Please cite this article in press as: Kester and van Oudenaarden, Single-Cell Transcriptomics Meets Lineage Tracing, Cell Stem Cell (2018), https://doi.org/10.1016/j.stem.2018.04.014

Cell Stem Cell
Review

Single-Cell Transcriptomics Meets Lineage Tracing

Lennart Kester[1] and Alexander van Oudenaarden[1,*]
[1]Oncode Institute, Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences (KNAW) and University Medical Center Utrecht, 3584 CT Utrecht, the Netherlands
*Correspondence: a.vanoudenaarden@hubrecht.eu
https://doi.org/10.1016/j.stem.2018.04.014

Reconstructing lineage relationships between cells within a tissue or organism is a long-standing aim in biology. Traditionally, lineage tracing has been achieved through the (genetic) labeling of a cell followed by the tracking of its offspring. Currently, lineage trajectories can also be predicted using single-cell transcriptomics. Although single-cell transcriptomics provides detailed phenotypic information, the predicted lineage trajectories do not necessarily reflect genetic relationships. Recently, techniques have been developed that unite these strategies. In this Review, we discuss transcriptome-based lineage trajectory prediction algorithms, single-cell genetic lineage tracing, and the promising combination of these techniques for stem cell and cancer research.

Understanding the lineage through which tissues and organisms are formed is one of the fundamental questions in biology. Identifying these relationships will provide invaluable information not only on normal tissue development and homeostasis, but also on developmental disorders and pathologies such as cancer. Historically, lineage tracing is accomplished through the introduction of a heritable mark in a cell, followed by the tracking of its progeny. The different cell types that comprise the progeny are developmentally related since all of these marked cells originate from the same founder cell. Furthermore, the variety of cell types found in the progeny represents the potential of the founder cell. In order to accurately predict the potential of the founder cell, lineage tracing requires accurate cell-type identification. Ideally, one would use as many markers as possible to achieve accurate and precise cell-type classification. However, cell-type identification is usually based on a limited number of markers, thereby potentially masking variability within a subpopulation of cells that express the selected marker genes. This approach might therefore give a biased view on organ complexity.

Recent advances in single-cell transcriptomics technologies now allow transcriptome profiling of thousands of single cells, giving unprecedented resolution in cell-type identification and deepening our understanding of tissue complexity (Grün et al., 2015; Haber et al., 2017; Jaitin et al., 2014; Klein et al., 2015; Macosko et al., 2015; Zeisel et al., 2015). The power of single-cell transcriptomics has led researchers to start large-scale sequencing projects such as "The Human Cell Atlas" and "Fly Cell Atlas," endeavors aimed toward sequencing all cell types present in the human body (Rozenblatt-Rosen et al., 2017) and fruit fly, respectively. In addition, the NIH Brain Initiative (Insel et al., 2013) is funding projects aimed at sequencing all cell types present in the human and rodent brain. In parallel, there have been considerable advances in computational methods aimed at performing lineage trajectory reconstruction based on single-cell transcriptomics data (Cannoodt et al., 2016), allowing researchers to sort the transcriptomes of single cells according to their differentiation status. Since lineage reconstruction based on single-cell transcriptomics is independent from the true genetic relationship between cells, we reserve the term "lineage tracing" for genetic lineage tracing and use the term "differentiation trajectories" for transcriptome-derived lineage predictions. However, new experimental techniques that combine single-cell transcriptome sequencing with genetic lineage labels provide information on the relationships between cells for lineage reconstruction along with detailed phenotypic information (Alemany et al., 2018; Frieda et al., 2017; Raj et al., 2018; Spanjaard et al., 2018; Yao et al., 2017). This integration of single-cell lineage tracing and transcriptomics will be incredibly powerful, as it allows coarse lineage reconstruction based on genetically heritable marks, followed by refinement based on the transcriptome-derived differentiation trajectories and the assessment of gene-expression changes as a function of the developmental cell state (Figure 1).

In this review, we discuss some of the numerous strategies used to predict differentiation trajectories based on single-cell transcriptomics and highlight their applications in various biological systems. Next, we discuss recent developments in prospective lineage tracing that entails the introduction of a heritable mark and retrospective lineage tracing, which exploits naturally occurring elements in the genome (Woodworth et al., 2017). Finally, we discuss some recent studies that successfully combine genetic lineage tracing with single-cell transcriptomics, highlighting the power of integrating these two techniques.

Advances in Single-Cell mRNA Sequencing Technologies

In the short time span since the first single-cell transcriptome sequencing technique was published in 2009 (Tang et al., 2009), an impressive amount of new techniques have become available (Hashimshony et al., 2012, 2016; Islam et al., 2014; Klein et al., 2015; Macosko et al., 2015; Muraro et al., 2016; Picelli et al., 2013). While manual picking and processing of each individual cell was initially required, now thousands to tens of thousands of cells are processed in parallel using microfluidic or robot-based approaches (Klein et al., 2015; Macosko et al., 2015; Muraro et al., 2016). Two recently published techniques completely eliminated the need for single-cell isolation by using a pool and split strategy in which cells are first labeled in multiple groups, followed by an iteration of pooling, splitting, and labeling to ensure each single-cell obtains a unique set of

Please cite this article in press as: Kester and van Oudenaarden, Single-Cell Transcriptomics Meets Lineage Tracing, Cell Stem Cell (2018), https://doi.org/10.1016/j.stem.2018.04.014

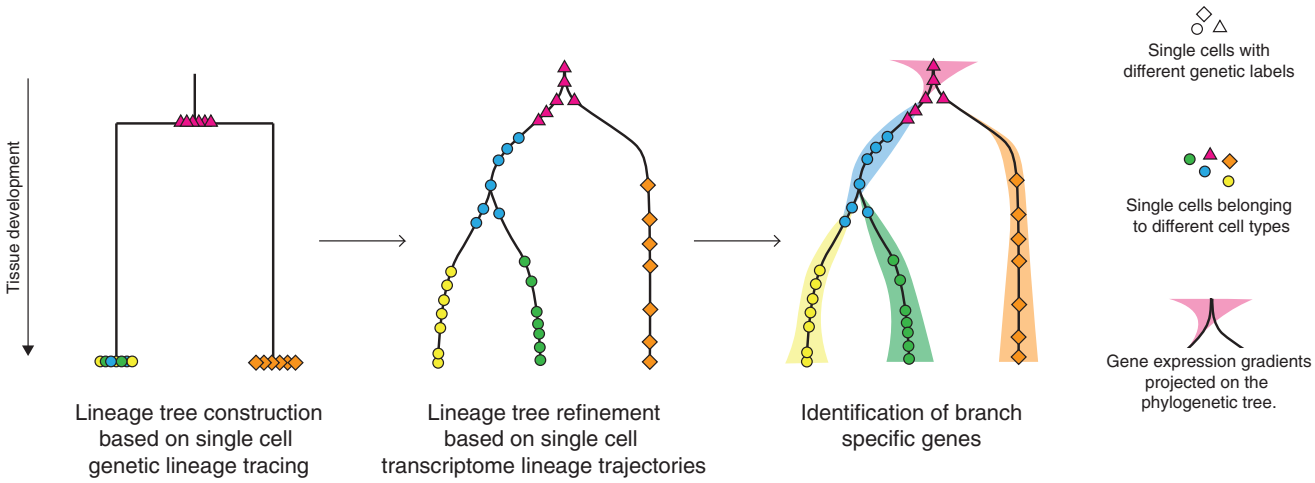


Figure 1. Combination of Single-Cell Genetic Lineage Tracing and Single-Cell Transcriptomics
First a phylogenetic tree is constructed based on genetic labels identified in single cells. This tree can then be refined using transcriptomics-based lineage reconstruction algorithms. Finally, gene-expression gradients can be projected onto the phylogenetic tree to identify gene-expression dynamics throughout the system.

labels (Cao et al., 2017; Rosenberg et al., 2018). A more detailed overview of the available single-cell mRNA sequencing techniques and their properties can be found in Haque et al. (2017) and Papalexi and Satija (2018).

