylation changes (epimutations) can serve as a molecular clock and therefore can be exploited as native barcodes for high-resolution lineage tracing, as the team previously showed in chronic lymphocytic leukemia cells. Gaiti focused on a collaborative project with Mario Suva at Massachusetts General Hospital to understand epigenetic encoding of cell states in human glioblastoma samples, demonstrating heritability of malignant cell states, with key differences in hierarchal versus plastic cell state architectures in IDH-mutant and IDH-wildtype glioblastoma, respectively.

Rinat Arbel-Goren from Joel Stavans's lab at the Weizmann Institute of Science presented work on understanding circadian clock–controlled processes in multicellular filaments of *Anabaena* cyanobacteria (Fig. 8). Although circadian circuits have been described in unicellular cyanobacteria, less in known about circadian clocks in multicellular cyanobacteria. Approximately 100 genes have been shown to exhibit oscillatory behavior in *Anabaena*.[98] Arbel-Goren has studied one such gene, *pecB*, which codes for a subunit of a photosynthetic pigment. Visualizing *pecB* expression at the single cell level in real time via fluorescence showed that *pecB* exhibits oscillatory expression

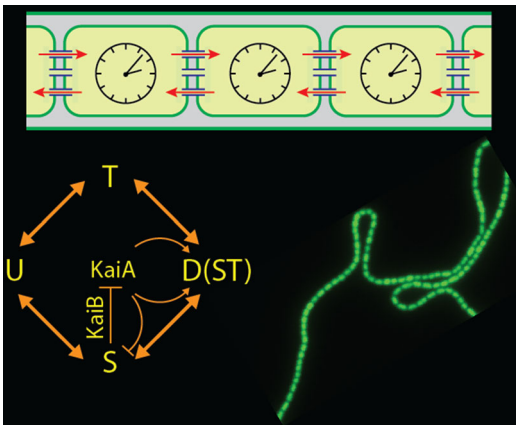

Figure 8. **Circadian clock–controlled processes in multicellular filaments of *Anabaena* cyanobacteria.**

patterns despite constant light conditions, suggesting an autonomous clock. *pecB* expression showed synchronization and spatial coherence along the multicellular filament despite cell-to-cell variability. Deleting the clock genes *kaiABC* abolished this oscillatory behavior and spatial coherence. Deleting proteins involved in cell–cell communication also reduced spatial coherence and synchronization. Arbel-Goren's work shows that circadian clocks in *Anabaena* are coupled via cell–cell communication, which allows oscillations in circadian genes to be highly synchronized and spatially coherent.[99]

Steffen Rulands from the Max Planck Institute presented work using methods from statistical physics to infer spatiotemporal processes from single cell sequencing data. As an example, Rulands focused on the establishment of the methylome that occurs early in development. To understand how the methylome is established, cells are cultured under conditions in which DNA methylation is erased and then released to conditions where DNA methylases are upregulated. The methylome, transcriptome, and local chromatin accessibility are assessed via whole-genome bisulfite sequencing and single cell NMT-sequencing. Rulands showed that although DNA methylation is established at different rates for different sites, these rates can be mathematically condensed, demonstrating that there is a single mechanism by which DNA methylation is established genome wide. Rulands's work highlights how analytical mathematical approaches from statistical physics can

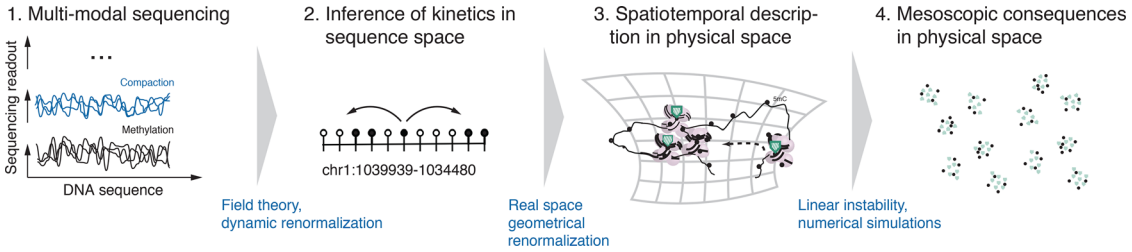

Figure 9. **Overview over a method to infer emergent, spatiotemporal phenomena from linear sequencing data using mathematical tools from theoretical physics (blue).**

complement computational tools in single cell genomics. Although they focused on the spatial and temporal rules that underlie the establishment of the embryonic methylome, the work can also be generalized to infer other emergent, spatiotemporal processes from single cell sequencing (Fig. 9).[100]

Fate tracing of single cells

The symposium ended with a session on fate tracing of single cells. Speakers demonstrated how lineage tracing can be used to delineate the origins of specific cell populations that arise because of injury or during development as well as how single cell technologies can be used to better understand cell reprogramming.

