How a cell decides its own fate: a single-cell view of molecular mechanisms and dynamics of cell-type specification

Mircea, M., Semrau, S.

Biochemical Society Transactions (2021) **49** 2509–2525
https://doi.org/10.1042/BST20210135

PORTLAND PRESS

Review Article

How a cell decides its own fate: a single-cell view of molecular mechanisms and dynamics of cell-type specification

Maria Mircea and Stefan Semrau

Leiden Institute of Physics, Leiden University, Leiden, The Netherlands

Correspondence: Stefan Semrau (semrau@physics.leidenuniv.nl)

On its path from a fertilized egg to one of the many cell types in a multicellular organism, a cell turns the blank canvas of its early embryonic state into a molecular profile fine-tuned to achieve a vital organismal function. This remarkable transformation emerges from the interplay between dynamically changing external signals, the cell's internal, variable state, and tremendously complex molecular machinery; we are only beginning to understand. Recently developed single-cell omics techniques have started to provide an unprecedented, comprehensive view of the molecular changes during cell-type specification and promise to reveal the underlying gene regulatory mechanism. The exponentially increasing amount of quantitative molecular data being created at the moment is slated to inform predictive, mathematical models. Such models can suggest novel ways to manipulate cell types experimentally, which has important biomedical applications. This review is meant to give the reader a starting point to participate in this exciting phase of molecular developmental biology. We first introduce some of the principal molecular players involved in cell-type specification and discuss the important organizing ability of biomolecular condensates, which has been discovered recently. We then review some of the most important single-cell omics methods and relevant findings they produced. We devote special attention to the dynamics of the molecular changes and discuss methods to measure them, most importantly lineage tracing. Finally, we introduce a conceptual framework that connects all molecular agents in a mathematical model and helps us make sense of the experimental data.

Introduction

What is a cell type, anyway? Traditionally, cell types have been defined by their function within an organism: neurons process and transmit information, macrophages remove harmful microorganisms, and podocytes are crucial for blood filtration in the kidney. As function can be difficult to ascertain, especially for subtle variants of cell types, cell morphology and the presence of certain marker genes are often used as proxies [1,2]. With the advent of single-cell omics technologies, cell types have increasingly come to be identified with their molecular profiles. While most cell types persist over long periods of time, often the entire life span of an adult organism, cells are found in short-lived, transient states such as different phases of the cell cycle, different metabolic states or multiple forms of the stress response. Here, we are only concerned with the specification of cell types, which occurs during embryonic development or regeneration of adult tissues. During development, pluripotent embryonic cells differentiate into progenitors with diminishing developmental potential, and eventually fully specified cell types [1,3–5]. In adult tissue, long-lived adult stem cells give rise to multiple types of descendants. These processes are collectively termed differentiation. Differentiation involves changes in gene expression (i.e. messenger RNA and protein levels), which are accompanied and guided by epigenetic changes. Broadly, the epigenetic profile of a cell encompasses any heritable

Received: 13 August 2021
Revised: 6 November 2021
Accepted: 8 November 2021

Version of Record published:
2 December 2021

Biochemical Society Transactions (2021) **49** 2509–2525
https://doi.org/10.1042/BST20210135

molecular mark, with the exception of changes in the DNA sequence [6,7]. Two of the most important epigenetic marks are DNA methylation and histone modifications. These marks are tightly linked to the accessibility of the DNA and thus influence the expression of specific genes. A comprehensive introduction to epigenetics can be found in [8].

Here, we will review few of the many cell-autonomous molecular mechanisms that make differentiation a reproducible process and ensure the long-term stability of cell types. Importantly, cells do not develop in isolation. Their communication with neighboring cells via chemical and mechanical signaling is an integral part of embryonic development and tissue regeneration, which we will not discuss here (for recent reviews see [9,10]). Equally important, but also outside the scope of this review, is the role of molecular noise, which can drive cell-type decisions but must also be controlled to ensure the stability of the fully differentiated state (for recent reviews see [11,12]). In this review, we will first introduce some of the most important molecular players, which can be used to define a cell type and discuss omics techniques that can measure molecular profiles comprehensively in single cells. We will then focus on the dynamics of differentiation and novel methods that allow the inference of the developmental lineage tree. Finally, we will discuss challenges arising in the analysis of data sets comprising multiple modalities and a conceptual framework that enables a quantitative understanding of differentiation (Figure 1).

Molecular embodiment of a cell type

Most, if not all, cell-type decisions involve specific transcription factors (TFs) [1,2,4,5,13,14]. These DNA binding proteins control a gene's transcription level by binding to *cis*-regulatory elements (CREs) in the DNA. Enhancers, CREs that can be found at large distances from the regulated gene, play a particularly important role in cell-type determination. Enhancers work in concert and physically interact with promoters, another type of CRE that is usually found near the regulated gene. TF binding of CREs not only depends on the presence of specific DNA sequence motifs but is also strongly modulated by the configuration of the chromatin (the complex of DNA, nucleosomes, and other associated proteins, see Figure 1, left). With the exception of so-called pioneer factors [15], TFs only bind accessible, nucleosome-free DNA [16]. Chromatin configuration

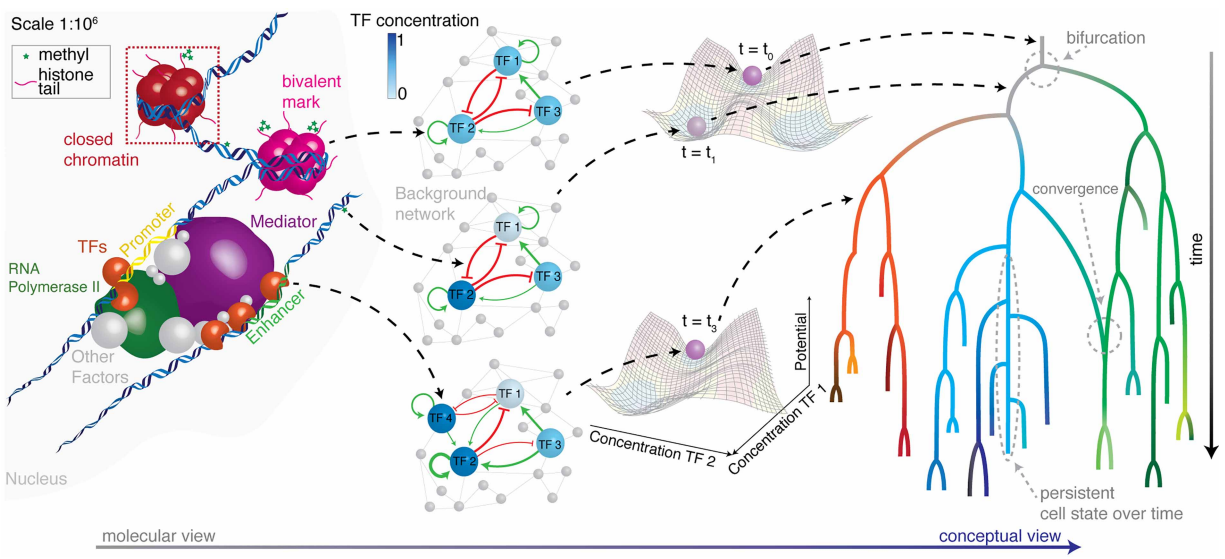

Figure 1. Conceptual framework for cell types and differentiation dynamics.

Left: TFs bind to enhancer and promoter regions depending on the chromatin state. All molecules are drawn at a scale of 1:10^6. Center left: A simplified gene regulatory network consisting of TFs (nodes of the network graph) and their interactions (edges of the network graph). Center right: A gene regulatory network can be modeled by a dynamical system, which can be represented by a potential energy landscape. The position in the landscape is determined by TF abundance. The shape of the landscape depends on the TF interactions. Cells follow the path of steepest descent to stable states, which correspond to cell types. Right: The collection of trajectories in the potential landscape form a cell-type decision tree, which highlights the hierarchical nature of differentiation.

and TF binding are affected by chemical modifications of histones (the components of a nucleosome), as well as the DNA [16]. Different histone modifications, or marks, are associated with different functions, broadly categorized as active and repressive, and the effect of DNA modifications strongly depends on the genomic context [17,18]. For example, DNA methylation at the enhancer regions of the pluripotency gene Sox2 results in silencing of its expression in embryonic stem cells [18]. Importantly, the interaction between TFs and chromatin configuration is reciprocal: TFs recruit enzymes that locally change the molecular make-up of the chromatin [16]. Both histone marks and DNA methylation are heritable molecular marks, as they are copied to the newly synthesized DNA during cell division [7,8,19]. They can therefore function as long-term memory of a cell's molecular profile and hence cell type. The pattern of chromatin accessibility and epigenetic marks can thus be used to identify a cell type and reveal relevant CREs [16,20].

Importantly, cell-type specification cannot be understood by studying individual TFs or epigenetic features in isolation. Cell types rather emerge from the complex interactions of several TFs. The presence of particular subsets of TFs has therefore been used to define a periodic table of cell types [4]. Together with their target genes, TFs form gene regulatory networks that establish and maintain cell identity [2,4], see Figure 1, middle. Regulatory interactions between TFs, in particular negative feedback loops, are crucial for the stability of molecular states. Due to the presence of fluctuations in the environment as well as the internal state of the cell, robustness is an important requirement for regulatory networks. At the same time, they need to be dynamic and react appropriately to external signaling inputs [1]. Mutual repression of TFs is one mechanism by which multiple, alternative cell types can be created. A prominent example is an interaction between the TFs GATA6 and NANOG, which governs the lineage decision between two of the earliest cell types in the mammalian embryo [21–23]. The conceptual framework discussed in the final section of this review explains how various stable cell types and unidirectional differentiation dynamics emerge from gene regulatory networks, see Figure 1, right.