With the massive increase in throughput of single-cell mRNA sequencing techniques, multidimensional data analysis becomes increasingly important. Numerous algorithms have been developed to cluster cells, extract significantly differentially expressed genes between clusters of cells, identify outlier cells, and allow visual representation of datasets in two dimensions. These algorithms along with other important aspects of the design and analysis of single-cell mRNA sequencing experiments will not be discussed in depth but are reviewed in Grün and van Oudenaarden (2015) and Yuan et al. (2017).

Differentiation Trajectory Reconstruction Algorithms

Single-cell transcriptomics allows one to investigate the transcriptional state of thousands of individual single cells thereby reliably capturing cell-type diversity in heterogeneous samples. When applied to a developing or differentiating biological system, many cells transition between different states. If sufficient amounts of cells in these transition states are captured, differentiation trajectories through which tissues are built or maintained can be accurately predicted. These differentiation trajectories can then be exploited to probe kinship among different cell types and to identify genes essential for transitions along these trajectories. However, the difficulty with single-cell transcriptomics data lies in its inherent noisiness and dropout effects (lowly expressed genes are difficult to detect due to technical limitations). Over the last few years, an impressive amount of computational methods have been developed to place cells on differentiation trajectories. Most of these methods rely on the assumption that cells with similar expression profiles arise from the same lineage and that cells with more similarity between their expression profiles are closely related. Here, we will discuss some of those algorithms and the biological systems they have been applied to.

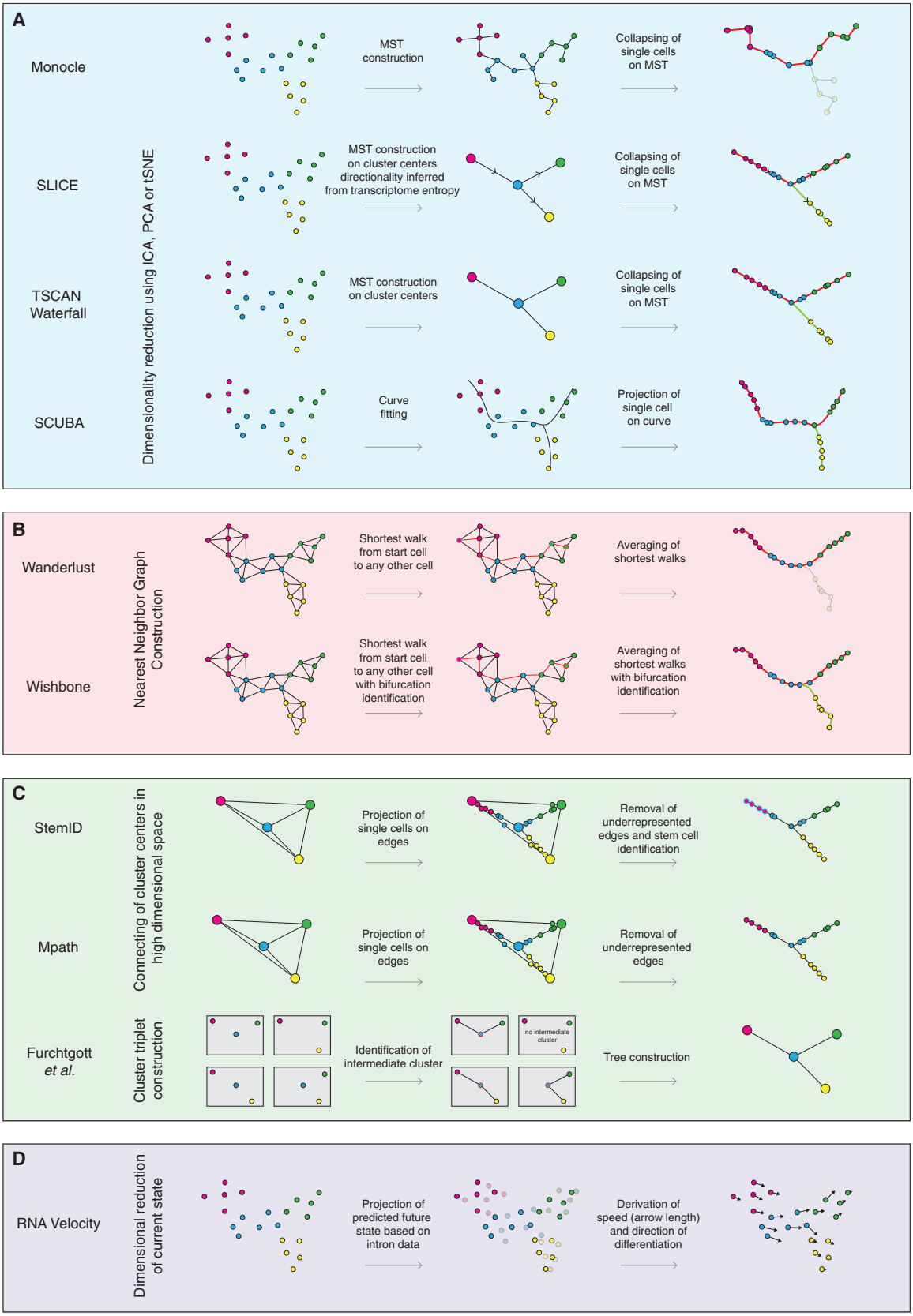
Dimensionality Reduction-Based Algorithms

The majority of the differentiation trajectory reconstruction methods rely on some form of dimensionality reduction. One of the first and commonly applied algorithms, Monocle (Trapnell et al., 2014), uses independent-component analysis (ICA) to project all cells in a low dimensional space (usually 2 dimensions) (Figure 2A). This is followed by Minimum Spanning Tree (MST) construction and the definition of a backbone connecting the most and least differentiated cells. All remaining cells are then projected on the backbone, resulting in a 1-dimensional ordering of all cells. This dimension is termed pseudotime, which represents the predicted lineage trajectory of the studied sample. In Monocle, the pseudotime ordering of the cells is not directional, meaning that high pseudotime could either mean most or least differentiated cells; however, users can define a root cell giving the MST a starting point. Monocle was originally used to reconstruct the differentiation trajectory of developing human skeletal muscle myoblasts. One limitation of Monocle is its inability to allow for bifurcations in the lineage prediction, so only linear differentiation systems can be analyzed. This issue has been resolved in the second edition, Monocle2 (Qiu et al., 2017). Monocle2 builds the lineage tree in a higher dimensional space, retaining more data for highly intricate differentiation trajectories. Since their publication, Monocle and Monocle2 have been used in numerous studies to predict differentiation trajectories of many developing or differentiating systems. Among these are studies that unravel the lineage trajectories of several types of neurons (Bardy et al., 2016; Camp et al., 2015; Dulken et al., 2017; Hanchate et al., 2015; Lacar et al., 2016; Lisi et al., 2017; Marques et al., 2016; Treutlein et al., 2016), progenitors and stem cells of the hematopoietic system (Mass et al., 2016; Qiu et al., 2017; Stubbington et al., 2016; Tang et al., 2017), the placenta (Tsang et al., 2017), and hair follicles (Joost et al., 2016).