Lineage tracing

Philipp Junker from the Max Delbrück Center for Molecular Medicine (MDC) presented work on lineage tracing in the regenerating zebrafish heart. The zebrafish heart is a model for understanding the mechanisms and consequences of cell activation in complex organs. After injury, multiple cell types in the heart are activated and produce pro-regenerative factors while cardiomyocytes dedifferentiate and proliferate, resulting in efficient regeneration and revascularization. Junker's group, in collaboration with Daniela Panáková, also at MDC, is working to identify the transient cell states in the regenerative niche and the cellular drivers for regeneration. ScRNA-seq data of the zebrafish heart revealed transient increases in fibroblasts and immune cells as well as dedifferentiated cardiomyocytes after injury. Junker identified three distinct fibroblast populations present only in the injured heart as drivers of regeneration on the basis of their location, gene expression, and timing.[101] Junker's group combined RNA velocity[72] with CRISPR lineage tracing to reconstruct a lineage tree[102] of these pro-regenerative fibroblasts. The results revealed that two of the transient pro-regenerative fibroblasts are derived from the epicardium, while the other is derived from the endocardium. Inhibiting WNT signaling depleted endocardial fibroblasts while delaying regeneration and cardiomyocyte dedifferentiation, suggesting that the epicardial and endocardial fibroblasts may respond to perturbations differently.[101] Junker's work shows how combining CRISPR lineage tracing, trajectory analysis, and perturbation analysis can identify the origin of disease-specific cell types.

Allon M. Klein from Harvard Medical School presented work on developing CoSpar, a coherent, sparse optimization approach that infers transition maps from clonal data. CoSpar provides a way to order events in differentiation from lineage tracing and single cell sequencing, which can be difficult becasue of the high dimensionality and sparse, noisy nature of the data. Although single cell landscapes can generate hypotheses about the dynamics of differentiation, there are often ambiguities. For example, during a bifurcation event, it can be difficult to know whether uncommitted cells begin to commit when they approach a branch point or whether they are strongly committed earlier. CoSpar infers transition maps by determining the probability that a cell in a given state will transition into another state using lineage-tracing scRNA-seq data. To reduce ambiguity, the map is constrained with regard to sparseness (i.e., most cells are only able to access a few states) and coherence (i.e., cells in similar states will have similar fates). Klein showed that CoSpar is robust even with data paucity and can rescue sparsely sampled trajectories. And it has been benchmarked with ground truth datasets and

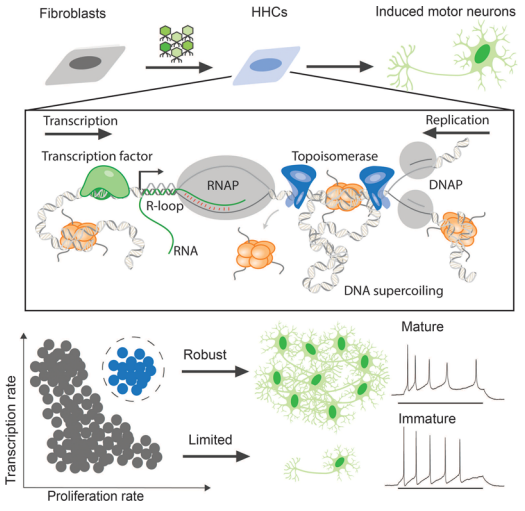

Figure 10. **A rare population of cells capable of high rates of transcription and proliferation reprogram at near-deterministic rates to generate more functionally mature induced motor neurons. Hypertranscribing/hyperproliferative cells (HHCs) mitigate genomic stress introduced by high rates of transcription and replication through the activity of topoisomerases. Topoisomerases resolve conflicts between transcription and replication machinery as well as curate DNA supercoils and R-loops to reduce transcription factor–induced genomic stress. Increasing the HHC population through a chemical and genetic cocktail increases reprogramming efficiency by two orders of magnitude. Adapted from Ref. 117.**

shown to predict heterogeneity associated with fate bias in the examples tested.[103]

John I. Murray from the University of Pennsylvania presented work on generating a gene expression map across *Caenorhabditis elegans* development. *C. elegans* is an ideal model for a single cell understanding of development because it follows a robust developmental process in which entire patterns of cell division have been mapped out.[104] This also makes *C. elegans* an excellent system for testing and validation of single cell methods. The Murray lab is using lineage-tracing methods to map gene expression of every gene in every cell during *C. elegans* development. Cell-tracking software tracks cells over time and quantifies gene expression within each cell.[105–107] Murray's group has generated single cell lineage-aligned expression data for over 250 genes.[106,108,109] Using scRNA-seq data, they have created an atlas of *C. elegans* embryos across development.[110] Such large-scale projects have revealed several lessons and themes, including several examples of transcriptional convergence, in which cells of different lineages converge to the same fate, multilineage priming, rapid transitions from lineage-correlated to cell type–correlated transcriptomes, as well as frequent discontinuities in projections between mother and daughter cells that correspond to rapid changes in gene expression in single cells.