Despite the fact that TFs always work in concert, some have a particularly large impact on lineage decisions: overexpression of certain TFs can revert a differentiated cell back to a pluripotent state (reprogramming) or convert one cell type into another (transdifferentiation) [24]. The remarkable power of these TFs, termed master TFs or master regulators, can be rationalized by their DNA binding patterns. Master TFs have been shown to bind clusters of enhancers, or super-enhancers, which drive high levels of key cell-type-specific genes [25,26], see Figure 2. Super-enhancers owe their special role to a high density of co-localized Mediator complex [25,26], a protein complex that links TF binding to the recruitment of the transcription machinery and

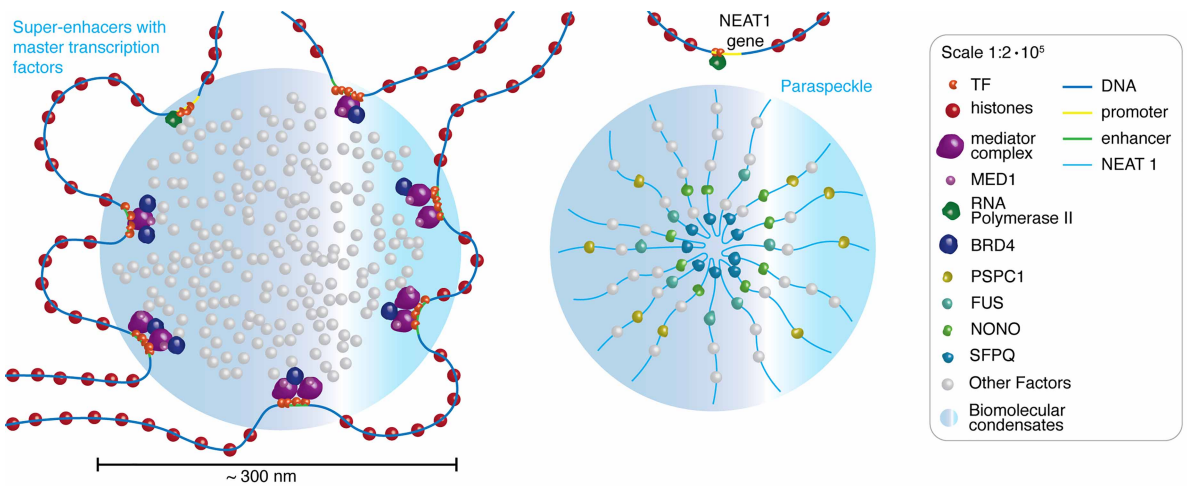

Figure 2. Master transcription factors and super-enhancers play a major role in guiding cell-type specification and are compartmentalized by biomolecular condensates.

Left: super-enhancer with bound master transcription factors in a large biomolecular condensate. Right: A paraspeckle that forms in the presence of the lncRNA NEAT1. All molecules are drawn at a scale of $1:2 \times 10^5$.

Biochemical Society Transactions (2021) 49 2509–2525
https://doi.org/10.1042/BST20210135

therefore gene expression. A well-studied example of master TFs that bind cell-type-specific super-enhancers are regulators of the pluripotent state in embryonic stem cells: NANOG, SOX2, and OCT4 [26].

TFs, CREs, epigenetic marks, and enzymes that modify chromatin state are just a small, albeit important, subset of the many molecular species that are involved in cell-type specification. It has long been unclear, how all of these mobile molecules, some of which are freely diffusing in the nucleus, can interact in an efficient manner. Recently, biomolecular condensates, which form through liquid–liquid phase separation (LLPS) [27–29], have been suggested as a possible answer to this question. Biomolecular condensates form according to well-known thermodynamic principles as a result of multivalent, homotypic interactions between molecules [30]. The high concentration of several molecular species in the condensed phase leads to increased interaction rates [29]. Examples of biomolecular condensates are the well-known membrane-less organelles, such as the nucleolus or Cajal bodies, as well as paraspeckles and many more [27,28,31]. It has been found that intrinsically disordered regions (IDRs) of proteins can lead to multivalent interactions that can cause condensates to form [32–34]. Interestingly, MED1, a member of the Mediator complex, and BRD4, a coactivator of transcription, have large IDRs and form condensates at super-enhancers [32], see Figure 2. Thus, phase-separated condensates likely concentrate components of the transcription apparatus and thereby ensure robust transcription of key cell-type-specific genes. Additionally, the large size of the Mediator cluster at super-enhancers enables contact with multiple promoter sites [35]. Therefore, biomolecular condensates are likely of crucial importance for establishing a cell type.

Molecular profiling

In recent years, omics technologies have emerged that measure one or multiple molecular species comprehensively in single cells (see Box 1 and Figure 3 for a selection of common methods). These technologies can reveal cell-type-specific molecular profiles in high throughput. With single-cell RNA sequencing (scRNA-seq), the transcriptomes of individual cells can be obtained [1,4,5], which enables the identification of new cell types and cell states in complex tissues [36]. Multiple large consortia are currently generating transcriptional atlases of entire organisms (reviewed in [37]). The human cell atlas [38] and Tabula Muris [39] are two prominent examples. Notwithstanding the great value of scRNA-seq measurements, gene expression should ideally be measured at the protein level. Numerous regulatory mechanisms at the translational and post-translational level make mRNA abundance just a proxy for protein abundance. Protein measurements have indeed revealed phenotypic features that could not be discerned with scRNA-seq alone [40,41].

As mentioned before, chromatin state is an important factor in gene regulation. Knowledge of the chromatin landscape can therefore improve the identification of cell types [42]. There is a growing variety of single-cell methods that measure chromatin features. For example, by using scATAC-seq [43,44], which reveals accessible chromatin regions in single cells, it is possible to identify cell-type-specific regulatory elements and candidate master TFs. By combining scATAC-seq and scRNA-seq, open chromatin regions can be associated with active transcription, which improves the identification of TFs and target genes compared with pure RNA measurements. It was also shown that considering the chromatin state of distal CREs significantly increased the power to predict cell-type specific gene expression, compared with using promoter chromatin state alone [45]. Similarly, a simultaneous measurement of histone modifications and transcriptome showed that active enhancers are epigenetically more variable across cell types than promoter regions [46]. A related finding resulted from the simultaneous measurement of DNA methylation and transcriptome (scM&T-seq [47]). The authors confirmed that promoter DNA methylation in mouse ESCs is typically correlated with reduced gene expression. In contrast, DNA methylation of distal enhancers is more often correlated with increased gene expression, compared with promoters.

Since active transcription typically requires the physical proximity of enhancers and promoters, knowledge of chromatin organization can be helpful to understand cell-type decision making.

scHi-C is a high-throughput method to reveal chromatin interactions throughout the genome in single cells [48,49]. In combination with DNA methylation measurements, cell-type-specific chromatin conformations can be obtained [50], which might help to clarify the role of biomolecular condensates [51,52]. In a recent study, a new variant of Hi-C [52] was used to determine the stability of chromatin interactions, which were revealed to vary substantially between organelles. Approaches to measure the spatial distribution of transcripts [53,54] and proteins [55] with sub-cellular resolution might lead to an even better understanding of cellular compartmentalization through biomolecular condensates.

Biochemical Society Transactions (2021) **49** 2509–2525
https://doi.org/10.1042/BST20210135

Box 1: Common single-cell omics and multi-omics techniques

Transcriptomics: The currently most prevalent single-cell omics method is single-cell RNA sequencing, which measures RNA abundance. Different experimental implementations of this method include: Smart-seq2 [127], Drop-seq [128], CEL-seq2 [129], and Sci-RNA-seq [130].

Proteomics: It is not yet feasible to measure every protein in a single cell. Antibody-based methods, for example, scCyTOF [41], can measure hundreds of proteins, but cannot easily be scaled to the whole proteome and rely on the existence of highly specific antibodies. Mass spectrometry-based proteomics methods, which recently became available, might soon produce high-quality proteomes of single cells. One example is SCoPE-MS [131], which detects around 1000 proteins per cell. Improvements to this method have been recently made in SCoPE2 [132] and another method [133]. By sequencing of DNA-tagged antibodies, the quantification of hundreds of proteins together with the transcriptome in single cells is possible with CITE-seq [134] and REAP-seq [135]. RAID [136] uses RNA-tagged antibodies for the same purpose.

Epigenomics: To gain insights into chromatin accessibility, one of the most prominent techniques is scATAC-seq [43,44], which uses transposons to barcode accessible DNA. scATAC-seq can be performed simultaneously with scRNA-seq, which was implemented, for example, by sci-CAR [45] and Paired-seq [88]. DNA methylation is measured by scBS-seq [137,138] and scRRBS [139]. Single-cell 5hmC-seq measures DNA hydroxymethylation [140]. Joint measurements of DNA methylation and the transcriptome are possible with, for example, scM&T-seq [47] and scMT-seq [141]. A method to measure all three molecular profiles (DNA methylation, transcriptome, and chromatin accessibility) is scNMT-seq [142]. Histone modifications can be measured, for example, with scChIP-seq [143] and scCUT&Tag [144]. New methods to measure the transcriptome jointly with histone modifications are CoTech [90] and PairedTag [46]. It is now also possible to study the chromatin conformation in every single cell with scHi-C [48,49]. Recent methods have allowed the capture of both chromatin conformation and DNA methylation (sn-m3C-seq [50] and methyl-HiC [145]).

Genomics: The DNA sequence of single cells can be measured by methods such as MALBAC [146] and NanoSeq [147]. NanoSeq has been designed to detect even small somatic mutations in single DNA molecules. Measurement methods that combine DNA sequencing with transcriptomics are, for example, G&T-seq [148] and TARGET-seq [149].

Multi-omics single-cell methods, like those presented here, promise to enable an improved mechanistic understanding of cell-type-specific gene regulation [56]. A more comprehensive discussion of these methods can be found in [57].

Dynamics of differentiation

During differentiation, the molecular profile of a cell is remodeled substantially. TFs are, unsurprisingly, important drivers of this transformation. As the majority of TFs bind to accessible chromatin regions, differentiation is accompanied by pervasive changes in chromatin accessibility [12,15,16]. One underlying mechanism involves pioneer TFs, which bind to nucleosome-associated DNA and create an open chromatin state [15,16]. These TFs can explore nucleosomal DNA through non-specific and transient binding, which in turn allows partial opening of the chromatin and other, non-pioneering factors to bind [15]. This mechanism has been recently validated, for example, for the pioneer factor PAX7 [58]. Another mechanism is a passive competition of TFs for DNA binding during short periods of local chromatin opening, which increases and stabilizes with higher TF concentrations [16,59].