Several other algorithms, including SLICE (Guo et al., 2017), TSCAN (Ji and Ji, 2016), Waterfall (Shin et al., 2015), SCUBA (Marco et al., 2014), and Slingshot (Street et al., 2017), employ a similar strategy as Monocle by first reducing dimensionality of

Please cite this article in press as: Kester and van Oudenaarden, Single-Cell Transcriptomics Meets Lineage Tracing, Cell Stem Cell (2018), https://doi.org/10.1016/j.stem.2018.04.014

CellPress



(legend on next page)

Please cite this article in press as: Kester and van Oudenaarden, Single-Cell Transcriptomics Meets Lineage Tracing, Cell Stem Cell (2018), https://doi.org/10.1016/j.stem.2018.04.014

the data through principal-component analysis (PCA), ICA, or t-stochastic neighbor embedding (tSNE) and then constructing a MST or fitting a smooth curve and finally projecting single cells on the pseudotime axis (Figure 2A). MST construction in SLICE, TSCAN, and Waterfall is less sensitive to outliers since these algorithms use predefined clusters of cells rather than single cells. SLICE uses transcriptome entropy as a measure for differentiation in order to identify the least differentiated cells in the population and thereby creates a pseudotime starting point. The application of TSCAN to a single-cell dataset containing primitive hematopoiesis cells revealed the importance of the HOPX transcription factor for blood formation (Palpant et al., 2017). SCUBA reduces dimensionality of the data using tSNE followed by the fitting of a smooth curve. SCUBA has been used to identify a population of cells in transition between $Lgr5^+$ stem cells and more mature cells in the mouse small intestine. On top of that, SCUBA identified a set of key genes that change expression during this transition (Kim et al., 2016). In contrast to SLICE, SCUBA, and TSCAN, Slingshot takes any form of dimensional reduction, constructs a MST, and then further refines this tree by fitting smooth curves through all of the major MST branches. Single cells are then projected onto their closest curve resulting in ordered lineage trajectories with bifurcations. Slingshot has recently been used to predict the cell-fate potentials and branch points in the lineage trajectories of olfactory stem cells (Fletcher et al., 2017).

Nearest Neighbor Graph-Based Algorithms

Another class of differentiation trajectory reconstruction algorithms is based on k-nearest neighbor graphs (k-NNGs). In k-NNGs, each cell is connected to its k nearest neighbors, thereby linking similar cells to each other. The first algorithm using k-NNGs was Wanderlust (Bendall et al., 2014), which represents cells as nodes in a collection of k-NNGs, each comprising a subset of the total population of cells (Figure 2B). A user-defined root cell is used to generate a collection of shortest walks from the root cell to all the other cells in each of the graphs. This process results in numerous possible differentiation trajectories, which are then averaged to select the most probable one. Similar to Monocle, Wanderlust only enables the study of linear trajectories, while its successor, Wishbone (Setty et al., 2016) allows bifurcations, expanding the repertoire of complex differentiation trajectories that can be studied (Figure 2B). Both Wanderlust and Wishbone were originally designed for Cytometry by Time Of Flight (CyTOF) data, but Wishbone has been adapted so it can be used with single-cell transcriptomics data. Wanderlust and Wishbone were developed to order cells along a developmental axis; however, there are other uses for k-NNGs in single-cell transcriptomics. Markov Affinity-based Graph Imputation of Cells (MAGIC), for instance, locally diffuses gene-expression values in the NNG, thereby smoothening gene expression across highly similar cells (van Dijk et al., 2017). This reduces the dropout effects often observed for lowly expressed genes.

Other graph-based algorithms include approximate graph abstraction (AGA) (Wolf et al., 2017) and population balance analysis (PBA) (Weinreb et al., 2018). AGA averages single cells on the NNG into clusters before constructing the differentiation tree and should therefore be less sensitive to outliers. PBA approximates the velocity of cell differentiation based on the local density of cells on the NNG. PBA also provides fate probabilities, predicting the direction of differentiation for the less differentiated cells.

Other Lineage Reconstruction Algorithms

StemID (Grün et al., 2016) and Mpath (Chen et al., 2016) belong to a third category of differentiation trajectory reconstruction algorithms and start by creating clusters of cells using k-medoid or hierarchical clustering, respectively (Figure 2C). All cluster centers are connected in a high dimensional space, creating a transition network. Single cells are then projected on the edges between the cluster centers, thereby populating the edges. Underpopulated edges are removed, resulting in a lineage trajectory network. StemID then identifies a stem cell population from the transcriptome entropy of each cell cluster, thereby providing a root for the network. StemID has been used to identify a potential stem cell population in the human pancreas (Grün et al., 2016).

Another approach, first developed for bulk mRNA sequencing by Heinäniemi et al. (2013), but later implemented for single-cell mRNA sequencing by Furchtgott et al. (2017), starts with the identification of cell clusters. Each triplet of clusters is then screened for differentially expressed transcription factors, which are subsequently used to infer whether the cell clusters in the triplet are related, and if so, to identify the intermediate cluster. This is done iteratively for all triplets of clusters followed by the construction of a lineage tree based on the relations between all triplets. This approach is useful for datasets with highly diverse systems containing many cell types, although in the current implementation, single cells are not projected back on the lineage tree. This approach has been used to infer the differentiation trees of hematopoietic cells (Furchtgott et al., 2017) and of an *in vitro* neural differentiation system (Yao et al., 2017).

Figure 2. Overview of Lineage Reconstruction Algorithms

(A) Lineage reconstruction algorithms based on dimensional reduction. Monocle uses independent-component analysis, followed by the construction of a minimum spanning tree (MST) connecting all cells. Connecting the two cells furthest away from each other identifies a backbone. Directionality can be provided by the user through the identification of a root cell. Large side branches are excluded, and remaining cells are projected onto the pseudotime backbone. SLICE constructs a MST of cluster centers, and directionality is inferred from transcriptome entropy. Single cells are projected on the edges connecting the cluster centers. TSCAN and Waterfall also constructs a MST based on the cluster centers, followed by projection of the single cells onto the edges to align cells in pseudotime. SCUBA uses tSNE for dimensionality reduction followed by the fitting of a smooth curve. Single cells are projected on the smooth curve to order them in pseudotime. Monocle, TSCAN, Waterfall, and SCUBA all require user input to infer directionality.

(B) Lineage reconstruction algorithms based on NNGs. Both Wanderlust and Wishbone start with the construction of a NNG. A collection of shortest walks, from a user-defined root cell to all other cells in the graph, is then used to construct the lineage trajectory. Wishbone has the added benefit that it can identify bifurcations in the lineage trajectory.

(C) Lineage reconstruction algorithms based on cluster networks. Both StemID and Mpath start by connecting all cluster centers in a high dimensional space. Single cells are then projected on the edges between the clusters, and underrepresented edges are removed from the graph. StemID identifies a potential stem cell population based on transcriptome entropy. The Furchgott method infers the intermediate cluster (if possible) from each triplet of clusters in the data, followed by tree construction based on the triplet relations.

(D) Graphical representation of RNA velocity. Current and future state of the cell are computed based on exon and intron data, respectively. Difference between current and future state determines the direction and speed of differentiation. No user input is required to infer these parameters.

Please cite this article in press as: Kester and van Oudenaarden, Single-Cell Transcriptomics Meets Lineage Tracing, Cell Stem Cell (2018), https://doi.org/10.1016/j.stem.2018.04.014

A recent and fundamentally different view on differentiation trajectory reconstruction was presented in La Manno et al. (2017). In this algorithm, named RNA velocity, transcripts only containing exons are selected to represent the current state of the cell. The fraction of unspliced transcripts across all genes defines an RNA velocity vector, pointing toward the future state of the cell. Both the current state and the future state can then be projected into a 2-dimensional space, showing the direction and the speed of differentiation (Figure 2D). The authors use this approach to show differentiation trajectories of astrocytes, oligodendrocyte precursors, dentate gyrus granule neurons, and pyramidal neurons in the developing mouse hippocampus. Furthermore, directionality from the RNA velocity vector has allowed for the identification of glia cells as the root of the differentiation trajectory, which confirms the current view on hippocampus development (Gebara et al., 2016).