Murray also described work on understanding how embryos achieve precise spatial, temporal, and dosage control in gene expression. Different species have evolved different mechanisms to achieve such transcriptional precision. Spatial averaging is seen in the fruit fly, while temporal averaging is seen in zebrafish. In *C. elegans* embryos, transcripts appear to accumulate rapidly to precise, high levels.[111,112] Murray's group is collaborating with Arjun Raj's group to determine absolute transcript levels in single cells during development, to understand how these levels are regulated, and to determine how defects in transcript dosage influence development.

Understanding cell reprogramming

Samantha A. Morris from Washington University in St. Louis and **Kate E. Galloway** from MIT presented work on understanding what features of a cell's state make it more likely to undergo reprogramming in order to use those insights to increase reprogramming efficiency. Morris's group has developed two new tools for understanding cell identity: Capybara, which measures cell identity and fate transitions from single cell data;[113] and CellOracle, which infers gene regulatory networks and predicts the outcomes of transcription factor perturbations.[114] CellOracle uses scRNA-seq and scATAC-seq to infer gene regulatory networks and then predicts how these networks change during reprogramming. They have used CellOracle to understand how transcription factors regulate cell identity during reprogramming by simulating the effects of transcription factor overexpression or knockout throughout a gene regulatory network to predict future gene expression values and the direction of future cell identity transition. CellOracle can also be used to link early cell states to their eventual fate. Morris showed that reprogramming mouse embryonic fibroblasts to hepatocytes is a very inefficient process.[115] Lineage tracing with CellTagging revealed two trajectories during the reprogramming process. A few, rare cells followed a reprogramming

trajectory while the majority followed a dead-end trajectory. CellTagging was able to identify which early cell states were predetermined to follow the reprogramming trajectory.[116] Characterizing the gene regulatory networks in the early cell states by CellOracle revealed that early networks are wired very differently from cells destined to reprogram and from those that follow dead-end trajectories. One key difference is that Fos is highly connected in gene regulatory networks in cells destined to reprogram. CellOracle simulations and experimental methods showed that Fos overexpression enhances reprogramming, moving cells from a dead-end state to a reprogramming state.[114,116]

Galloway's group has identified global increases in both proliferation and transcription as characteristics of reprogrammable cells. Among fibroblasts induced to reprogram into motor neurons, only a rare population of hypertranscribing/hyperproliferative cells (HHCs) reprogram at high rates. Expanding this population increases reprogramming efficiency. Galloway's group has worked to understand how this rare population develops, with the hope of using this insight to design genetic circuits that facilitate reprogramming. High transcription and replication induce high rates of genomic stress on a cell. ScRNA-seq showed that HHCs have high expression of topoisomerases, which reduce genomic stress by resolving R-loops and DNA supercoils induced by transcription and replication (Fig. 10). Galloway proposed that topoisomerases support HHCs by mitigating the stress of high transcription and replication rates that promote reprogramming.[117] She argued that this has implications for synthetic biology as the field expands beyond designs focused on logical functions to include how the physical arrangement of elements in circuits supports or impedes requisite transcriptional dynamics.[118,119] Understanding how the structure of genetic programs designed to promote reprogramming integrates into the chromatin structure will be key to ensuring their function.[120,121]

The meeting ended with talks highlighting unpublished work on understanding cell fate and differentiation.

Michael Ratz from Jonas Frisén's lab at the Karolinska Institute presented unpublished work on combining *in vivo* clonal tracking and scRNA-seq to understand neurogenesis in the mouse brain.[122] The approach utilizes ultrasound-guided *in utero* injections of complex lentivirus barcode libraries into the developing mouse brain to label progenitor cells. Ratz reconstructed thousands of clones to uncover the existence of fate-restricted progenitor cells in the early hippocampal neuroepithelium and showed that microglia are derived from few primitive myeloid precursors that massively expand to generate widely dispersed progeny. By coupling spatial transcriptomics with clonal barcoding, Ratz revealed migration patterns and gene expression of clonally related cells in densely labeled tissue sections. Compared to classical fate mapping, their approach enables high-throughput dense reconstruction of cell phenotypes and clonal relations at the single cell and tissue level in individual animals and provides an integrated approach for understanding tissue architecture.

Merrit Romeike from Christa Bücker's lab at Max Perutz Labs described work on deconvoluting effects due to differentiation delays from direct effects in cells with impaired differentiation.

Acknowledgments

M.S. Balzer is supported by German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) Grant BA 6205/2-1. M. Ratz is supported by a DFG Research Fellowship, Grant RA 2889/1-1. J. Fan is supported by the National Science Foundation, Grant No. 2047611.

Competing interests

The authors declare no competing interests.