Chromatin state is also influenced by chromatin remodelers that are recruited by TFs [15,16,60] and bind to different histone marks [60,61]. This is one mechanism by which epigenetic marks strongly impact chromatin accessibility. Importantly, activating and repressing histone marks can also occur simultaneously, on the same nucleosome. These bivalent domains play a particularly important role in cell-type decisions [62] and are more

Biochemical Society Transactions (2021) **49** 2509–2525
https://doi.org/10.1042/BST20210135

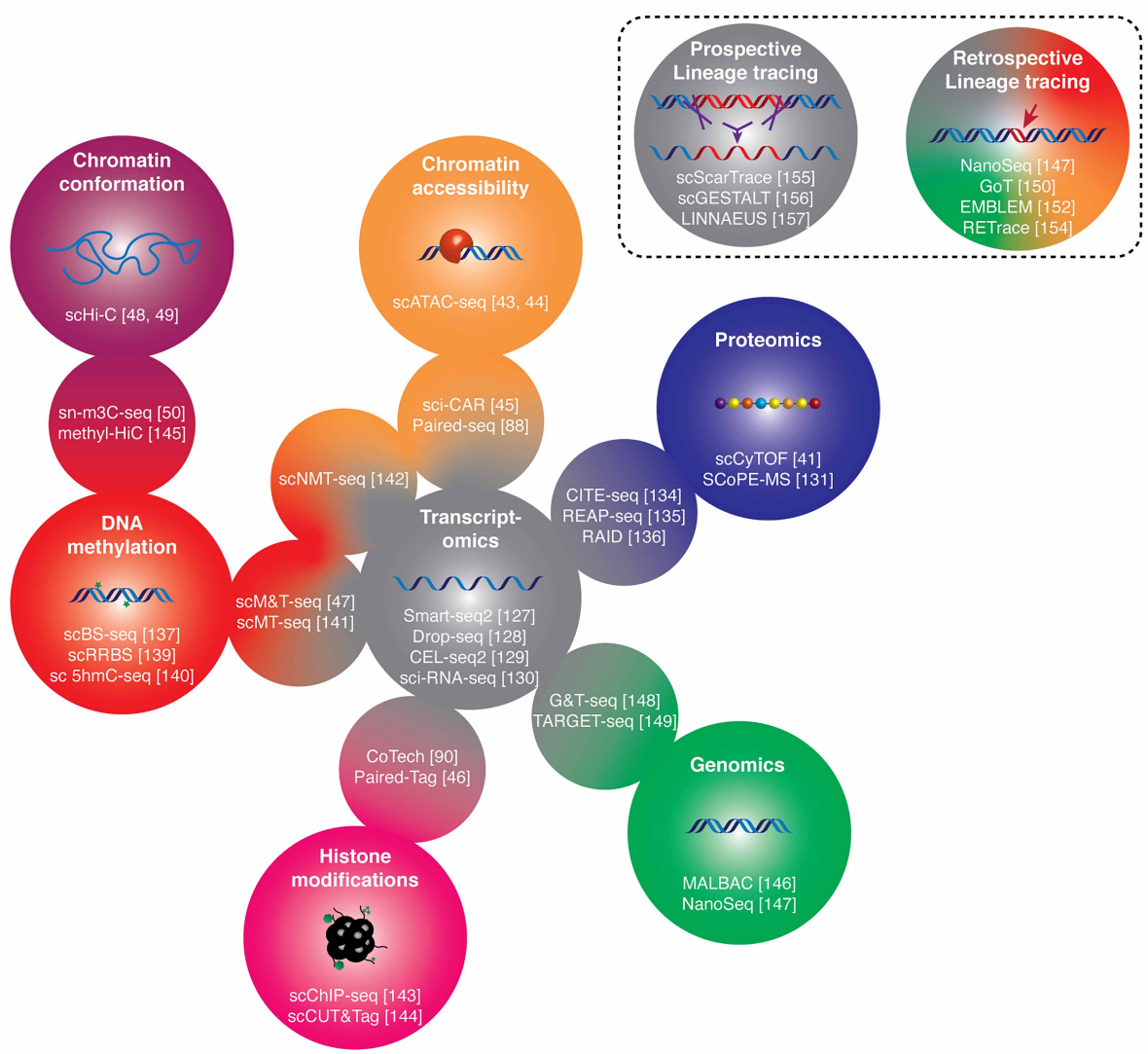

Figure 3. A selection of single-cell omics and multi-omics techniques useful for studying cell-type specification.
Colors indicate the measurement of different molecular species in a cell (green: DNA sequence, grey: RNA abundance, blue: protein abundance, orange: chromatin accessibility, red: DNA methylation, pink: histone modifications, violet: chromatin conformation). Circles with color gradients contain techniques that measure several types of molecules at the same time. Dashed lines envelop techniques that aim at understanding similar concepts.

abundant in embryonic stem cells (ESCs) than in adult tissues. A prominent example is the combination of H3K4me3 (trimethylation of histone H3 on lysine 4) which is associated with active transcription and H3K27me3 (trimethylation of histone H3 on lysine 27), which causes chromatin compaction and is thus a repressive mark. It has been shown that bivalent domains are positioned at key TF genes that are important for development [62–64]. Enhancers and promoters with bivalent marks are thought to be in a poised state that can be quickly resolved to either activation or repression. This effect can be mediated by multiprotein complexes composed of polycomb group (PcG) proteins. These proteins cause gene silencing by, for example, catalyzing the methylation of H3K27 [65,66]. In ESCs, the occupancy of PcG proteins at bivalent histone marks can change during differentiation, which results in altered gene expression [64,67]. Poised enhancers have been found to be necessary, for example, for the differentiation into specific neural cell types [68]. The examples mentioned here are only a few of many epigenetic mechanisms that drive dynamic chromatin remodeling during stem cell differentiation (reviewed in [69]).

Epigenetic marks are not homogeneously distributed in the nucleus, but rather need to be localized at important regulatory sites, which might be promoted by biomolecular condensates. The formation of biomolecular condensates during differentiation has been linked to different long non-coding RNAs (lncRNA) and RNA-binding proteins (RBPs) [31,70]. For example, the lncRNA DIGIT forms biomolecular condensates together with the RNA binding protein BRD3, which contains an IDR [70]. BRD3 is recruited to sites of the activating histone mark H3K18ac (acetylation of histone H3 on lysine 18). Paraspeckles are another important class of biomolecular condensates defined by the presence of the lncRNA NEAT1, which recruits several RBPs [71], see Figure 2. These condensates can, for example, influence transcriptional regulation via associated RBPs [71–73]. NEAT1 has been found to physically interact with EZH2, a PcG protein, which is involved in catalyzing histone methylation [74]. Interestingly, paraspeckles were found to be involved in slowing down the differentiation process and their number changes dynamically during differentiation to several lineages [31,72]. Another example of dynamic transcriptional regulation through biomolecular condensates is the association of RNA polymerase II with Mediator condensates (see Figure 2). It has been found that, upon phosphorylation, RNA polymerase II transitions from condensates involved in transcription initiation to condensates involved in RNA splicing at genes associated with super-enhancers [34].

Measurement of differentiation dynamics

In simple organisms, the entire lineage tree can be assembled using a microscope [75]. In larger organisms that becomes unfeasible and scRNA-seq data of developing tissues have been used instead to infer lineage relationships [76]. Due to asynchrony in embryonic development or regeneration of adult tissues, a single scRNA-seq measurement can capture cells in different stages of differentiation [42,77] and developmental order, or pseudotime, can be inferred by computational methods (reviewed in [78], see also the section on data analysis below). If the developmental process is sufficiently accessible for repeated sampling, scRNA-seq measurements at several time points can be used to resolve developmental dynamics [13,79–82]. This approach improves the temporal resolution and revealed that cells with different lineage histories can converge to globally similar cells [82]. However, combining multiple data sets to infer the correct developmental trajectory is challenging.

Lineage reconstruction has also been performed based on protein measurements in single cells at different time points. In a recent study [83], 27 proteins, of which 16 were TFs, were measured over a time course of 22 days during hematopoietic differentiation. This study showed that, at the protein level, cell-type decisions are accompanied by gradual changes in lineage-specific TFs, as no abrupt switches in TF levels were observed.

To reveal the gene regulatory programs that cause gene expression changes, chromatin conformation measurements during development can be used. Bulk methods have been used extensively to measure epigenetic changes and chromatin accessibility of cell populations [84], which produced many important insights. However, cell-to-cell variability and rare cell populations can only be distinguished with single-cell methods. Therefore, time-resolved single-cell chromatin accessibility measurements can be very informative [85], in particular in combination with transcriptomics [45,86–88]. One study found a class of genes with a high number of putative enhancers whose chromatin accessibility is predictive of gene expression [87]. These genes are enriched in TFs that regulate cell-type-specific gene expression. These findings suggest participation in super-enhancers and a central role in cell-type specification. Additionally, it was observed that the expression of TFs precedes the accessibility of their target sites, which might indicate a causal role of TFs in chromatin remodeling, possibly through additional epigenetic mechanisms [86].

Another interesting case is the simultaneous measurement of chromatin accessibility, DNA methylation, and transcriptome (scNMT [89]) at several time points in mouse development. The authors were able to study the dynamic changes of all three profiles in time and confirmed ectoderm, one of the three embryonic germ layers, as the default developmental pathway.

Specific histone marks have also been measured during differentiation and development. For example, the co-occurrence of H3K4me3 and H3K27me3 (bivalent mark) was measured in mouse ESCs together with scRNA-seq. The authors calculated a bivalency score along an RNA-based pseudotime trajectory and were able to classify genes by trends in bivalency dynamics [90]. A similar method found a significant overlap between H3K27ac (acetylation of histone H3 on lysine 27) and H3K27me3 in the adult mouse brain at CREs related to forebrain development [46].

An entirely different approach to study developmental dynamics is used in lineage tracing techniques [91–93] (see Box 2), which aim to find the correct phylogenetic tree [94,95] from pluripotent cells to fully specified cell

Biochemical Society Transactions (2021) **49** 2509–2525
https://doi.org/10.1042/BST20210135

Box 2: Lineage tracing

There are two, conceptionally distinct approaches to lineage tracing: retrospective and prospective. In retrospective lineage tracing, lineage relationships are inferred from naturally occurring somatic mutations. These mutations can be traced using DNA sequencing methods [147]. In a recent study, such mutations were linked with scRNA-seq data to investigate clonal relationships and cell types in humans [150]. Mitochondrial DNA has a ∼10 fold higher mutation rate than nuclear DNA [151,152], which makes it a good candidate for retrospective lineage tracing. Interestingly, these mutations can be tracked with ATAC-seq measurements because mitochondrial DNA is accessible [152]. DNA methylation also undergoes stochastic changes during cell division known as epimutations, which allow tracking of lineage histories through measurements of DNA methylation [151,153]. Coupling genomics to DNA methylation measurements allows both lineage tracing and the study of cell-type-specific methylation patterns [154]. However, naturally occurring mutations are rare, which requires highly accurate and sensitive measurement techniques and computational methods. In prospective lineage tracing, heritable markers are introduced that are read out at a later time point. The most recently developed dynamic lineage tracing methods insert 'scars' into the DNA at random or pre-determined locations, resulting in a large variety of different markers [91,92]. In some cases, these markers, or barcodes, are also transcribed, so that scRNA-seq is able to capture transcriptomes and lineage information simultaneously. Different omics technologies have been used in the context of lineage tracing (see Box 1 for a list of omics techniques). For retrospective lineage tracing, NanoSeq [147] has been used to track even small somatic mutations and GoT [150] linked transcriptomics to genotyping. scATAC-seq has been used to track mutations in mitochondrial DNA [152] and scRRBS has been used together with DNA sequencing to track DNA mutations together with DNA methylation [154]. Examples of prospective lineage tracing techniques that use transcriptomics measurements are scScarTrace [155], scGESTALT [156] and LINNAEUS [157].

types. Lineage tracing methods have produced a large number of valuable insights. A recent study used lineage tracing to reveal early biases toward particular cell types [96] that are not resolved with transcriptomics: transcriptionally similar cells were found to be committed to particular cell types prior to the divergence of their transcriptional profiles [91,97]. Importantly, such cells can easily be mistaken for multipotent progenitors. Coupling lineage tracing with epigenomics or proteomics measurements might help to avoid some of these biases and pinpoint the correct sequence of transcriptional and epigenetic changes during development. Lineage tracing experiments also seem to indicate that cell fate decisions occur in a more continuous manner rather than abruptly, as previously believed [98]. Finally, lineage tracing made it possible to observe the convergence of differentiation trajectories from distinct developmental origins [97].