Besides these algorithms, many publications present their own differentiation trajectory reconstruction algorithms due to unsatisfactory performance of previous strategies. This raises the question about whether the robustness of each differentiation trajectory reconstruction algorithm depends on the biological system of interest or technical differences, such as the experimental protocols.

Reconstructing differentiation trajectories from single-cell transcriptomics data heavily relies on sufficient sampling of cells that transition between the different states in the lineage trajectory. In the scenario of little or no cell differentiation, predictions on lineage relations become incredibly challenging. This will be particularly problematic in pathologies, such as cancer, where tumors often have multiple transcriptionally distinct populations of cells (Brady et al., 2017; Chung et al., 2017; Patel et al., 2014; Tirosh et al., 2016a, 2016b), and transitioning cells connecting these populations are frequently absent. Finally, and most importantly, differentiation trajectories based on single-cell transcriptomics are purely phenotypic: they do not necessarily reflect the genetic relations between the true lineages of the cells.

Single-Cell Genetic Lineage Tracing

Reconstructing lineage relationships between cells in developing tissues or organisms, or between stem cells and their mature progeny, has been a long-standing interest in biology. Lineage tracing was initially performed through the visual tracking of cells over time (Deppe et al., 1978), utilizing several strategies: the creation of chimeric embryos (Mintz, 1967), the engraftment of cells from one species to another (Le Douarin and Teillet, 1973), and the injection of dyes into a single founder cell (Lawson et al., 1986; Pedersen et al., 1986). The development of fluorescent-activated cell sorting (FACS) allowed isolation of single cells and the transplantation of these single cells to assess their potential. This has been at the basis of studies on hematopoietic stem cells (Osawa et al., 1996) and cancer initiation (Quintana et al., 2008). In recent years, lineage tracing is mainly performed through tracking of genetic features, either introduced experimentally or intrinsic to the system, termed prospective and retrospective lineage tracing, respectively. Prospective lineage tracing relies on the introduction of reporter transgenes, such as β-galactosidase or GFP (Frank and Sanes, 1991; Turner and Cepko, 1987). Transgenes integrate into the genome through viruses, with the goal of labeling single founder cells. All descendants of the founder cell inherit the transgenes. Since the transgenes encode a protein easily visualized using microscopy, the clone and the cell-fate choices made by the descendants of the founder cell become traceable. Another frequently used strategy is MARCM in *Drosophila* (Lee and Luo, 1999) or MADM in mice (Zong et al., 2005), where a fluorescent protein is expressed in a few founder cells upon induction of a recombinase that either removes an inhibiting factor or repairs a deprecated fluorescent protein. All descendants of the founder cells are labeled, allowing lineage tracing and fate determination of their progenies. However, this strategy only detects clones sparsely distributed across a sample, otherwise distinguishing clones from one another become impossible.

Viral Barcoding-Based Lineage Tracing

To overcome the limitations associated with imaging-based lineage tracing, viral strategies have been extended to include cell-specific DNA barcodes. The DNA barcodes usually consist of a stretch of random nucleotides, providing almost limitless complexity (Lu et al., 2011). Each founder cell contains a unique DNA barcode, meaning that cells with different barcodes originated from different founder cells even if they occur in close proximity. The DNA barcodes can be identified through (next-generation) sequencing, allowing the identification of the clones and the assessment of the clonal complexity of the studied system (Figure 3A).

The viral barcoding approach has been used extensively over the last few years to perform lineage tracing in a wide variety of tissues. In the hematopoietic system, viral barcoding has delineated how HSPCs differentiate into the wide variety of mature cells that comprise the blood (Gerlach et al., 2013; Gerrits et al., 2010; Lu et al., 2011; Naik et al., 2013; Rodriguez-Fraticelli et al., 2018; Schepers et al., 2008; van Heijst et al., 2009; Verovskaya et al., 2013). Viral barcoding has also successfully elucidated aspects of cancer biology. The landmark paper Eirew et al. (2015) demonstrates with viral barcoding that clonal selection upon xenograft transplantation is highly variable from tumor to tumor. Tumors that experience moderate initial selection upon xenotransplantation have long-lasting clonal dynamics during serial xenotransplantations. Furthermore, they show that clonal dynamics patterns are reproducible when performing parallel xenotransplantations from the same tumor. A second publication had established that the number of cells with tumor initiating capacity varies from 1 in 10 to 1 in 10,000 cells upon xenotransplantation (Nguyen et al., 2014). Viral barcoding strategies have also been employed to study drug resistance in *in vitro* models of cancer (Bhang et al., 2015).

Cre-Lox-Based Lineage Tracing

The use of fluorescent proteins for lineage tracing has immensely expanded with the introduction of color switching techniques such as Brainbow and its successors (Weissman and Pan, 2015). These techniques rely on a Cre-Lox system, which, upon activation, marks a founder cell with a single or an assortment of fluorescent proteins. The advantage of this approach over viral labeling is that cell labeling can be induced at any point in time and in a wide variety of tissues. These techniques are frequently implemented in model organisms to study the dynamics of tissue development and tissue maintenance and