References

1. Zhu, R., J.M. del Rio-Salgado, J. Garcia-Ojalvo, *et al.* 2021. Synthetic multistability in mammalian cells. bioRxiv. https://doi.org/10.1101/2021.02.10.430659.
2. Su, C.J., A. Murugan, J.M. Linton, *et al.* 2020. Ligand-receptor promiscuity enables cellular addressing. bioRxiv. https://doi.org/10.1101/2020.12.08.412643.
3. Ma, Y., M.W. Budde, M.N. Mayalu, *et al.* 2020. Synthetic mammalian signaling circuits for robust cell population control. bioRxiv. https://doi.org/10.1101/2020.09.02.278564.
4. Zeng, W., R.M. Pirzgalska, M.M.A. Pereira, *et al.* 2015. Sympathetic neuro-adipose connections mediate leptin-driven lipolysis. *Cell* **163:** 84–94.
5. Pirzgalska, R.M., E. Seixas, J.S. Seidman, *et al.* 2017. Sympathetic neuron-associated macrophages contribute to

obesity by importing and metabolizing norepinephrine. *Nat. Med.* **23:** 1309–1318.
6. Mahú, I., A. Barateiro, E. Rial-Pensado, *et al.* 2020. Brain-sparing sympathofacilitators mitigate obesity without adverse cardiovascular effects. *Cell Metab.* **31:** 1120–1135.e7.
7. Domingos, A.I. 2020. Leptin: a missing piece in the immunometabolism puzzle. *Nat. Rev. Immunol.* **20:** 3.
8. Habib, N., C. McCabe, S. Medina, *et al.* 2020. Disease-associated astrocytes in Alzheimer's disease and aging. *Nat. Neurosci.* **23:** 701–706.
9. Cain, A., M. Taga, C. McCabe, *et al.* 2020. Multi-cellular communities are perturbed in the aging human brain and with Alzheimer's disease. bioRxiv. https://doi.org/10.1101/2020.12.22.424084.
10. Moor, A.E. & S. Itzkovitz. 2017. Spatial transcriptomics: paving the way for tissue-level systems biology. *Curr. Opin. Biotechnol.* **46:** 126–133.
11. Moor, A.E., Y. Harnik, S. Ben-Moshe, *et al.* 2018. Spatial reconstruction of single enterocytes uncovers broad zonation along the intestinal villus axis. *Cell* **175:** 1156–1167.e15.
12. Baron, M., M. Tagore, M.V. Hunter, *et al.* 2020. The stress-like cancer cell state is a consistent component of tumorigenesis. *Cell Syst.* **11:** 536–546.e7.
13. Rao, A., D. Barkley, G.S. França, *et al.* 2021. Exploring tissue architecture using spatial transcriptomics. *Nature* **596:** 211–220.
14. Moncada, R., D. Barkley, F. Wagner, *et al.* 2020. Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat. Biotechnol.* **38:** 333–342.
15. Barkley, D. & I. Yanai. 2019. Plasticity and clonality of cancer cell states. *Trends Cancer* **5:** 655–656.
16. Lieberman-Aiden, E., N.L. van Berkum, L. Williams, *et al.* 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326:** 289–293.
17. Beagrie, R.A., A. Scialdone, M. Schueler, *et al.* 2017. Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* **543:** 519–524.
18. Quinodoz, S.A., N. Ollikainen, B. Tabak, *et al.* 2018. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* **174:** 744–757.e24.
19. Beliveau, B.J., E.F. Joyce, N. Apostolopoulos, *et al.* 2012. Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc. Natl. Acad. Sci. USA* **109:** 21301–21306.
20. Payne, A.C., Z.D. Chiang, P.L. Reginato, *et al.* 2021. *In situ* genome sequencing resolves DNA sequence and structure in intact biological samples. *Science* **371**. https://doi.org/10.1126/science.aay3446
21. van Steensel, B. & S. Henikoff. 2000. Identification of *in vivo* DNA targets of chromatin proteins using tethered dam methyltransferase. *Nat. Biotechnol.* **18:** 424–428.