Data analysis

Many single-cell methods involve advanced data analysis (see Box 3 and Figure 4 for a selection of computational methods). In scRNA-seq data, cell types can in principle be identified by clustering similar transcriptomes [99] and the underlying gene regulatory networks can be inferred [100–103]. However, both cell-type identification and network inference are improved by integrating multiple omics data sets [40,104]. Integration methods typically aim to extract variations common to all measured modalities [105–107]. That is even possible if molecular species are not measured simultaneously in the same cell, as shown, for example, for DNA methylation and transcriptome measurements [108].

Trajectory inference algorithms seek to reconstruct gene expression dynamics from scRNA-seq measurements of developing tissues. Many of these methods use the similarity of transcriptomes to estimate temporal proximity, which comes with many challenges and limitations [77,96,109,110]: for example, the starting point of a differentiation trajectory has to be provided by the user, because most methods cannot infer directionality. One exception is RNA velocity [111,112], which exploits RNA splicing dynamics to infer gene expression

Biochemical Society Transactions (2021) **49** 2509–2525
https://doi.org/10.1042/BST20210135

Box 3: A selection of computational methods for the analysis of single-cell omics data

Clustering methods have been used extensively for transcriptomics data (reviewed in [99]), where they partition cells based on the similarity of their transcriptomes. Clustering is now also applied to a combination of different omics data sets (reviewed in [158–160]). Clusters can be first obtained separately for each modality and then combined, or the different data sets are integrated prior to clustering. Popular examples of integration methods are WNN [40], totalVI [107], MOFA+ [105] and LIGER [108].

Inference of gene regulatory networks has been frequently performed using scRNA-seq data [102]. Examples of existing algorithms are GENIE3 [100], SCENIC [101], SINCERITIES [103], and Scribe [161]. A new method, CellOracle [104], allows the identification of gene regulatory networks from a combination of scRNA-seq and scATAC-seq data. Symphony [162] provides multi-omics clustering as well as gene regulatory network inference. Importantly, these methods often rely on the proper identification of TFs and CREs.

Inference of differentiation trajectories was first introduced for transcriptomics data. These methods make use of the asynchrony during differentiation and order cells by developmental progress (pseudotime). Examples of trajectory inference methods are PAGA [163], DPT [164], Monocle3 [165], FateID [166] and Palantir [167] (reviewed in [78]). A pseudotime method that makes use of spliced and unspliced RNA is RNA velocity [111,112], which has also been expanded to include protein dynamics [168]. To combine several transcriptomics data sets and recreate the differentiation trajectory, optimal transport theory has been applied [13,81]. A new, interesting method is MATCHER [169], which infers pseudotime based on multi-omics assays.

Reconstruction of lineage trees is the goal of dynamic lineage tracing techniques, where barcodes are introduced randomly during a short period of time. Classic reconstruction methods, like neighbor joining [170] are not robust enough for this purpose. Several studies therefore designed custom-made methods [157,171] and additionally, a new, more robust inference method has been proposed, Cassiopeia [113]. Building on the neighbor-joining algorithm, CLiNC [95] tries to discover inconsistencies within the phylogenic tree.

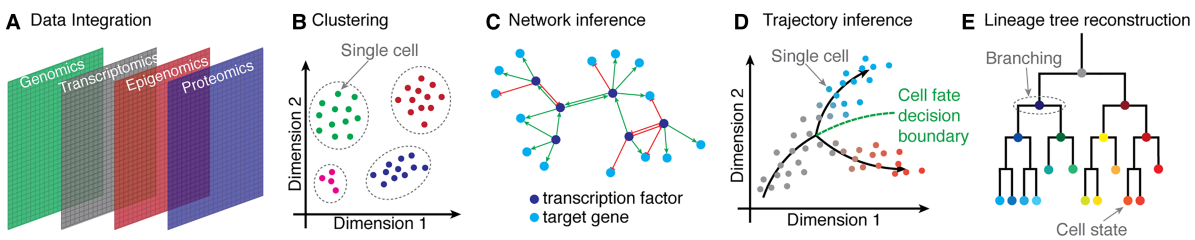

Figure 4. Common elements of single-cell omics data analysis.

(**a**) Each rectangle represents a data matrix from an omics technology with cells in columns and features in rows. Data Integration methods can be used to combine these data sets. (**b**) Each circle represents a cell projected into a two-dimensional space. Cells that are closer have more similar molecular profiles. Different colors indicate different clusters of cells, which have been determined through a clustering algorithm. Clusters are usually identified with cell types. (**c**) Inferred gene regulatory network. Circles show transcription factors and their target genes. Edges correspond to the interactions between them. (**d**) Each circle represents a cell in a two-dimensional space. Developmental dynamics, indicated by a black line, are revealed with trajectory inference methods. The cell fate decision boundary separates molecular profiles that develop into different cell types. (**e**) Each node represents a cell type. The leaves of the tree (lowest nodes) are the observed final cell types, whereas the other cell types are transient. Lines indicate the lineage relationships.

Biochemical Society Transactions (2021) **49** 2509–2525
https://doi.org/10.1042/BST20210135

dynamics and directionality. Time-resolved measurements can be analyzed with optimal transport theory to infer probabilities for the transitions between the observed cell types [13,81].

Prospective lineage tracing presents a completely different set of challenges for data analysis. The increase in data complexity caused by randomly inserted barcodes necessitates the development of novel algorithms to infer the underlying phylogenetic tree [96,113], which captures the hierarchy and relationship of cells during differentiation. As barcoding is often limited to a short period of time, it becomes difficult to infer the lineage tree beyond the point where barcoding has stopped. However, a new method [95] leverages covariances between barcodes to transcend this limitation. An interesting concept in this regard is phylodynamics [94], which studies how the cell-type distribution changes over time, given an observed lineage tree. For example, it has been shown that a model with a constant cell division rate can result in a skewed lineage tree that appears like earlier generations were dividing more rapidly [94].

The algorithms mentioned here are just a small selection of the many tools that have been developed specifically to deal with the challenges arising in single-cell methods. We refer the reader to [114–116] for a much more comprehensive overview.

Conceptual framework

Even with appropriate data analysis algorithms in place, we still need a conceptual framework for the quantitative understanding of cell types and their formation. The challenge is to reveal, how gene regulatory networks with certain topologies give rise to the observed cell types and molecular dynamics during differentiation. Dynamical systems theory has been used extensively to model gene regulatory networks quantitatively. In this framework, cell types can be understood as stable states in a system of coupled differential equations [117]. The number, position, and robustness of these stable states all depend on parameters that reflect the interactions between TFs and other members of the regulatory network. These parameters can be difficult to infer from experiments, except for (unrealistically) small networks. Nevertheless, dynamical systems describe key properties of the differentiation process. They explain how the interactions between several TFs jointly give rise to cell types that are robust up to a certain level of perturbation [118]. They also explain how a change in TF interactions causes cell types to destabilize [119]. Finally, unstable, intermediate cell states can be found, depending on the parameters of the system [120,121].

A dynamical systems model can be represented by a potential energy landscape, where a cell follows the path of steepest decent into locally stable states, that correspond to cell types [121–123], see Figure 1, middle right. This potential energy landscape is closely related to Waddington's epigenetic landscape [124], a pioneering metaphor that abstracted from molecular details to conceptualize embryonic development. Importantly, the shape of Waddington's landscape is constant in time and a location in the landscape corresponds to the complete molecular profile of a cell. In contrast, most dynamical system models identify the state of a cell by its transcriptome or even just the expression levels of the TFs in a gene regulatory network (see Figure 1, middle left). The shape of the potential landscape is then defined by the gene regulatory network, most importantly the interactions between TFs and their target genes [117]. Changes in the epigenetic state and other gene regulatory molecules can modulate the strength of those interactions (i.e. the parameters of the dynamic system) and thereby cause different stable and unstable states to appear or disappear [121,122]. Defining the gene expression profile as the state of the cell and modeling the epigenetic profile as parameters of the gene regulatory network has certain conceptual advantages. For example, at critical points, which have been studied extensively by catastrophe theory, small changes of the parameters can cause large changes in the stable states of a dynamical system [117,121]. Lineage decisions might thus be driven by dynamic epigenetic changes around critical points. Importantly, Waddington's landscape implies a strict hierarchy of differentiation, leading from multipotent to more and more specified, unipotent states (see Figure 1, right).

Despite its many advantages, the landscape model also has clear drawbacks, including its inability to describe periodic trajectories, for example, caused by the cell cycle [77,122]. Therefore, many other ways to conceptualize differentiation have been devised. For example, spin glass, a model that originated in physics, describes a system of interacting particles that can have stable low energy states corresponding to different cell types [117,125]. It accommodates different strengths of interactions between TFs, can describe symmetry breaking events and is scalable to larger numbers of TFs. However, it is often simplified by the usage of binary TF expression (on/off) and symmetric interactions for mathematical tractability.

The concepts discussed here are just a few examples of the many models that are currently being developed. More comprehensive overviews can be found in [117,126].

Perspectives

- Importance of the field: To discover the molecular underpinnings of cell types and their formation is of fundamental interest in developmental and stem cell biology. It is equally important for the understanding of diseases such as cancer, where cell types lose their stability and are transformed to malignant states.

- Summary of current thinking: New single-cell measurement techniques have given us unprecedented insights into the interactions and dynamics of the relevant molecular agents. In the current paradigm, transcription factors, regulatory DNA elements and other classes of molecules form a regulatory network from which cell types emerge.

- Comment on future directions: In the future, lineage tracing and other quantitative methods will be leveraged to reveal the complete lineage tree and infer a predictive mathematical model of the underlying gene regulatory network. Such a model would allow us to manipulate cell types at will, which has numerous medical applications.

Glossary

Biomolecular condensates: Droplets of a condensed liquid phase formed in cells by homotypic, multivalent interactions (i.e. interactions between identical molecules that involve multiple binding sites). One example is membrane-less organelles.
Bivalent domain: Chromatin domain that carries both activating and repressing histone marks.
Chromatin: The complex of nucleosomes, DNA, and other associated proteins.
Chromatin remodeler: Protein complexes that catalyze molecular changes of the chromosome, such as nucleosome removal.
Cis-regulatory elements: Sequences of non-coding DNA, which regulate the transcription of genes.
Coupled differential equations: Differential equations describe the temporal evolution of a system. They are coupled if variables appear in several equations. Such equations can have multiple stable solutions, which do not evolve in time unless perturbed.
Critical point: A point in parameter space where the number or stability of solutions to a dynamical system change abruptly.
Histone: Proteins that are crucial for the organization of DNA in the nucleus. DNA is tightly wound around nucleosome core particles which consist of eight histones.
Intrinsically disordered regions: Segments of a protein that do not form a stable three-dimensional structure.
Liquid–liquid phase separation: De-mixing of a homogeneous liquid into two distinct liquid phases.
Master transcription factor/master regulator: A transcription factor that affects the transcription of multiple downstream genes and is essential for cell-type specification.
Mediator: A multiprotein complex that coactivates transcription by interacting with TFs and RNA polymerase II.
Nucleosome: Smallest unit of DNA organization. Consists of a DNA wound around eight histones.
Paraspeckle: A biomolecular condensate that forms in the presence of the long non-coding RNA NEAT1 and several RNA binding proteins.
Pioneer factor: A transcription factor that can bind to nucleosome-bound DNA.
Pluripotency: Ability of a cell to give rise to multiple cell types.
Regulatory network: A system of interacting molecules that regulate each other's gene expression as well as a set of target genes.
RNA Polymerase II: A multiprotein complex that transcribes DNA into messenger RNA.
Single-cell omics technologies: Experimental methods to measure the entire genome, epigenome, transcriptome, proteome, etc. of a cell in high throughput.