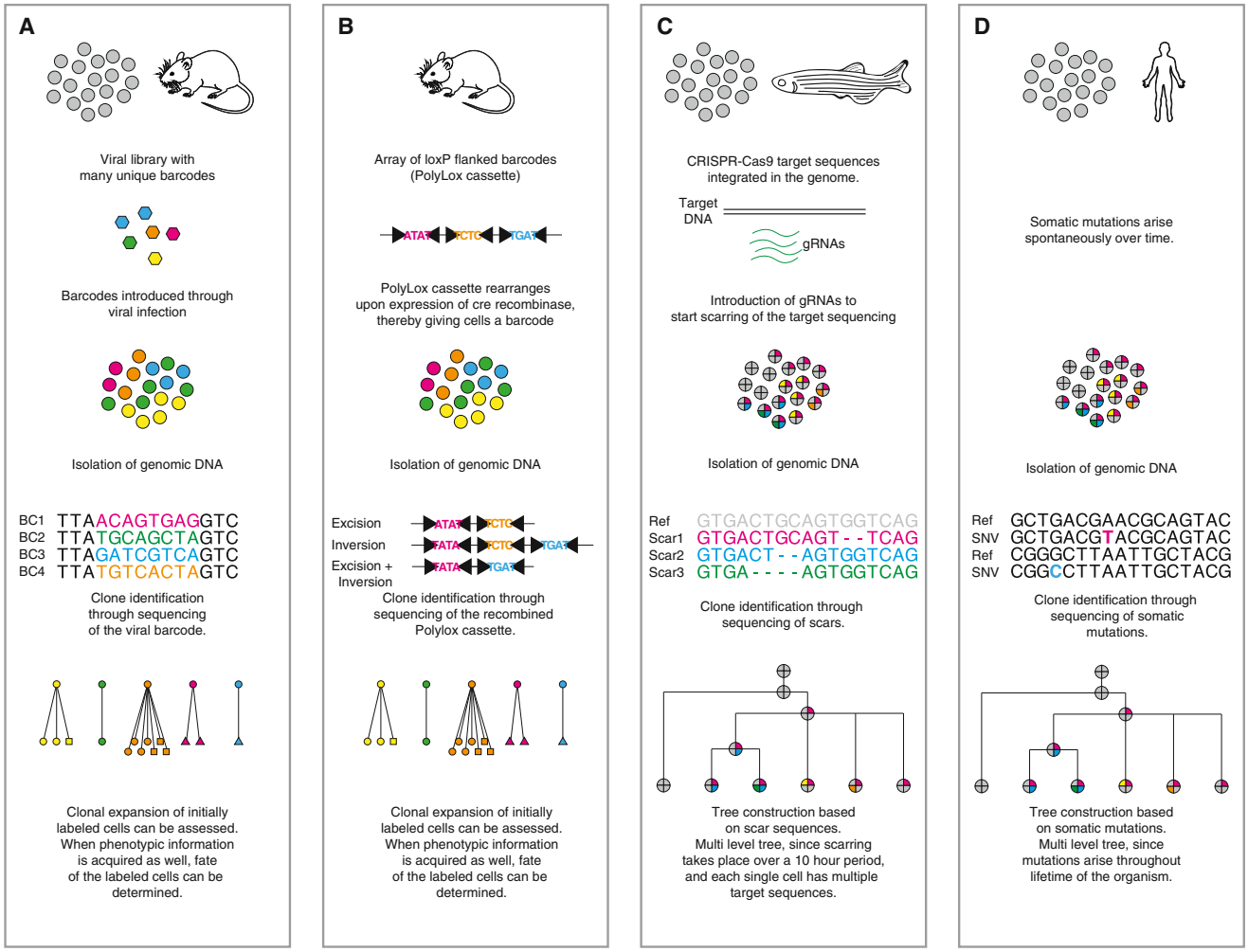


Figure 3. Overview of Genetic Lineage Tracing Strategies

(A) Lineage tracing through viral barcoding can be performed in a wide variety of model systems, among which are cultured cells and mice. Cells are infected with a virus library containing many different barcodes. After a period of time, cells are harvested, genomic DNA is isolated, and barcodes are sequenced. Clones can be identified through the barcode sequence. Clonal expansion of initially labeled cells can be assessed. When phenotypic information is acquired (shape), it is possible to determine the fate of the initially labeled cells.

(B) The Polylox mouse model was created by the introduction of a set of barcodes interspersed with loxP sites. Upon activation of the Cre recombinase, the Polylox cassette recombines, thereby producing unique cellular barcodes via a combination of losses and inversions of single barcodes. After a period of time, cells are harvested, genomic DNA is isolated, and barcodes are sequenced. Clone identification is done through the assessment of the combination of losses and inversions of barcodes. Clonal expansion of initially labeled cells can be assessed. When phenotypic information is acquired (shape), it is possible to determine the fate of the initially labeled cells.

(C) Lineage tracing using CRISPR-Cas9 can be done in cultured cells and zebrafish. Introduction of CRISPR-Cas9 and gRNA into the cells results in the scarring of the target sequence during a given time window. After a period of time cells are harvested, genomic DNA is isolated, and scars are sequenced. The combination of scars in each cell produces a unique barcode, and the construction of multi-level phylogenetic trees is possible since scarring takes place over a long period of time and a portion of the scars will be shared between clones.

(D) Tracking of somatic mutations can be performed in any model organism. Somatic mutations arise spontaneously and accumulate over time, thereby marking single cells and all their progeny. Clones can be identified through whole-genome or targeted sequencing. Construction of multi-level phylogenetic trees is possible since mutations arise over a long period of time.

the initiation and progression of cancer. However, the maximum number of traceable unique clones with this system is ~50-100, depending on the technique and microscopy setup. Given the relatively limited number of fluorescent proteins, the number of clones that can be traced simultaneously is unlikely to increase dramatically. In a recently published technique, the fluorescent proteins have been replaced by DNA barcodes, using the flexibility of induction of the Cre-Lox system, but dramatically increasing the maximum number of clones that can be traced in parallel (Figure 3B) (Pei et al., 2017). In this work, a mouse model was created harboring an array of barcodes interspersed by LoxP sites (Polylox cassette). Recombination with the Cre recombinase results in the deletion or inversion of some barcodes, thereby introducing a cell-specific genetic label. Theoretical diversity of barcodes achieved through this recombination is greater than 10^6. Since this mouse model is easily bred with all sorts of tissue-specific and inducible Cre mouse models, researchers can induce labeling of thousands of unique clones

in a tissue and time-specific manner. Similarly to viral barcoding, the induction of the Polylox rearrangement can occur only once, therefore hindering the construction of multi-level phylogenetic trees based on these techniques.

CRISPR-Cas9 Genome Editing-Based Lineage Tracing

Another recently developed method for prospective lineage tracing uses CRISPR-Cas9-directed genome editing. Targeting of the Cas9 nuclease to a specific locus often results in the introduction of small insertions or deletions (indels) (Jao et al., 2013; Varshney et al., 2015). All the descendants of the founder cell inherit these indels, which therefore act as traceable elements. This approach has been used in zebrafish, in which 1-cell embryos were injected with CRISPR-Cas9 and guide RNAs (gRNAs) targeting previously introduced transgenes (Junker et al., 2017; McKenna et al., 2016). After injection of Cas9 mRNA into the yolk of the one cell embryo, scarring occurs for a period of 12 hr (Alemany et al., 2018), creating complex scarring patterns in many cells. This work aimed to study the lineage contribution of early embryonic cells to adult zebrafish organs, revealing that most organs derive from relatively few progenitors (Figure 3C) (Junker et al., 2017; McKenna et al., 2016). Since scarring transpires over the course of several hours, complex multi-level trees can be built. The multiple rounds of scarring increase the amount of information used for phylogenetic tree construction. Although scars created through CRISPR-Cas9 are highly complex, there exists a limited number of scars. This led researchers to the idea of designing a CRISPR-Cas9 system that targets its own gRNA sequence (Kalhor et al., 2017; Perli et al., 2016). This homing CRISPR barcode increases the complexity by several orders of magnitude as scarring occurs over a longer period of time. Using this self-targeting form of CRISPR-Cas9, a mouse model has been created to study clonal dynamics during early mouse development (Kalhor et al., 2018). Another technique, named CRISPR-UMI, combines the CRISPR-Cas9 approach with cellular barcoding in cell culture, which not only allows lineage tracing, but also creates mutations in a specific set of targeted genes. This enables researchers to study the effect of gene knockouts at a single-cell level and distinguishes between frequently occurring mutations with minor effects compared to infrequently occurring mutations with significant effects (Michlits et al., 2017).

Retrospective Lineage Tracing

All prospective lineage-tracing techniques require the introduction of exogenous material into the cells. Such techniques are impossible to implement in a completely wild-type setting, such as in human tissues. Furthermore, prospective lineage tracing can only start after the introduction of the genetic labels. These limitations of prospective lineage tracing can be overcome by using naturally occurring mutations (such as somatic mutations or copy number variations [CNVs]) or other inheritable elements that accumulate over time, eliminating the need for researchers to actively induce labeling. The significant advances in next-generation sequencing technologies can now harness naturally occurring mutations to perform lineage tracing. Similar to genetic barcodes or CRISPR-Cas9-induced scars, spontaneously occurring somatic mutations or CNVs are inherited by all the progeny of the mutated cell. Furthermore, the amount of mutations increases over time, providing an additional layer of information to the data (Figure 3D). However, naturally occurring mutations generally arise in low frequencies, thereby requiring high coverage sequencing data, especially for less frequent clones (Nik-Zainal et al., 2012).