22. Kind, J., L. Pagie, S.S. de Vries, *et al.* 2015. Genome-wide maps of nuclear lamina interactions in single human cells. *Cell* **163:** 134–147.
23. Altemose, N., A. Maslan, C. Rios-Martinez, *et al.* 2020. μDamID: a microfluidic approach for joint imaging and sequencing of protein-DNA interactions in single cells. *Cell Syst.* **11:** 354–366.e9.
24. Chen, T.N., A. Gupta, M.D. Zalavadia, *et al.* 2020. μCB-seq: microfluidic cell barcoding and sequencing for high-resolution imaging and sequencing of single cells. *Lab. Chip* **20:** 3899–3913.
25. Takei, Y., J. Yun, S. Zheng, *et al.* 2021. Integrated spatial genomics reveals global architecture of single nuclei. *Nature* **590:** 344–350.
26. Femino, A.M., F.S. Fay, K. Fogarty, *et al.* 1998. Visualization of single RNA transcripts *in situ*. *Science* **280:** 585–590.
27. Raj, A., P. van den Bogaard, S.A. Rifkin, *et al.* 2008. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **5:** 877–879.
28. Lubeck, E., A.F. Coskun, T. Zhiyentayev, *et al.* 2014. Single-cell *in situ* RNA profiling by sequential hybridization. *Nat. Methods* **11:** 360–361.
29. Burgess, D.J. 2019. Spatial transcriptomics coming of age. *Nat. Rev. Genet.* **20:** https://doi.org/10.1038/s41576-019-0129-z.
30. Shah, S., Y. Takei, W. Zhou, *et al.* 2018. Dynamics and spatial genomics of the nascent transcriptome by intron seqFISH. *Cell* **174:** 363–376.e16.
31. Eng, C.-H.L., M. Lawson, Q. Zhu, *et al.* 2019. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature* **568:** 235–239.
32. Stergachis, A.B., B.M. Debo, E. Haugen, *et al.* 2020. Single-molecule regulatory architectures captured by chromatin fiber sequencing. *Science* **368:** 1449–1454.
33. Akhtar, W., A.V. Pindyurin, J. de Jong, *et al.* 2014. Using TRIP for genome-wide position effect analysis in cultured cells. *Nat. Protoc.* **9:** 1255–1281.
34. Serra, D., U. Mayr, A. Boni, *et al.* 2019. Self-organization and symmetry breaking in intestinal organoid development. *Nature* **569:** 66–72.
35. Lukonin, I., D. Serra, L. Challet Meylan, *et al.* 2020. Phenotypic landscape of intestinal organoid regeneration. *Nature* **586:** 275–280.
36. Bryant, D.M., K. Sousounis, D. Payzin-Dogru, *et al.* 2017. Identification of regenerative roadblocks via repeat deployment of limb regeneration in axolotls. *NPJ Regen. Med.* **2:** 30.
37. Sousounis, K., D.M. Bryant, Jez. Fernandez, *et al.* 2020. Eya2 promotes cell cycle progression by regulating DNA damage response during vertebrate limb regeneration. *eLife* **9:** e51217.
38. Leigh, N.D., G.S. Dunlap, K. Johnson, *et al.* 2018. Transcriptomic landscape of the blastema niche in regenerating adult axolotl limbs at single-cell resolution. *Nat. Commun.* **9:** 5153.
39. Johnson, K., J. Bateman, T. DiTommaso, *et al.* 2018. Systemic cell cycle activation is induced following complex tissue injury in axolotl. *Dev. Biol.* **433:** 461–472.
40. Keren, L., M. Bosse, S. Thompson, *et al.* 2019. MIBI-TOF: a multiplexed imaging platform relates cellular phenotypes and tissue structure. *Sci. Adv.* **5**. https://doi.org/10.1126/sciadv.aax5851.
41. Keren, L., M. Bosse, D. Marquez, *et al.* 2018. A structured tumor-immune microenvironment in triple negative breast