Biochemical Society Transactions (2021) **49** 2509–2525
https://doi.org/10.1042/BST20210135

Stable states: A solution of a dynamical system that is a local minimum of the corresponding potential landscape.
Super-enhancer: A group of multiple enhancers in close proximity characterized by high levels of Mediator complex, which strongly drives gene expression of its target genes.
Transcription factor: A protein that binds to specific DNA sequences and regulates transcription.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

M.M. and S.S. were supported by the Netherlands Organization for Scientific Research (NWO/OCW, www.nwo.nl), as part of the Frontiers of Nanoscience (NanoFront) program. We acknowledge funding by an NWO/OCW Vidi grant (016.Vidi.189.007) for S.S.

Author Contributions

S.S. and M.M. wrote the manuscript. M.M. created the figures.

Abbreviations

CREs, *cis*-regulatory elements; ESCs, embryonic stem cells; IDRs, intrinsically disordered regions; RBPs, RNA-binding proteins; TFs, transcription factors.

References

1 Morris, S.A. (2019) The evolving concept of cell identity in the single cell era. *Development* **146**, e169748 https://doi.org/10.1242/dev.169748
2 Arendt, D., Musser, J.M., Baker, C.V.H., Bergman, A., Cepko, C., Erwin, D.H. et al. (2016) The origin and evolution of cell types. *Nat. Rev. Genet.* **17**, 744–757 https://doi.org/10.1038/nrg.2016.127
3 Savulescu, A.F., Jacobs, C., Negishi, Y., Davignon, L. and Mhlanga, M.M. (2020) Pinpointing cell identity in time and space. *Front. Mol. Biosci.* **7**, 209 https://doi.org/10.3389/fmolb.2020.00209
4 Xia, B. and Yanai, I. (2019) A periodic table of cell types. *Development* **146**, e169854 https://doi.org/10.1242/dev.169854
5 Almeida, N., Chung, M.W.H., Drudi, E.M., Engquist, E.N., Hamrud, E., Isaacson, A. et al. (2021) Employing core regulatory circuits to define cell identity. *EMBO J.* **40**, e106785 https://doi.org/10.15252/embj.2020106785
6 Dupont, C., Armant, D.R. and Brenner, C.A. (2009) Epigenetics: definition, mechanisms and clinical perspective. *Semin. Reprod. Med.* **27**, 351 https://doi.org/10.1055/S-0029-1237423
7 Allis, C.D. and Jenuwein, T. (2016) The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* **17**, 487–500 https://doi.org/10.1038/nrg.2016.59
8 Paro, R., Grossniklaus, U., Santoro, R. and Wutz, A. (2021) *Introduction to Epigenetics*, Springer International Publishing, Cham
9 Armingol, E., Officer, A., Harismendy, O. and Lewis, N.E. (2021) Deciphering cell–cell interactions and communication from gene expression. *Nat. Rev. Genet.* **22**, 71–88 https://doi.org/10.1038/s41576-020-00292-x
10 Almet, A.A., Cang, Z., Jin, S. and Nie, Q. (2021) The landscape of cell–cell communication through single-cell transcriptomics. *Curr. Opin. Syst. Biol.* **26**, 12–23 https://doi.org/10.1016/J.COISB.2021.03.007
11 Eling, N., Morgan, M.D. and Marioni, J.C. (2019) Challenges in measuring and understanding biological noise. *Nat. Rev. Genet.* **20**, 536–548 https://doi.org/10.1038/s41576-019-0130-6
12 Guillemin, A. and Stumpf, M.P.H. (2021) Noise and the molecular processes underlying cell fate decision-making. *Phys. Biol.* **18**, 11002 https://doi.org/10.1088/1478-3975/abc9d1
13 Mittnenzweig, M., Mayshar, Y., Cheng, S., Ben-Yair, R., Hadas, R., Rais, Y. et al. (2021) A single-embryo, single-cell time-resolved model for mouse gastrulation. *Cell* **184**, 1–18 https://doi.org/10.1016/j.cell.2021.04.004
14 Stadhouders, R., Filion, G.J. and Graf, T. (2019) Transcription factors and 3D genome conformation in cell-fate decisions. *Nature* **569**, 345–354 https://doi.org/10.1038/s41586-019-1182-7
15 Zaret, K.S. (2020) Pioneer transcription factors initiating gene network changes. *Annu. Rev. Genet.* **54**, 367–385 https://doi.org/10.1146/annurev-genet-030220-015007
16 Klemm, S.L., Shipony, Z. and Greenleaf, W.J. (2019) Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* **20**, 207–220 https://doi.org/10.1038/s41576-018-0089-8
17 Stelzer, Y., Shivalila, C.S., Soldner, F., Markoulaki, S. and Jaenisch, R. (2015) Tracing dynamic changes of DNA methylation at single-cell resolution. *Cell* **163**, 218–229 https://doi.org/10.1016/J.CELL.2015.08.046
18 Song, Y., van den Berg, P.R., Markoulaki, S., Soldner, F., Dall'Agnese, A., Henninger, J.E. et al. (2019) Dynamic enhancer DNA methylation as basis for transcriptional and cellular heterogeneity of ESCs. *Mol. Cell* **75**, 905–920.e6 https://doi.org/10.1016/J.MOLCEL.2019.06.045
19 Riggs, A.D., Martienssen, R.A. and Russo, V.E. (1996) *Epigenetic Mechanisms of Gene Regulation*, Cold Spring Harbor Laboratory Press
20 Ludwig, C.H. and Bintu, L. (2019) Mapping chromatin modifications at the single cell level. *Development* **146**, e170217 https://doi.org/10.1242/dev.170217

21 Cang, Z., Wang, Y., Wang, Q., Cho, K.W.Y., Holmes, W. and Nie, Q. (2021) A multiscale model via single-cell transcriptomics reveals robust patterning mechanisms during early mammalian embryo development. *PLoS Comput Biol* **17**, e1008571 https://doi.org/10.1371/journal.pcbi.1008571
22 Bessonnard, S., De, M.L., Gonze, D., Barriol, M., Dennis, C., Goldbeter, A. et al. (2014) Gata6, nanog and Erk signaling control cell fate in the inner cell mass through a tristable regulatory network. *Development* **141**, 3637–3648 https://doi.org/10.1242/dev.109678
23 Schrode, N., Saiz, N., Di, T.S. and Hadjantonakis, A.-K. (2014) GATA6 levels modulate primitive endoderm cell fate choice and timing in the mouse blastocyst. *Dev. Cell* **29**, 454–467 https://doi.org/10.1016/J.DEVCEL.2014.04.011
24 Graf, T. and Enver, T. (2009) Forcing cells to change lineages. *Nature* **462**, 587–594 https://doi.org/10.1038/nature08533
25 Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A. et al. (2013) Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934 https://doi.org/10.1016/j.cell.2013.09.053
26 Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H. et al. (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 https://doi.org/10.1016/j.cell.2013.03.035
27 Razin S, V. and Gavrilov, A.A. (2020) The role of liquid–liquid phase separation in the compartmentalization of cell nucleus and spatial genome organization. *Biochemistry* **85**, 643–650 https://doi.org/10.1134/S0006297920060012
28 Lyon, A.S., Peeples, W.B. and Rosen, M.K. (2021) A framework for understanding the functions of biomolecular condensates across scales. *Nat. Rev. Mol. Cell Biol.* **22**, 215–235 https://doi.org/10.1038/s41580-020-00303-z
29 Banani, S.F., Lee, H.O., Hyman, A.A. and Rosen, M.K. (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 https://doi.org/10.1038/nrm.2017.7
30 Gomes, E. and Shorter, J. (2019) The molecular language of membraneless organelles. *J. Biol. Chem.* **294**, 7115–7127 https://doi.org/10.1074/jbc.TM118.001192
31 Grosch, M., Ittermann, S., Shaposhnikov, D. and Drukker, M. (2020) Chromatin-associated membraneless organelles in regulation of cellular differentiation. *Stem Cell Rep* **15**, 1220–1232 https://doi.org/10.1016/j.stemcr.2020.10.011
32 Sabari, B.R., Dall'Agnese, A., Boija, A., Klein, I.A., Coffey, E.L., Shrinivas, K. et al. (2018) Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 https://doi.org/10.1126/science.aar3958
33 Esposito, M., Fang, C., Cook, K.C., Park, N., Wei, Y., Spadazzi, C. et al. (2021) TGF-β-induced DACT1 biomolecular condensates repress Wnt signalling to promote bone metastasis. *Nat. Cell Biol.* **23**, 257–267 https://doi.org/10.1038/s41556-021-00641-w
34 Guo, Y.E., Manteiga, J.C., Henninger, J.E., Sabari, B.R., Dall'Agnese, A., Hannett, N.M. et al. (2019) Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* **572**, 543–548 https://doi.org/10.1038/s41586-019-1464-0
35 Cho, W.K., Spille, J.H., Hecht, M., Lee, C., Li, C., Grube, V. et al. (2018) Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412–415 https://doi.org/10.1126/science.aar4199
36 Hochane, M., van den Berg, P.R., Fan, X., Bérenger-Currias, N., Adegeest, E., Bialecka, M. et al. (2019) Single-cell transcriptomics reveals gene expression dynamics of human fetal kidney development. *PLoS Biol.* **17**, e3000152 https://doi.org/10.1371/journal.pbio.3000152
37 Ando, Y., Kwon, A.T.J. and Shin, J.W. (2020) An era of single-cell genomics consortia. *Exp. Mol. Med.* **52**, 1409–1418 https://doi.org/10.1038/s12276-020-0409-x
38 Regev, A., Teichmann, S.A., Lander, E.S., Amit, I., Benoist, C., Birney, E. et al. (2017) The human cell atlas. *eLife* **6**, e27041 https://doi.org/10.7554/eLife.27041
39 Schaum, N., Karkanias, J., Neff, N.F., May, A.P., Quake, S.R., Wyss-Coray, T. et al. (2018) Single-cell transcriptomics of 20 mouse organs creates a tabula muris. *Nature* **562**, 367–372 https://doi.org/10.1038/s41586-018-0590-4
40 Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A. et al. (2021) Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 https://doi.org/10.1016/j.cell.2021.04.048
41 Bendall, S.C., Simonds, E.F., Qiu, P., Amir, E.D., Krutzik, P.O., Finck, R. et al. (2011) Single-Cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332**, 687–696 https://doi.org/10.1126/SCIENCE.1198704
42 Tanay, A. and Regev, A. (2017) Scaling single-cell genomics from phenomenology to mechanism. *Nature* **541**, 331–338 https://doi.org/10.1038/nature21350
43 Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P. et al. (2015) Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* **523**, 486–490 https://doi.org/10.1038/nature14590
44 Cusanovich, D.A., Daza, R., Adey, A., Pliner, H.A., Christiansen, L., Gunderson, K.L. et al. (2015) Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* **348**, 910–914 https://doi.org/10.1126/science.aab1601
45 Cao, J., Cusanovich, D.A., Ramani, V., Aghamirzaie, D., Pliner, H.A., Hill, A.J. et al. (2018) Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science* **361**, 1380–1385 https://doi.org/10.1126/science.aau0730
46 Zhu, C., Zhang, Y., Li, Y.E., Lucero, J., Behrens, M.M. and Ren, B. (2021) Joint profiling of histone modifications and transcriptome in single cells from mouse brain. *Nat. Methods* **18**, 283–292 https://doi.org/10.1038/s41592-021-01060-3
47 Angermueller, C., Clark, S.J., Lee, H.J., Macaulay, I.C., Teng, M.J., Hu, T.X. et al. (2016) Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat. Methods* **13**, 229–232 https://doi.org/10.1038/nmeth.3728
48 Ramani, V., Deng, X., Qiu, R., Gunderson, K.L., Steemers, F.J., Disteche, C.M. et al. (2017) Massively multiplex single-cell Hi-C. *Nat. Methods* **14**, 263–266 https://doi.org/10.1038/nmeth.4155
49 Nagano, T., Lubling, Y., Stevens, T.J., Schoenfelder, S., Yaffe, E., Dean, W. et al. (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59–64 https://doi.org/10.1038/nature12593
50 Lee, D.-S., Luo, C., Zhou, J., Chandran, S., Rivkin, A., Bartlett, A. et al. (2019) Simultaneous profiling of 3D genome structure and DNA methylation in single human cells. *Nat. Methods* **16**, 999–1006 https://doi.org/10.1038/s41592-019-0547-z
51 Ulianov S, V., Velichko, A.K., Magnitov, M.D., Luzhin A, V., Golov, A.K., Ovsyannikova, N. et al. (2021) Suppression of liquid–liquid phase separation by 1,6-hexanediol partially compromises the 3D genome organization in living cells. *Nucleic Acids Res.* **49**, 10524–10541 https://doi.org/10.1093/NAR/GKAB249
52 Belaghzal, H., Borrman, T., Stephens, A.D., Lafontaine, D.L., Venev S, V., Weng, Z. et al. (2021) Liquid chromatin Hi-C characterizes compartment-dependent chromatin interaction dynamics. *Nat. Genet.* **53**, 367–378 https://doi.org/10.1038/s41588-021-00784-4