Retrospective Lineage Tracing through CNVs

The most readily accessible genetic elements to use for single-cell retrospective lineage tracing are CNV. This is primarily due to the fact that relatively shallow sequencing data can easily identify CNVs (<1×), given that there is equal coverage across the genome with minimal amplification biases. Although CNVs are relatively uncommon in healthy tissue, several studies show the presence of CNVs in skin and brain (Cai et al., 2014; Knouse et al., 2016; Knouse et al., 2014; McConnell et al., 2013). Most CNVs found in these studies were unique to a single cell, although small clones of cells sharing CNVs were present, suggesting that these CNVs appeared during development of the tissue. Nevertheless, the low frequency of CNVs found in healthy somatic tissue does not allow for the construction of a phylogenetic tree that accurately represents tissue development. In contrast to healthy tissue, CNVs are often abundant in cancer and change dramatically during tumor progression, which makes an ideal system for lineage tracing purposes (Demeulemeester et al., 2016; Gawad et al., 2014; Navin et al., 2011; Wang et al., 2014).

Using single-cell DNA sequencing of CNVs, Navin et al. showed the presence of several genetically distinct clones in breast tumors, which probably arose during consecutive waves of clonal expansion (Navin et al., 2011; Wang et al., 2014). In these studies, enough CNVs were identified to allow accurate reconstruction of phylogenetic trees. A subsequent study from the Navin group incorporated a spatial component to single-cell whole-genome sequencing (scWGS) (Casasent et al., 2018). Here, using laser microdissection, the authors showed distinct CNV profiles between ductal carcinoma *in situ* (DCIS) and invasive ductal carcinoma (IDC), indicating genomic evolution precedes acquisition of invasive properties. In tumors with multiple IDC lesions, cells harboring the same set of CNVs were generally not spatially restricted, confirming the migratory nature of IDC cells.

Retrospective Lineage Tracing through SNV, Indels, and Repeat Regions

Another class of genetic elements typically used to perform lineage tracing include single-nucleotide variants (SNVs) and small indels. Both frequently exist in non-coding regions of the genome, exerting little to no effect on the function of the cell. Since SNVs are faithfully copied during DNA replication, SNVs have been successfully used for the reconstruction of phylogenetic trees of tumors from bulk DNA (Abbosh et al., 2017; Gao et al., 2016; Ju et al., 2017; Martincorena et al., 2015; Navin et al., 2010). In recent years, several studies have detected SNVs by scWGS or single-cell whole-exome sequencing (scWES) (Hou et al., 2012; Lodato et al., 2015; Wang et al., 2014; Xu et al., 2012; Zafar et al., 2016). The difficulty with using scWGS or scWES for lineage tracing lies in the sparse distribution of SNVs throughout the genome, which renders a substantial portion of the generated data unusable for lineage tracing due to the absence of SNVs. Furthermore, confident detection of SNVs is hampered by whole-genome amplification methods, which are error prone and fail to provide equal coverage across

Please cite this article in press as: Kester and van Oudenaarden, Single-Cell Transcriptomics Meets Lineage Tracing, Cell Stem Cell (2018), https://doi.org/10.1016/j.stem.2018.04.014

the genome (Gawad et al., 2016). Despite these difficulties, there are several examples where researchers managed to construct phylogenetic trees based on scWGS or scWES data from single cells (Ross and Markowetz, 2016; Xu et al., 2012). Additionally, the SNV detection problems can be circumvented by targeted sequencing on a predefined set of genomic loci that potentially harbor somatic mutations. With this approach, more single cells can be processed for the same cost as scWGS or scWES techniques, since the amount of data with extraneous information is drastically reduced. This approach has been used to reconstruct lineages from mature neurons (Lodato et al., 2015) and metastatic colorectal carcinoma (Leung et al., 2017).

Besides CNVs and SNVs, retrotransposon elements, such as LINE1 elements, and microsatellite repeats have also been used to perform lineage tracing. Up to 17% of the human genome consists of LINE1 elements, where most are stably integrated in the genome and a small, mobile fraction can jump from its original locus to a new one (Muotri et al., 2005; Ostertag and Kazazian, 2001). Attempts to trace the lineage of these LINE elements have been unsuccessful, as studies in neurons estimate the frequency of these events to be less than 1 event per cell (Evrony et al., 2015). Microsatellite regions are repeats of small DNA elements that can either grow or shrink in size due to DNA polymerase slippage during DNA replication. A number of studies have exploited the variation in the size of each of these microsatellite repeats as a heritable mark for lineage tracing (Frumkin et al., 2005; Reizel et al., 2011; Reizel et al., 2012; Salipante and Horwitz, 2006). A major benefit of using microsatellite repeats for lineage tracing is their well-defined positions along the genome, alleviating the need for expensive techniques such as scWGS or scWES.

Retrospective Lineage Tracing through Epigenetic Marks

Growing evidence for the use of epigenetic elements, such as DNA methylation and DNA hydroxymethylation as a marker for lineage tracing, has accumulated in recent years. DNA methylation is a chemical modification that mostly occurs on CpG dinucleotides in the genome; CpG methylation is faithfully maintained during cell division. DNA methylation has been established as an important regulator of gene expression with dramatic differences between different tissues (Smith and Meissner, 2013). There is evidence that DNA methylation patterns can change slowly over time, in a fashion that might be unrelated to cell function (Jones et al., 2015). Therefore, such genomic regions are well suited for lineage tracing studies.

Another epigenetic mark that could have potential for lineage tracing is DNA hydroxymethylation. After DNA replication, hydroxymethylation is not copied onto the newly synthesized strand. Subsequently, one daughter cell inherits the old highly hydroxymethylated strand, while the other daughter cell inherits the new lowly hydroxymethylated strand. In a recent study, Mooijman et al. (2016) have shown that strand bias, the amount of hydroxymethylation on the positive strand of the DNA divided by the total amount of hydroxymethylation, can be used as a marker for lineage tracing in early mouse embryos. They found that strand bias patterns were faithfully mirrored in sister cells, which can be explained by the slow conversion of 5-methylcytosine into 5-hydroxymethylcytosine.

Combinatorial Techniques for Transcriptomics and Lineage Tracing

All of the above methods for genetic lineage tracing have provided invaluable new information on the development of tissue in health and disease. However, the main focus of these techniques is the genotypic relation of cells to each other. In other words, these methods address the question of whether cells originate from the same ancestor but lack any information on the phenotypic state of the cell. Some of the techniques measure a few markers that reveal the phenotypic state of the cell; however, this cannot rival the richness of information acquired through single-cell transcriptomics. Ideally one would measure both the genotypic and phenotypic state of the cell by combining the techniques described above. This will allow for refinement of the phylogenetic trees based on genetic labels by incorporation of differentiation trajectories acquired through transcriptomics. Furthermore, this will provide a detailed insight of transcriptional changes that occur as cells move along the branches of the phylogenetic trees (Figure 1). In the last two years, a handful of studies have been published that combine lineage tracing with single-cell transcriptomics. These studies serve as an example of the power achieved by the convergence between single-cell transcriptomics and single-cell lineage tracing techniques.

Short-Term Lineage Tracing and Transcriptomics to Assess Cell-State Transitions

The first two examples of these combinatorial studies use imaging to perform lineage tracing and obtain phenotypic information from either single-cell mRNA sequencing or single-molecule fluorescent *in situ* hybridization (smFISH) (Hormoz et al., 2016; Kimmerling et al., 2016). In the first of these studies, the authors were able to assess the embryonic stem cell (ESC) state and quantify cell-state transitions between different ESC states (Hormoz et al., 2016). In the second study, the biggest determinant of transcriptome heterogeneity was the amount of time since the last cell division rather than the lineage relation between cells (Kimmerling et al., 2016). This suggests that, at least on a short timescale, cell intrinsic processes, such as cell division, confound the influence of genotypic relations on the transcriptome level.