cancer revealed by multiplexed ion beam imaging. *Cell* **174:** 1373–1387.e19.
42. Blise, K.E., S. Sivagnanam, G.L. Banik, *et al.* 2021. Single-cell spatial proteomics analyses of head and neck squamous cell carcinoma reveal tumor heterogeneity and immune architectures associated with clinical outcome. bioRxiv. https://doi.org/10.1101/2021.03.10.434649.
43. Stoeckius, M., C. Hafemeister, W. Stephenson, *et al.* 2017. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14:** 865–868.
44. Gayoso, A., Z. Steier, R. Lopez, *et al.* 2021. Joint probabilistic modeling of single-cell multi-omic data with totalVI. *Nat. Methods* **18:** 272–282.
45. Adler, M., A. Mayo, X. Zhou, *et al.* 2020. Principles of cell circuits for tissue repair and fibrosis. *iScience* **23:** 100841.
46. Duffield, J.S., S.J. Forbes, C.M. Constandinou, *et al.* 2005. Selective depletion of macrophages reveals distinct, opposing roles during liver injury and repair. *J. Clin. Invest.* **115:** 56–65.
47. Ide, M., M. Kuwamura, T. Kotani, *et al.* 2005. Effects of gadolinium chloride (GdCl(3)) on the appearance of macrophage populations and fibrogenesis in thioacetamide-induced rat hepatic lesions. *J. Comp. Pathol.* **133:** 92–102.
48. Pradere, J.-P., J. Kluwe, S. De Minicis, *et al.* 2013. Hepatic macrophages but not dendritic cells contribute to liver fibrosis by promoting the survival of activated hepatic stellate cells in mice. *Hepatology* **58:** 1461–1473.
49. Sunami, Y., F. Leithäuser, S. Gul, *et al.* 2012. Hepatic activation of IKK/NFκB signaling induces liver fibrosis via macrophage-mediated chronic inflammation. *Hepatology* **56:** 1117–1128.
50. Wynn, T.A. & L. Barron. 2010. Macrophages: master regulators of inflammation and fibrosis. *Semin. Liver Dis.* **30:** 245–257.
51. Ziegler, C.G.K., S.J. Allon, S.K. Nyquist, *et al.* 2020. SARS-CoV-2 receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is detected in specific cell subsets across tissues. *Cell* **181:** 1016–1035.e19.
52. Muus, C., M.D. Luecken, G. Eraslan, *et al.* 2021. Single-cell meta-analysis of SARS-CoV-2 entry genes across tissues and demographics. *Nat. Med.* **27:** 546–559.
53. Lukassen, S., R.L. Chua, T. Trefzer, *et al.* 2020. SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. *EMBO J.* **39:** e105114.
54. Ziegler, C.G.K., V.N. Miao, A.H. Owings, *et al.* 2021. Impaired local intrinsic immunity to SARS-CoV-2 infection in severe COVID-19. bioRxiv. https://doi.org/10.1101/2021.02.20.431155.
55. Combes, A.J., T. Courau, N.F. Kuhn, *et al.* 2021. Global absence and targeting of protective immune states in severe COVID-19. *Nature* **591:** 124–130.
56. Bastard, P., L.B. Rosen, Q. Zhang, *et al.* 2020. Autoantibodies against type I IFNs in patients with life-threatening COVID-19. *Science* **370**. https://doi.org/10.1126/science.abd4585.
57. Zhang, Q., P. Bastard, Z. Liu, *et al.* 2020. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* **370**. https://doi.org/10.1126/science.abd4570.
58. Delorey, T.M., C.G.K. Ziegler, G. Heimberg, *et al.* 2021. A single-cell and spatial atlas of autopsy tissues reveals pathology and cellular targets of SARS-CoV-2. bioRxiv. https://doi.org/10.1101/2021.02.25.430130.
59. Hamel, R., L. Peruzzotti-Jametti, K. Ridley, *et al.* 2020. Time-resolved single-cell RNAseq profiling identifies a novel Fabp5-expressing subpopulation of inflammatory myeloid cells in chronic spinal cord injury. bioRxiv. https://doi.org/10.1101/2020.10.21.346635.
60. Pfau, S.J., U.H. Langen, T.M. Fisher, *et al.* 2021. Vascular and perivascular cell profiling reveals the molecular and cellular bases of blood–brain barrier heterogeneity. bioRxiv. https://doi.org/10.1101/2021.04.26.441465
61. Ganong, W.F. 2000. Circumventricular organs: definition and role in the regulation of endocrine and autonomic function. *Clin. Exp. Pharmacol. Physiol.* **27:** 422–427.
62. Kaur, C. & E.-A. Ling. 2017. The circumventricular organs. *Histol. Histopathol.* **32:** 879–892.
63. O'Brown, N.M., S.J. Pfau & C. Gu. 2018. Bridging barriers: a comparative look at the blood–brain barrier across organisms. *Genes Dev.* **32:** 466–478.
64. Emert, B.L., C. Cote, E.A. Torre, *et al.* 2020. Variability within rare cell states enables multiple paths towards drug resistance. bioRxiv. https://doi.org/10.1101/2020.03.18.996660.
65. Chen, M.B., X. Jiang, S.R. Quake, *et al.* 2020. Persistent transcriptional programmes are associated with remote memory. *Nature* **587:** 437–442.
66. Lui, J.H., N.D. Nguyen, S.M. Grutzner, *et al.* 2021. Differential encoding in prefrontal cortex projection neuron classes across cognitive tasks. *Cell* **184:** 489–506.e26.
67. Tabula Muris Consortium, Overall coordination, Logistical coordination, *et al.* 2018. Single-cell transcriptomics of 20 mouse organs creates a *Tabula Muris*. *Nature* **562:** 367–372.
68. Tabula Muris Consortium, A.O. Pisco, A. McGeever, *et al.* 2020. A single cell transcriptomic atlas characterizes aging tissues in the mouse. bioRxiv. https://doi.org/10.1101/661728.
69. Wu, C.-Y., B.T. Lau, H.S. Kim, *et al.* 2021. Alleloscope: integrative single cell analysis of allele-specific copy number alterations and chromatin accessibility in cancer. bioRxiv. https://doi.org/10.1101/2020.10.23.349407.
70. Chen, K.H., A.N. Boettiger, J.R. Moffitt, *et al.* 2015. RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348:** aaa6090.
71. Xia, C., J. Fan, G. Emanuel, *et al.* 2019. Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression. *Proc. Natl. Acad. Sci. USA* **116:** 19490–19499.
72. La Manno, G.. , R. Soldatov, A. Zeisel, *et al.* 2018. RNA velocity of single cells. *Nature* **560:** 494–498.
73. Atta, L., A. Sahoo & J. Fan. 2021. VeloViz: rNA-velocity informed embeddings for visualizing cellular trajectories. bioRxiv. https://doi.org/10.1101/2021.01.28.425293.
74. Tarashansky, A.J., Y. Xue, P. Li, *et al.* 2019. Self-assembling manifolds in single-cell RNA sequencing data. *eLife* **8:** e48994. https://doi.org/10.7554/eLife.48994.