53 Holler, K., Neuschulz, A., Drewe-Boß, P., Mintcheva, J., Spanjaard, B., Arsiè, R. et al. (2021) Spatio-temporal mRNA tracking in the early zebrafish embryo. *Nat. Commun.* **12**, 3358 https://doi.org/10.1038/s41467-021-23834-1
54 Eng, C.-H.L., Lawson, M., Zhu, Q., Dries, R., Koulena, N., Takei, Y. et al. (2019) Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. *Nature* **568**, 235–239 https://doi.org/10.1038/s41586-019-1049-y
55 Lundberg, E. and Borner, G.H.H. (2019) Spatial proteomics: a powerful discovery tool for cell biology. *Nat. Rev. Mol. Cell Biol.* **20**, 285–302 https://doi.org/10.1038/s41580-018-0094-y
56 Shema, E., Bernstein, B.E. and Buenrostro, J.D. (2019) Single-cell and single-molecule epigenomics to uncover genome regulation at unprecedented resolution. *Nat. Genet.* **51**, 19–25 https://doi.org/10.1038/s41588-018-0290-x
57 Lee, J., Hyeon, D.Y. and Hwang, D. (2020) Single-cell multiomics: technologies and data analysis methods. *Exp. Mol. Med.* **52**, 1428–1442 https://doi.org/10.1038/s12276-020-0420-2
58 Mayran, A., Khetchoumian, K., Hariri, F., Pastinen, T., Gauthier, Y., Balsalobre, A. et al. (2018) Pioneer factor Pax7 deploys a stable enhancer repertoire for specification of cell fate. *Nat. Genet.* **50**, 259–269 https://doi.org/10.1038/s41588-017-0035-2
59 Joseph, S.R., Pálfy, M., Hilbert, L., Kumar, M., Karschau, J., Zaburdaev, V. et al. (2017) Competition between histone and transcription factor binding regulates the onset of transcription in zebrafish embryos. *eLife* **6**, e23326 https://doi.org/10.7554/ELIFE.23326
60 Clapier, C.R. and Cairns, B.R. (2009) The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* **78**, 273–304 https://doi.org/10.1146/ANNUREV.BIOCHEM.77.062706.153223
61 Tyagi, M., Imam, N., Verma, K. and Patel, A.K. (2016) Chromatin remodelers: we are the drivers!!. *Nucleus* **7**, 388 https://doi.org/10.1080/19491034.2016.1211217
62 Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jørgensen, H.F., John, R.M. et al. (2006) Chromatin signatures of pluripotent cell lines. *Nat. Cell Biol.* **8**, 532–538 https://doi.org/10.1038/ncb1403
63 Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J. et al. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 https://doi.org/10.1016/J.CELL.2006.02.041
64 Blanco, E., González-Ramírez, M., Alcaine-Colet, A., Aranda, S. and Di, C.L. (2020) The bivalent genome: characterization, structure, and regulation. *Trends Genet.* **36**, 118–131 https://doi.org/10.1016/J.TIG.2019.11.004
65 Margueron, R., Justin, N., Ohno, K., Sharpe, M.L., Son, J., Drury, III, W.J. et al. (2009) Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* **461**, 762–767 https://doi.org/10.1038/nature08398
66 Di Croce, L. and Helin, K. (2013) Transcriptional regulation by polycomb group proteins. *Nat. Struct. Mol. Biol.* **20**, 1147–1155 https://doi.org/10.1038/nsmb.2669
67 Kundu, S., Ji, F., Sunwoo, H., Jain, G., Lee, J.T., Sadreyev, R.I. et al. (2017) Polycomb repressive complex 1 generates discrete compacted domains that change during differentiation. *Mol. Cell* **65**, 432–446.e5 https://doi.org/10.1016/J.MOLCEL.2017.01.009
68 Cruz-Molina, S., Respuela, P., Tebartz, C., Kolovos, P., Nikolic, M., Fueyo, R. et al. (2017) PRC2 facilitates the regulatory topology required for poised enhancer function during pluripotent stem cell differentiation. *Cell Stem Cell* **20**, 689–705.e9 https://doi.org/10.1016/J.STEM.2017.02.004
69 Atlasi, Y. and Stunnenberg, H.G. (2017) The interplay of epigenetic marks during stem cell differentiation and development. *Nat. Rev. Genet.* **18**, 643–658 https://doi.org/10.1038/nrg.2017.57
70 Daneshvar, K., Ardehali, M.B., Klein, I.A., Hsieh, F.K., Kratkiewicz, A.J., Mahpour, A. et al. (2020) lncRNA DIGIT and BRD3 protein form phase-separated condensates to regulate endoderm differentiation. *Nat. Cell Biol.* **22**, 1211–1222 https://doi.org/10.1038/s41556-020-0572-2
71 Modic, M., Grosch, M., Rot, G., Schirge, S., Lepko, T., Yamazaki, T. et al. (2019) Cross-Regulation between TDP-43 and paraspeckles promotes pluripotency-differentiation transition. *Mol. Cell* **74**, 951–965.e13 https://doi.org/10.1016/j.molcel.2019.03.041
72 Grosch, M., Ittermann, S., Rusha, E., Greisle, T., Ori, C., Truong, D.J.J. et al. (2020) Nucleus size and DNA accessibility are linked to the regulation of paraspeckle formation in cellular differentiation. *BMC Biol.* **18**, 1–19 https://doi.org/10.1186/s12915-020-00770-y
73 Knott, G.J., Bond, C.S. and Fox, A.H. (2016) The DBHS proteins SFPQ, NONO and PSPC1: a multipurpose molecular scaffold. *Nucleic Acids Res.* **44**, 3989–4004 https://doi.org/10.1093/NAR/GKW271
74 Wang, S., Zuo, H., Jin, J., Lv, W., Xu, Z., Fan, Y. et al. (2019) Long noncoding RNA Neat1 modulates myogenesis by recruiting Ezh2. *Cell Death Dis.* **10**, 1–15 https://doi.org/10.1038/s41419-019-1742-7
75 Sulston, J.E. and Horvitz, H.R. (1977) Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev Biol.* **56**, 110–156 https://doi.org/10.1016/0012-1606(77)90158-0
76 Sagar, G.D. (2020) Deciphering cell fate decision by integrated single-cell sequencing analysis. *Annu. Rev. Biomed. Data Sci.* **3**, 1–22 https://doi.org/10.1146/annurev-biodatasci-111419-091750
77 Weinreb, C., Wolock, S., Tusi, B.K., Socolovsky, M. and Klein, A.M. (2018) Fundamental limits on dynamic inference from single-cell snapshots. *Proc. Natl Acad. Sci. U.S.A.* **115**, E2467–E2476 https://doi.org/10.1073/pnas.1714723115
78 Saelens, W., Cannoodt, R., Todorov, H. and Saeys, Y. (2019) A comparison of single-cell trajectory inference methods. *Nat. Biotechnol.* **37**, 547–554 https://doi.org/10.1038/s41587-019-0071-9
79 Pijuan-Sala, B., Griffiths, J.A., Guibentif, C., Hiscock, T.W., Jawaid, W., Calero-Nieto, F.J. et al. (2019) A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature* **566**, 490–495 https://doi.org/10.1038/s41586-019-0933-9
80 Delile, J., Rayon, T., Melchionda, M., Edwards, A., Briscoe, J. and Sagner, A. (2019) Single cell transcriptomics reveals spatial and temporal dynamics of gene expression in the developing mouse spinal cord. *Development* **146**, dev173807 https://doi.org/10.1242/dev.173807
81 Schiebinger, G., Shu, J., Tabaka, M., Cleary, B., Subramanian, V., Solomon, A. et al. (2019) Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. *Cell* **176**, 928–943.e22 https://doi.org/10.1016/j.cell.2019.01.006
82 Nowotschin, S., Setty, M., Kuo, Y.-Y., Liu, V., Garg, V., Sharma, R. et al. (2019) The emergent landscape of the mouse gut endoderm at single-cell resolution. *Nature* **569**, 361–367 https://doi.org/10.1038/s41586-019-1127-1
83 Palii, C.G., Cheng, Q., Gillespie, M.A., Shannon, P., Mazurczyk, M., Napolitani, G. et al. (2019) Single-Cell proteomics reveal that quantitative changes in co-expressed lineage-specific transcription factors determine cell fate. *Cell Stem Cell* **24**, 812–820.e5 https://doi.org/10.1016/j.stem.2019.02.006