Memory by engineered mutagenesis with optical *in situ* readout (MEMOIR) is another technique that combines the power of genetic lineage tracing with the acquisition of phenotypic information (Frieda et al., 2017). Here, a collection of transgenes that consist of multiple gRNA target sequences is integrated into the genome. These transgenes are expressed in cells and can be visualized using seqFISH, a high throughput adaptation of smFISH (Lubeck et al., 2014; Shah et al., 2016). Upon the expression of a collection of gRNAs targeting each of the transgenes, some of the transgenes will collapse and therefore can no longer be identified through seqFISH. However, there will be cell-to-cell variability in which target sequences collapse. Therefore the combination of collapsed and uncollapsed gRNA sequences in each single cell acts as a heritable barcode, which can be used for lineage tracing. MEMOIR can be combined with seqFISH for the detection of regular RNA molecules, allowing the acquisition of phenotypic information. The advantage of MEMOIR is that readout of the lineage and the transcriptome occurs *in situ*, which maintains the spatial relation of cells, thereby adding another dimension of complexity to the data.

Please cite this article in press as: Kester and van Oudenaarden, Single-Cell Transcriptomics Meets Lineage Tracing, Cell Stem Cell (2018), https://doi.org/10.1016/j.stem.2018.04.014

Investigating Reprogramming Efficiency by Combining Viral Barcoding with Single-Cell Transcriptomics

As mentioned earlier, lineage tracing is often done through viral barcoding of cell populations . The first study to combine viral barcoding with single-cell transcriptomics utilized a 10-nt random lineage barcode in the 3′ UTR of a tdTomato gene that is easily detected by single-cell mRNA sequencing (Yao et al., 2017). Viruses harboring the barcoded tdTomato gene were used to infect cells at day 26 of an *in vitro* model for neural differentiation. At day 54, the barcodes showed a clonal segregation of cortical precursors and mid/hindbrain precursors. This suggests that specification of those cell types occurred prior to labeling (day 26), which is confirmed by a transcriptome-derived lineage tree that also predicts segregation of cortical and mid/hindbrain precursors occurring prior to day 26. In a similar study (Biddy et al., 2017), endodermal reprogramming of fibroblast cells was followed over a period of 4 weeks. Reprogramming of these fibroblasts was shown to be highly heterogeneous. Many different end states of reprogramming were observed even within a single clone. Despite the successful reprogramming of many cells in the concluding experiment, the clonal diversity was limited, indicating that most of the reprogrammed cells came from the same founder cell.

Delineating Zebrafish Development through a Combination of Single-Cell Transcriptomics and CRISPR-Cas9-Based Lineage Tracing

As described above, CRISPR-Cas9-induced scars can be used as a heritable barcode throughout development. The first three studies uniting the power of CRISPR-Cas9-induced scars with single-cell transcriptomics for lineage tracing have recently been published.

In the first of these studies, Spanjaard et al. (2018) induce scars in a red fluorescent protein (RFP) transgene in zebrafish embryos. Once most of the organs have formed 5 days post-fertilization, scars and transcriptomes are acquired using single-cell mRNA sequencing. Phylogenetic trees can then be constructed based on the distribution of scars across all cells, i.e., cells sharing multiple scars are probably more closely related than cells not sharing any scars. Cell-type inference is performed based on the transcriptome of the same single cells, allowing the authors to investigate the genetic relation between cells from different germ layers, organs, and even cell types within an organ. This study confirms the early split of endoderm, mesoderm, and ectoderm-derived lineages during zebrafish development. Furthermore, in depth analysis of hearth development showed an early separation in the lineages contributing to the myocardial and endocardial lineages.

In a similar study, Raj et al. (2018) investigate the development of the zebrafish brain. The authors developed a heat-shock-inducible version of the GESTALT fish (McKenna et al., 2016), termed scGESTALT, allowing scars to be introduced at multiple time points. The first iteration of scarring takes place upon injection of a set of 4 gRNAs (gRNAs 1–4) into the fertilized oocyte. The second iteration, takes place upon heat-shock induction of gRNAs 5–9. The two iterations of scarring allow for increased resolution throughout development. Scars are identified from the mRNA of the transgene harboring the 9 gRNA target sequences. Phylogenetic tree reconstruction shows that the major areas of the brain (left and right forebrain, left and right hindbrain) contain distinct clones, indicating early cell-fate specification for those areas. Applying Monocle2 to the data reveals a trajectory from neural progenitors into cerebral granule cells. The overlaying of the genetic lineage information from the scars shows that cells from the same clone populate the entire trajectory, indicating that the clones that gave rise to the mature cells are still present in the pool of progenitors.

In a third study, Alemany et al. (2018) use a similar system to study the clonal history of the hematopoietic system, different regions of the brain, the eyes, and the caudal fin. The authors inject either Cas9 protein or Cas9 mRNA into the fertilized embryo, inducing scars until 3 or 12 hr post fertilization (hpf) respectively. In the hematopoietic system, most clones are dispersed over all cell types, indicating that progenitors of the hematopoietic system arise once scarring is finished after 12 hpf. Furthermore, only a handful of distinct clones contribute to the hematopoietic system. This finding indicates that at the end of scarring (12 hpf), only a small proportion of cells give rise to the hematopoietic system but not to any of the other organs. In contrast to the hematopoietic system, where clones are shared across all cell types, clones in the midbrain are specific to either the left or the right half when scarring occurs until 12 hpf, whereas clones are shared between the brain halves when scarring occurs until 3 hpf. For the eyes, there is no overlap in clones between left and right eye independent of protein or mRNA Cas9 injection. This suggests that there is a difference in timing between specification of cells contributing to the eyes, the midbrain, and the hematopoietic system. Another interesting finding is that a portion of the macrophages in the zebrafish fin contain no overlapping scars with the rest of the hematopoietic system but do share scars with mesenchymal and epidermal cell types, suggesting that these resident macrophages might be of non-hematopoietic origin. These three studies are the first *in vivo* examples that combine single-cell lineage tracing of many parallel lineages with single-cell transcriptomics and provide a glimpse of what can be achieved with integrated techniques.

In vivo genetic lineage tracing through CRISPR-Cas9 in combination with single-cell transcriptomics has so far only been described in zebrafish. However, a mouse model harboring the homing CRISPR lineage tracing system has been developed (MARC1 mouse) (Kalhor et al., 2018). Although single-cell scar detection and transcriptomics have not been achieved, this will most likely be developed in the near future. The Polylox system is also promising as it can be easily combined with existing Cre lines to perform lineage tracing in mouse models.

The combinatorial strategies mentioned here all use prospective lineage tracing, which requires genome editing to introduce the traceable elements, limiting the use of these technologies to model organisms or *in vitro* systems. Therefore, the next big step in lineage tracing will result from the combination of retrospective genetic lineage tracing techniques with single-cell transcriptomics. This will allow researchers to perform lineage tracing and transcriptomics without altering the genome of the study subject, which makes it possible to use human samples and opens up many new avenues for studying human development and disease.

Future Directions

Combinatorial methods for single-cell transcriptomics and genetic lineage tracing are only just starting to be used. In the

near future, their use will probably expand from cell culture models and zebrafish to other model organisms and human samples. One of the many fields where this combinatorial technique could play an important role is in cancer research. As described above, tumors are heterogeneous in their gene-expression profiles (Patel et al., 2014; Tirosh et al., 2016a, 2016b) and often contain multiple distinct genetic clones (Casasent et al., 2018; Gao et al., 2016; Navin et al., 2011; Wang et al., 2014). However, the extent to which distinct genetic clones differ in their transcriptome remains largely unknown. Furthermore, combinatorial techniques might prove greatly useful in understanding clonal dynamics, drug resistance, and metastatic properties of cancer cells while relating those properties to the cancer cells transcriptome.