75. Tarashansky, A.J., J.M. Musser, M. Khariton, *et al.* 2021. Mapping single-cell atlases throughout Metazoa unravels cell type evolution. *eLife* **10**. e66747.
76. Srivatsan, S.R., J.L. McFaline-Figueroa, V. Ramani, *et al.* 2020. Massively multiplex chemical transcriptomics at single-cell resolution. *Science* **367:** 45–51.
77. Greenwald, N.F., G. Miller, E. Moen, *et al.* 2021. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. bioRxiv. https://doi.org/10.1101/2021.03.01.431313.
78. Vento-Tormo, R., M. Efremova, R.A. Botting, *et al.* 2018. Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563:** 347–353.
79. Efremova, M., M. Vento-Tormo, S.A. Teichmann, *et al.* 2020. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat. Protoc.* **15:** 1484–1506.
80. Wang, W., F. Vilella, P. Alama, *et al.* 2020. Single-cell transcriptomic atlas of the human endometrium during the menstrual cycle. *Nat. Med.* **26:** 1644–1653.
81. Garcia-Alonso, L., L.-F. Handfield, K. Roberts, *et al.* 2021. Mapping the temporal and spatial dynamics of the human endometrium *in vivo* and *in vitro*. bioRxiv. https://doi.org/10.1101/2021.01.02.425073.
82. Gunne-Braden, A., A. Sullivan, B. Gharibi, *et al.* 2020. GATA3 mediates a fast, irreversible commitment to BMP4-driven differentiation in human embryonic stem cells. *Cell Stem Cell* **26:** 693–706.e9.
83. Spencer, S.L., S.D. Cappell, F.-C. Tsai, *et al.* 2013. The proliferation-quiescence decision is controlled by a bifurcation in CDK2 activity at mitotic exit. *Cell* **155:** 369–383.
84. Tian, C., C. Yang & S.L. Spencer. 2020. Elliptrack: a global-local cell-tracking pipeline for 2D fluorescence time-lapse microscopy. *Cell Rep*. **32:** 107984.
85. Yang, C., C. Tian, T.E. Hoffman, *et al.* 2021. Melanoma subpopulations that rapidly escape MAPK pathway inhibition incur DNA damage and rely on stress signalling. *Nat. Commun.* **12:** 1747.
86. Garcia, H.G. & R. Phillips. 2011. Quantitative dissection of the simple repression input-output function. *Proc. Natl. Acad. Sci. USA* **108:** 12173–12178.
87. Brewster, R.C., F.M. Weinert, H.G. Garcia, *et al.* 2014. The transcription factor titration effect dictates level of gene expression. *Cell* **156:** 1312–1323.
88. Weinert, F.M., R.C. Brewster, M. Rydenfelt, *et al.* 2014. Scaling of gene expression with transcription-factor fugacity. *Phys. Rev. Lett.* **113:** 258101.
89. Garcia, H.G., M. Tikhonov, A. Lin, *et al.* 2013. Quantitative imaging of transcription in living *Drosophila embryos* links polymerase activity to patterning. *Curr. Biol.* **23:** 2140–2145.
90. Lucas, T., T. Ferraro, B. Roelens, *et al.* 2013. Live imaging of Bicoid-dependent transcription in *Drosophila embryos*. *Curr. Biol.* **23:** 2135–2139.
91. Bothma, J.P., H.G. Garcia, E. Esposito, *et al.* 2014. Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living *Drosophila embryos*. *Proc. Natl. Acad. Sci. USA* **111:** 10598–10603.
92. Lammers, N.C., V. Galstyan, A. Reimer, *et al.* 2020. Multimodal transcriptional control of pattern formation in embryonic development. *Proc. Natl. Acad. Sci. USA* **117:** 836–847.
93. Berrocal, A., N.C. Lammers, H.G. Garcia, *et al.* 2020. Kinetic sculpting of the seven stripes of the *Drosophila* even-skipped gene. *eLife* **9**. e61635.
94. Hilsenbeck, O., M. Schwarzfischer, S. Skylaki, *et al.* 2016. Software tools for single-cell tracking and quantification of cellular and molecular properties. *Nat. Biotechnol.* **34:** 703–706.
95. Ryu, H., M. Chung, M. Dobrzyński, *et al.* 2015. Frequency modulation of ERK activation dynamics rewires cell fate. *Mol. Syst. Biol.* **11:** 838.
96. Kim, N., J.M. Kim, M. Lee, *et al.* 2014. Spatiotemporal control of fibroblast growth factor receptor signals by blue light. *Chem. Biol.* **21:** 903–912.
97. Gaiti, F., R. Chaligne, H. Gu, *et al.* 2019. Epigenetic evolution and lineage histories of chronic lymphocytic leukaemia. *Nature* **569:** 576–580.
98. Kushige, H., H. Kugenuma, M. Matsuoka, *et al.* 2013. Genome-wide and heterocyst-specific circadian gene expression in the filamentous *Cyanobacterium Anabaena* sp. strain PCC 7120. *J. Bacteriol.* **195:** 1276–1284.
99. Arbel-Goren, R., V. Buonfiglio, F. Di Patti, *et al.* 2021. Robust, coherent, and synchronized circadian clock-controlled oscillations along *Anabaena* filaments. *eLife* **10:** e64348. https://doi.org/10.7554/eLife.64348.
100. Olmeda, F., T. Lohoff, S.J. Clark, *et al.* 2021. Inference of emergent spatio-temporal processes from single-cell sequencing reveals feedback between *de novo* DNA methylation and chromatin condensates. bioRxiv. https://doi.org/10.1101/2020.12.30.424823.
101. Hu, B., S. Lelek, B. Spanjaard, *et al.* 2021. Cellular drivers of injury response and regeneration in the adult zebrafish heart. bioRxiv. https://doi.org/10.1101/2021.01.07.425670.
102. Spanjaard, B., B. Hu, N. Mitic, *et al.* 2018. Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. *Nat. Biotechnol.* **36:** 469–473.
103. Wang, S.-W. & A.M. Klein. 2021. Learning dynamics by computational integration of single cell genomic and lineage information. bioRxiv. https://doi.org/10.1101/2021.05.06.443026.
104. Sulston, J.E., E. Schierenberg, J.G. White, *et al.* 1983. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100:** 64–119.
105. Bao, Z., J.I. Murray, T. Boyle, *et al.* 2006. Automated cell lineage tracing in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* **103:** 2707–2712.
106. Murray, J.I., Z. Bao, T.J. Boyle, *et al.* 2008. Automated analysis of embryonic gene expression with cellular resolution in *C. elegans*. *Nat. Methods* **5:** 703–709.
107. Richards, J.L., A.L. Zacharias, T. Walton, *et al.* 2013. A quantitative model of normal *Caenorhabditis elegans* embryogenesis and its disruption after stress. *Dev. Biol.* **374:** 12–23.
108. Murray, J.I. & Z. Bao. 2012. Automated lineage and expression profiling in live *Caenorhabditis elegans* embryos.