84 Carter, B. and Zhao, K. (2020) The epigenetic basis of cellular heterogeneity. *Nat. Rev. Genet.* **22**, 235–250 https://doi.org/10.1038/s41576-020-00300-0
85 Preissl, S., Fang, R., Huang, H., Zhao, Y., Raviram, R., Gorkin, D.U. et al. (2018) Single-nucleus analysis of accessible chromatin in developing mouse forebrain reveals cell-type-specific transcriptional regulation. *Nat. Neurosci.* **21**, 432–439 https://doi.org/10.1038/s41593-018-0079-3
86 Jia, G., Preussner, J., Chen, X., Guenther, S., Yuan, X., Yekelchyk, M. et al. (2018) Single cell RNA-seq and ATAC-seq analysis of cardiac progenitor cell transition states and lineage settlement. *Nat. Commun* **9**, 1–17 https://doi.org/10.1038/s41467-018-07307-6
87 Trevino, A.E., Müller, F., Andersen, J., Sundaram, L., Kathiria, A., Shcherbina, A. et al. (2020) Chromatin and gene-regulatory dynamics of the developing human cerebral cortex at single-cell resolution. *Cell* **184**, 5053–5069.e23 https://doi.org/10.1016/J.CELL.2021.07.039
88 Zhu, C., Yu, M., Huang, H., Juric, I., Abnousi, A., Hu, R. et al. (2019) An ultra high-throughput method for single-cell joint analysis of open chromatin and transcriptome. *Nat. Struct. Mol. Biol.* **26**, 1063–1070 https://doi.org/10.1038/s41594-019-0323-x
89 Argelaguet, R., Clark, S.J., Mohammed, H., Stapel, L.C., Krueger, C., Kapourani, C.A. et al. (2019) Multi-omics profiling of mouse gastrulation at single-cell resolution. *Nature* **576**, 487–491 https://doi.org/10.1038/s41586-019-1825-8
90 Xiong, H., Luo, Y., Wang, Q., Yu, X. and He, A. (2021) Single-cell joint detection of chromatin occupancy and transcriptome enables higher-dimensional epigenomic reconstructions. *Nat. Methods*, **18**, 1–9 https://doi.org/10.1038/s41592-021-01129-z
91 Wagner, D.E. and Klein, A.M. (2020) Lineage tracing meets single-cell omics: opportunities and challenges. *Nat. Rev. Genet.* **21**, 410–427 https://doi.org/10.1038/s41576-020-0223-2
92 Baron, C.S. and van Oudenaarden, A. (2019) Unravelling cellular relationships during development and regeneration using genetic lineage tracing. *Nat. Rev. Mol. Cell Biol.* **20**, 753–765 https://doi.org/10.1038/s41580-019-0186-3
93 McKenna, A. and Gagnon, J.A. (2019) Recording development with single cell dynamic lineage tracing. *Development* **146**, e169730 https://doi.org/10.1242/dev.169730
94 Stadler, T., Pybus, O.G. and Stumpf, M.P.H. (2021) Phylodynamics for cell biologists. *Science* **371**, eaah6266 https://doi.org/10.1126/science.aah6266
95 Weinreb, C. and Klein, A.M. (2020) Lineage reconstruction from clonal correlations. *Proc. Natl Acad. Sci. U.S.A.* **117**, 17041–17048 https://doi.org/10.1073/pnas.2000238117
96 Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F.D. and Klein, A.M. (2020) Lineage tracing on transcriptional landscapes links state to fate during differentiation. *Science* **367**, eaaw3381 https://doi.org/10.1126/science.aaw3381
97 Wagner, D.E., Weinreb, C., Collins, Z.M., Briggs, J.A., Megason, S.G. and Klein, A.M. (2018) Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **360**, 981–987 https://doi.org/10.1126/science.aar4362
98 Rodriguez-Fraticelli, A.E., Wolock, S.L., Weinreb, C.S., Panero, R., Patel, S.H., Jankovic, M. et al. (2018) Clonal analysis of lineage fate in native haematopoiesis. *Nature* **553**, 212–216 https://doi.org/10.1038/nature25168
99 Krzak, M., Raykov, Y., Boukouvalas, A., Cutillo, L. and Angelini, C. (2019) Benchmark and parameter sensitivity analysis of single-cell RNA sequencing clustering methods. *Front. Genet.* **10**, 1253 https://doi.org/10.3389/fgene.2019.01253
100 Huynh-Thu, V.A., Irrthum, A., Wehenkel, L. and Geurts, P. (2010) Inferring regulatory networks from expression data using tree-based methods. *PLoS ONE* **5**, 12776 https://doi.org/10.1371/journal.pone.0012776
101 Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G. et al. (2017) SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 https://doi.org/10.1038/nmeth.4463
102 Pratapa, A., Jalihal, A.P., Law, J.N., Bharadwaj, A. and Murali, T.M. (2020) Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. *Nat. Methods* **17**, 147–154 https://doi.org/10.1038/s41592-019-0690-6
103 Papili Gao, N., Ud-Dean, S.M.M., Gandrillon, O. and Gunawan, R. (2018) Sincerities: inferring gene regulatory networks from time-stamped single cell transcriptional expression profiles. *Bioinformatics* **34**, 258–266 https://doi.org/10.1093/bioinformatics/btx575
104 Kamimoto, K., Hoffmann, C.M. and Morris, S.A. (2020) Celloracle: dissecting cell identity via network inference and in silico gene perturbation. *bioRxiv* https://doi.org/10.1101/2020.02.17.947416
105 Argelaguet, R., Arnol, D., Bredikhin, D., Deloro, Y., Velten, B., Marioni, J.C. et al. (2020) MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol.* **21**, 1–17 https://doi.org/10.1186/S13059-020-02015-1
106 Butler, A., Hoffman, P., Smibert, P., Papalexi, E. and Satija, R. (2018) Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 https://doi.org/10.1038/nbt.4096
107 Gayoso, A., Steier, Z., Lopez, R., Regier, J., Nazor, K.L., Streets, A. et al. (2021) Joint probabilistic modeling of single-cell multi-omic data with totalVI. *Nat. Methods* **18**, 272–282 https://doi.org/10.1038/s41592-020-01050-x
108 Welch, J.D., Kozareva, V., Ferreira, A., Vanderburg, C., Martin, C. and Macosko, E.Z. (2019) Single-cell multi-omic integration compares and contrasts features of brain cell identity. *Cell* **177**, 1873–1887.e17 https://doi.org/10.1016/J.CELL.2019.05.006
109 Tritschler, S., Büttner, M., Fischer, D.S., Lange, M., Bergen, V., Lickert, H. et al. (2019) Concepts and limitations for learning developmental trajectories from single cell genomics. *Development* **146**, dev170506 https://doi.org/10.1242/dev.170506
110 Weinreb, C., Wolock, S. and Klein, A.M. (2018) SPRING: a kinetic interface for visualizing high dimensional single-cell expression data. *Bioinformatics* **34**, 1246–1248 https://doi.org/10.1093/bioinformatics/btx792
111 Bergen, V., Lange, M., Peidli, S., Wolf, F.A. and Theis, F.J. (2020) Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 https://doi.org/10.1038/s41587-020-0591-3
112 La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V. et al. (2018) RNA velocity of single cells. *Nature* **560**, 494–498 https://doi.org/10.1038/s41586-018-0414-6
113 Jones, M.G., Khodaverdian, A., Quinn, J.J., Chan, M.M., Hussmann, J.A., Wang, R. et al. (2020) Inference of single-cell phylogenies from lineage tracing data using cassiopeia. *Genome Biol.* **21**, 1–27 https://doi.org/10.1186/S13059-020-02000-8
114 Nguyen, H., Tran, D., Tran, B., Pehlivan, B. and Nguyen, T. (2021) A comprehensive survey of regulatory network inference methods using single cell RNA sequencing data. *Brief. Bioinform* **22**, 1–15 https://doi.org/10.1093/BIB/BBAA190
115 Andrews, T.S., Kiselev, V.Y., McCarthy, D. and Hemberg, M. (2020) Tutorial: guidelines for the computational analysis of single-cell RNA sequencing data. *Nat. Protoc.* **16**, 1–9 https://doi.org/10.1038/s41596-020-00409-w
116 Stuart, T. and Satija, R. (2019) Integrative single-cell analysis. *Nat. Rev. Genet.* **20**, 257–272 https://doi.org/10.1038/s41576-019-0093-7

117 Teschendorff, A.E. and Feinberg, A.P. (2021) Statistical mechanics meets single-cell biology. *Nat. Rev. Genet.* **22**, 1–18 https://doi.org/10.1038/s41576-021-00341-z
118 Furusawa, C. and Kaneko, K. (2012) A dynamical-systems view of stem cell biology. *Science* **338**, 215–217 https://doi.org/10.1126/SCIENCE.1224311
119 Tyson, J.J. and Novak, B. (2020) A dynamical paradigm for molecular cell biology. *Trends Cell Biol* **30**, 504–515 https://doi.org/10.1016/J.TCB.2020.04.002
120 MacLean, A.L., Hong, T. and Nie, Q. (2018) Exploring intermediate cell states through the lens of single cells. *Curr. Opin. Syst. Biol.* **9**, 32–41 https://doi.org/10.1016/j.coisb.2018.02.009
121 Guillemin, A. and Stumpf, M.P.H. (2020) Non-equilibrium statistical physics, transitory epigenetic landscapes, and cell fate decision dynamics. *Math Biosci. Eng.* **17**, 7916–7930 https://doi.org/10.3934/mbe.2020402
122 Greulich, P., Smith, R. and MacArthur, B.D. (2020) The physics of cell fate. *Phenotypic Switch*, 189–206 https://doi.org/10.1016/b978-0-12-817996-3.00003-7
123 Xu, L. and Wang, J. (2020) Quantifying waddington landscapes, paths, and kinetics of cell fate decision making of differentiation/development. *Phenotypic Switch*, 157–187 https://doi.org/10.1016/b978-0-12-817996-3.00002-5
124 Waddington, C.H. (1975) *The Strategy of the Genes : A Discussion of Some Aspects of Theoretical Biology*, Allen and Unwin, London
125 Lang, A.H., Li, H., Collins, J.J. and Mehta, P. (2014) Epigenetic landscapes explain partially reprogrammed cells and identify key reprogramming genes. *PLoS Comput. Biol.* **10**, e1003734 https://doi.org/10.1371/JOURNAL.PCBI.1003734
126 Cahan, P., Cacchiarelli, D., Dunn, S.-J., Hemberg, M., de S Lopes, S.M.C., Morris, S.A. et al. (2021) Computational stem cell biology: open questions and guiding principles. *Cell Stem Cell* **28**, 20–32 https://doi.org/10.1016/J.STEM.2020.12.012
127 Picelli, S., Björklund, Å.K, Faridani, O.R., Sagasser, S., Winberg, G. and Sandberg, R. (2013) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1098 https://doi.org/10.1038/nmeth.2639
128 Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M. et al. (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 https://doi.org/10.1016/J.CELL.2015.05.002
129 Hashimshony, T., Senderovich, N., Avital, G., Klochendler, A., de Leeuw, Y., Anavy, L. et al. (2016) CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol.* **17**, 1–7 https://doi.org/10.1186/S13059-016-0938-8
130 Cao, J., Packer, J.S., Ramani, V., Cusanovich, D.A., Huynh, C., Daza, R. et al. (2017) Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**, 661–667 https://doi.org/10.1126/SCIENCE.AAM8940
131 Budnik, B., Levy, E., Harmange, G. and Slavov, N. (2018) SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol.* **19**, 1–12 https://doi.org/10.1186/S13059-018-1547-5
132 Specht, H., Emmott, E., Petelski, A.A., Huffman, R.G., Perlman, D.H., Serra, M. et al. (2021) Single-cell proteomic and transcriptomic analysis of macrophage heterogeneity using SCoPE2. *Genome Biol.* **22**, 1–27 https://doi.org/10.1186/S13059-021-02267-5
133 Schoof, E.M., Furtwängler, B., Üresin, N., Rapin, N., Savickas, S., Gentil, C. et al. (2021) Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. *Nat. Commun.* **12**, 1–15 https://doi.org/10.1038/s41467-021-23667-y
134 Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P.K., Swerdlow, H. et al. (2017) Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 https://doi.org/10.1038/nmeth.4380
135 Peterson, V.M., Zhang, K.X., Kumar, N., Wong, J., Li, L., Wilson, D.C. et al. (2017) Multiplexed quantification of proteins and transcripts in single cells. *Nat. Biotechnol.* **35**, 936–939 https://doi.org/10.1038/nbt.3973
136 Gerlach, J.P., van Buggenum, J.A.G., Tanis, S.E.J., Hogeweg, M., Heuts, B.M.H., Muraro, M.J. et al. (2019) Combined quantification of intracellular (phospho-)proteins and transcriptomics from fixed single cells. *Sci. Rep.* **9**, 1–10 https://doi.org/10.1038/s41598-018-37977-7
137 Smallwood, S.A., Lee, H.J., Angermueller, C., Krueger, F., Saadeh, H., Peat, J. et al. (2014) Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat. Methods* **11**, 817–820 https://doi.org/10.1038/nmeth.3035
138 Farlik, M., Sheffield, N.C., Klughammer, J. and Bock Correspondence, C. (2015) Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics. *Cell Rep.* **10**, 1386–1397 https://doi.org/10.1016/j.celrep.2015.02.001
139 Guo, H., Zhu, P., Wu, X., Li, X., Wen, L. and Tang, F. (2013) Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res.* **23**, 2126–2135 https://doi.org/10.1101/GR.161679.113
140 Mooijman, D., Dey, S.S., Boisset, J.C., Crosetto, N. and Van Oudenaarden, A. (2016) Single-cell 5hmC sequencing reveals chromosome-wide cell-to-cell variability and enables lineage reconstruction. *Nat. Biotechnol.* **34**, 852–856 https://doi.org/10.1038/nbt.3598
141 Hu, Y., Huang, K., An, Q., Du, G., Hu, G., Xue, J. et al. (2016) Simultaneous profiling of transcriptome and DNA methylome from a single cell. *Genome Biol.* **17**, 1–11 https://doi.org/10.1186/S13059-016-0950-Z
142 Clark, S.J., Argelaguet, R., Kapourani, C.-A., Stubbs, T.M., Lee, H.J., Alda-Catalinas, C. et al. (2018) scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat. Commun.* **9**, 1–9 https://doi.org/10.1038/s41467-018-03149-4
143 Rotem, A., Ram, O., Shoresh, N., Sperling, R.A., Goren, A., Weitz, D.A. et al. (2015) Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat. Biotechnol.* **33**, 1165–1172 https://doi.org/10.1038/nbt.3383
144 Kaya-Okur, H.S., Wu, S.J., Codomo, C.A., Pledger, E.S., Bryson, T.D., Henikoff, J.G. et al. (2019) CUT&tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* **10**, 1–10 https://doi.org/10.1038/s41467-019-09982-5
145 Li, G., Liu, Y., Zhang, Y., Kubo, N., Yu, M., Fang, R. et al. (2019) Joint profiling of DNA methylation and chromatin architecture in single cells. *Nat. Methods* **16**, 991–993 https://doi.org/10.1038/s41592-019-0502-z
146 Zong, C., Lu, S., Chapman, A.R. and Xie, X.S. (2012) Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* **338**, 1622–1626 https://doi.org/10.1126/SCIENCE.1229164
147 Abascal, F., Harvey, L.M.R., Mitchell, E., Lawson, A.R.J., Lensing S, V., Ellis, P. et al. (2021) Somatic mutation landscapes at single-molecule resolution. *Nature* **593**, 405–410 https://doi.org/10.1038/s41586-021-03477-4
148 Macaulay, I.C., Haerty, W., Kumar, P., Li, Y.I., Hu, T.X., Teng, M.J. et al. (2015) G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat. Methods* **12**, 519–522 https://doi.org/10.1038/nmeth.3370
149 Rodriguez-Meira, A., Buck, G., Clark, S.A., Povinelli, B.J., Alcolea, V., Louka, E. et al. (2019) Unravelling intratumoral heterogeneity through high-sensitivity single-cell mutational analysis and parallel RNA sequencing. *Mol. Cell* **73**, 1292–1305.e8 https://doi.org/10.1016/J.MOLCEL.2019.01.009

150 Nam, A.S., Kim, K.T., Chaligne, R., Izzo, F., Ang, C., Taylor, J. et al. (2019) Somatic mutations and cell identity linked by genotyping of transcriptomes. *Nature* **571**, 355–360 https://doi.org/10.1038/s41586-019-1367-0
151 Abyzov, A. and Vaccarino, F.M. (2020) Cell lineage tracing and cellular diversity in humans. *Annu. Rev. Genomics Hum. Genet.* **21**, 101–116 https://doi.org/10.1146/ANNUREV-GENOM-083118-015241
152 Xu, J., Nuno, K., Litzenburger, U.M., Qi, Y., Corces, M.R., Majeti, R. et al. (2019) Single-cell lineage tracing by endogenous mutations enriched in transposase accessible mitochondrial DNA. *eLife* **8**, e45105 https://doi.org/10.7554/ELIFE.45105
153 Gaiti, F., Chaligne, R., Gu, H., Brand, R.M., Kothen-Hill, S., Schulman, R.C. et al. (2019) Epigenetic evolution and lineage histories of chronic lymphocytic leukaemia. *Nature* **569**, 576–580 https://doi.org/10.1038/s41586-019-1198-z
154 Wei, C.J.-Y. and Zhang, K. (2020) RETrace: simultaneous retrospective lineage tracing and methylation profiling of single cells. *Genome Res.* **30**, gr.255851.119 https://doi.org/10.1101/GR.255851.119
155 Alemany, A., Florescu, M., Baron, C.S., Peterson-Maduro, J. and Van Oudenaarden, A. (2018) Whole-organism clone tracing using single-cell sequencing. *Nature* **556**, 108–112 https://doi.org/10.1038/nature25969
156 Raj, B., Gagnon, J.A. and Schier, A.F. (2018) Large-scale reconstruction of cell lineages using single-cell readout of transcriptomes and CRISPR–Cas9 barcodes by scGESTALT. *Nat. Protoc.* **13**, 2685–2713 https://doi.org/10.1038/s41596-018-0058-x
157 Spanjaard, B., Hu, B., Mitic, N., Olivares-Chauvet, P., Janjuha, S., Ninov, N. et al. (2018) Simultaneous lineage tracing and cell-type identification using crIsPr-Cas9-induced genetic scars. *Nat. Biotechnol.* **36**, 469–473 https://doi.org/10.1038/nbt.4124
158 Rappoport, N. and Shamir, R. (2018) Multi-omic and multi-view clustering algorithms: review and cancer benchmark. *Nucleic Acids Res.* **46**, 10546 https://doi.org/10.1093/NAR/GKY889
159 Colomé-Tatché, M. and Theis, F.J. (2018) Statistical single cell multi-omics integration. *Curr. Opin. Syst. Biol.* **7**, 54–59 https://doi.org/10.1016/J.COISB.2018.01.003
160 Efremova, M. and Teichmann, S.A. (2020) Computational methods for single-cell omics across modalities. *Nat. Methods* **17**, 14–17 https://doi.org/10.1038/s41592-019-0692-4
161 Qiu, X., Rahimzamani, A., Wang, L., Ren, B., Mao, Q., Durham, T. et al. (2020) Inferring causal gene regulatory networks from coupled single-cell expression dynamics using scribe. *Cell Syst* **10**, 265–274.e11 https://doi.org/10.1016/j.cels.2020.02.003
162 Burdziak, C., Azizi, E., Prabhakaran, S. and Pe'er, D. (2019) A nonparametric multi-view model for estimating cell type-specific gene regulatory networks. ArXiv
163 Wolf, F.A., Hamey, F.K., Plass, M., Solana, J., Dahlin, J.S., Göttgens, B. et al. (2019) PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* **20**, 1–9 https://doi.org/10.1186/S13059-019-1663-X
164 Haghverdi, L., Büttner, M., Wolf, F.A., Buettner, F. and Theis, F.J. (2016) Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods* **13**, 845–848 https://doi.org/10.1038/nmeth.3971
165 Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J. et al. (2019) The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 https://doi.org/10.1038/s41586-019-0969-x
166 Herman, J.S., Sagar Grün, D. (2018) FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat. Methods* **15**, 379–386 https://doi.org/10.1038/nmeth.4662
167 Setty, M., Kiseliovas, V., Levine, J., Gayoso, A., Mazutis, L. and Pe'er, D. (2019) Characterization of cell fate probabilities in single-cell data with Palantir. *Nat. Biotechnol.* **37**, 451–460 https://doi.org/10.1038/s41587-019-0068-4
168 Gorin, G., Svensson, V. and Pachter, L. (2020) Protein velocity and acceleration from single-cell multiomics experiments. *Genome Biol.* **21**, 39 https://doi.org/10.1186/s13059-020-1945-3
169 Welch, J.D., Hartemink, A.J. and Prins, J.F. (2017) MATCHER: manifold alignment reveals correspondence between single cell transcriptome and epigenome dynamics. *Genome Biol.* **18**, 1–19 https://doi.org/10.1186/S13059-017-1269-0
170 Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 https://doi.org/10.1093/OXFORDJOURNALS.MOLBEV.A040454
171 Chan, M.M., Smith, Z.D., Grosswendt, S., Kretzmer, H., Norman, T.M., Adamson, B. et al. (2019) Molecular recording of mammalian embryogenesis. *Nature* **570**, 77–82 https://doi.org/10.1038/s41586-019-1184-5