Cold Spring Harb. Protoc. https://doi.org/10.1101/pdb.prot070615.
109. Araya, C.L., T. Kawli, A. Kundaje, *et al.* 2014. Regulatory analysis of the *C. elegans* genome with spatiotemporal resolution. *Nature* **512:** 400–405.
110. Packer, J.S., Q. Zhu, C. Huynh, *et al.* 2019. A lineage-resolved molecular atlas of *C. elegans* embryogenesis at single-cell resolution. *Science* **365:** eaax1971. https://doi.org/10.1126/science.aax1971.
111. Nair, G., T. Walton, J.I. Murray, *et al.* 2013. Gene transcription is coordinated with, but not dependent on, cell divisions during *C. elegans* embryonic fate specification. *Development* **140:** 3385–3394.
112. Raj, A., S.A. Rifkin, E. Andersen, *et al.* 2010. Variability in gene expression underlies incomplete penetrance. *Nature* **463:** 913–918.
113. Kong, W., Y.C. Fu & S.A. Morris. 2020. Capybara: a computational tool to measure cell identity and fate transitions. bioRxiv. https://doi.org/10.1101/2020.02.17.947390.
114. Kamimoto, K., C.M. Hoffmann & S.A. Morris. 2020. CellOracle: dissecting cell identity via network inference and in silico gene perturbation. bioRxiv. https://doi.org/10.1101/2020.02.17.947416.
115. Morris, S.A., P. Cahan, H. Li, *et al.* 2014. Dissecting engineered cell types and enhancing cell fate conversion via CellNet. *Cell* **158:** 889–902.
116. Biddy, B.A., W. Kong, K. Kamimoto, *et al.* 2018. Single-cell mapping of lineage and identity in direct reprogramming. *Nature* **564:** 219–224.
117. Babos, K.N., K.E. Galloway, K. Kisler, *et al.* 2019. Mitigating antagonism between transcription and proliferation allows near-deterministic cellular reprogramming. *Cell Stem Cell* **25:** 486–500.e9.
118. Johnstone, C.P. & K.E. Galloway. 2021. Engineering cellular symphonies out of transcriptional noise. *Nat. Rev. Mol. Cell Biol.* 22: 369–370.
119. Wang, N.B., A.M. Beitz & K.E. Galloway. 2020. Engineering cell fate: applying synthetic biology to cellular reprogramming. *Curr. Opin. Syst. Biol.* **24:** 18–31.
120. Johnstone, C.P., N.B. Wang, S.A. Sevier, *et al.* 2020. Understanding and engineering chromatin as a dynamical system across length and timescales. *Cell Syst*. **11:** 424–448.
121. Desai, R.V., X. Chen, B. Martin, *et al.* 2021. A DNA-repair pathway can affect transcriptional noise to promote cell fate transitions. *Science* **373:** eabc6506. https://doi.org/10.1126/science.abc6506.
122. Ratz, M., L. von Berlin, L. Larsson, *et al.* 2021. Cell types and clonal relations in the mouse brain revealed by single-cell and spatial transcriptomics. bioRxiv. https://doi.org/10.1101/2021.08.31.458418.