Another field where combinatorial methods could provide invaluable new insights is in understanding the dynamics of tissue homeostasis. For instance, the classical differentiation tree through which hematopoietic stem cells differentiate into the mature cell types of the blood has recently come under debate. Both single-cell transcriptomics studies (Jaitin et al., 2014; Paul et al., 2015) and single-cell lineage tracing studies (Rodriguez-Fraticelli et al., 2018; Sun et al., 2014), have shown that the classical lineage tree, where stem cells have unlimited potential and the multiple progenitor populations each have a predetermined fate, may be inaccurate. It is becoming increasingly apparent that, even within the stem cell compartment, cells have a bias toward a certain fate, while progenitor populations are not as restricted in their fate as previously thought. Therefore, the idea of a differentiation tree in which all of the stem and progenitor populations are separated and differentiation occurs as discrete steps along the tree is changing to a tree where differentiation is a continuous process, with stem and progenitor cells being biased toward a certain fate (Laurenti and Göttgens, 2018). Furthermore, both single-cell transcriptomics (Jaitin et al., 2014; Paul et al., 2015) and single-cell genetic lineage tracing studies (Rodriguez-Fraticelli et al., 2018; Sun et al., 2014) have strongly suggested that hematopoietic differentiation in steady-state homeostasis is different from differentiation upon transplantation (Carrelha et al., 2018). Combinatorial methods will aid further investigation, since the single-cell gene-expression profiles allow for detailed cell-type classification while the genetic lineage tracing reveals the clonal relationships of the cells.

Concluding Remarks

The development of next-generation sequencing technology has allowed for tremendous advances both in the field of lineage tracing and the field of single-cell transcriptomics. Acquiring single-cell transcriptome profiles from thousands of cells can be accomplished within just a few days. To gain knowledge from these datasets, many algorithms have been developed that predict lineage trajectories based on single-cell transcriptome profiles. This advance allows for a more detailed reconstruction of tissue development and homeostasis despite not necessarily reflecting the underlying genetic relationships between cells. Advances in single-cell genetic lineage tracing now allow for the tracking of thousands of single cells in several different model organisms. Over the past year, several publications showcase the true potential of combining transcriptomics with genetic lineage tracing at single-cell resolution. These techniques are powerful new additions to the biologists' toolbox and will contribute to better understanding of embryonic development, tissue homeostasis, and pathologies.

ACKNOWLEDGMENTS

Research in the authors' lab is supported by a European Research Council Advanced grant (ERC-AdG 742225-IntScOmics), Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) TOP award (NWO-CW 714.016.001), and a Dutch Cancer Society project grant (KWF project 10158/2016-1). This work is part of the Oncode Institute, which is partly financed by the Dutch Cancer Society. We'd like to thank M. Sen and A. Alemany for the careful reading and correcting of the manuscript.

REFERENCES

Abbosh, C., Birkbak, N.J., Wilson, G.A., Jamal-Hanjani, M., Constantin, T., Salari, R., Le Quesne, J., Moore, D.A., Veeriah, S., Rosenthal, R., et al.; TRACERx consortium; PEACE consortium (2017). Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. Nature *545*, 446–451.

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.

Bardy, C., van den Hurk, M., Kakaradov, B., Erwin, J.A., Jaeger, B.N., Hernandez, R.V., Eames, T., Paucar, A.A., Gorris, M., Marchand, C., et al. (2016). Predicting the functional states of human iPSC-derived neurons with single-cell RNA-seq and electrophysiology. Mol. Psychiatry *21*, 1573–1588.

Bendall, S.C., Davis, K.L., Amir, A.D., Tadmor, M.D., Simonds, E.F., Chen, T.J., Shenfeld, D.K., Nolan, G.P., and Pe'er, D. (2014). Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. Cell *157*, 714–725.

Bhang, H.E., Ruddy, D.A., Krishnamurthy Radhakrishna, V., Caushi, J.X., Zhao, R., Hims, M.M., Singh, A.P., Kao, I., Rakiec, D., Shaw, P., et al. (2015). Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. Nat. Med. *21*, 440–448.

Biddy, B.A., Waye, S.E., Sun, T., and Morris, S.A. (2017). Single-cell analysis of clonal dynamics in direct lineage reprogramming: A combinatorial indexing method for lineage tracing. bioRxiv. Published online April 28, 2017. https://doi.org/10.1101/127860.

Brady, S.W., McQuerry, J.A., Qiao, Y., Piccolo, S.R., Shrestha, G., Jenkins, D.F., Layer, R.M., Pedersen, B.S., Miller, R.H., Esch, A., et al. (2017). Combating subclonal evolution of resistant cancer phenotypes. Nat. Commun. *8*, 1231.

Cai, X., Evrony, G.D., Lehmann, H.S., Elhosary, P.C., Mehta, B.K., Poduri, A., and Walsh, C.A. (2014). Single-cell, genome-wide sequencing identifies clonal somatic copy-number variation in the human brain. Cell Rep. *8*, 1280–1289.

Camp, J.G., Badsha, F., Florio, M., Kanton, S., Gerber, T., Wilsch-Bräuninger, M., Lewitus, E., Sykes, A., Hevers, W., Lancaster, M., et al. (2015). Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. Proc. Natl. Acad. Sci. USA *112*, 15672–15677.

Cannoodt, R., Saelens, W., and Saeys, Y. (2016). Computational methods for trajectory inference from single-cell transcriptomics. Eur. J. Immunol. *46*, 2496–2506.

Cao, J., Packer, J.S., Ramani, V., Cusanovich, D.A., Huynh, C., Daza, R., Qiu, X., Lee, C., Furlan, S.N., Steemers, F.J., et al. (2017). Comprehensive single-cell transcriptional profiling of a multicellular organism. Science *357*, 661–667.

Carrelha, J., Meng, Y., Kettyle, L.M., Luis, T.C., Norfo, R., Alcolea, V., Boukarabila, H., Grasso, F., Gambardella, A., Grover, A., et al. (2018). Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. Nature *554*, 106–111.

Casasent, A.K., Schalck, A., Gao, R., Sei, E., Long, A., Pangburn, W., Casasent, T., Meric-Bernstam, F., Edgerton, M.E., and Navin, N.E. (2018). Multiclonal invasion in breast tumors identified by topographic single cell sequencing. Cell *172*, 205–217.

Chen, J., Schlitzer, A., Chakarov, S., Ginhoux, F., and Poidinger, M. (2016). Mpath maps multi-branching single-cell trajectories revealing progenitor cell progression during development. Nat. Commun. *7*, 11988.

Chung, W., Eum, H.H., Lee, H.O., Lee, K.M., Lee, H.B., Kim, K.T., Ryu, H.S., Kim, S., Lee, J.E., Park, Y.H., et al. (2017). Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. Nat. Commun. *8*, 15081.

Demeulemeester, J., Kumar, P., Møller, E.K., Nord, S., Wedge, D.C., Peterson, A., Mathiesen, R.R., Fjelldal, R., Zamani Esteki, M., Theunis, K., et al. (2016). Tracing the origin of disseminated tumor cells in breast cancer using single-cell sequencing. Genome Biol. *17*, 250.

Deppe, U., Schierenberg, E., Cole, T., Krieg, C., Schmitt, D., Yoder, B., and von Ehrenstein, G. (1978). Cell lineages of the embryo of the nematode Caenorhabditis elegans. Proc. Natl. Acad. Sci. USA *75*, 376–380.

Dulken, B.W., Leeman, D.S., Boutet, S.C., Hebestreit, K., and Brunet, A. (2017). Single-cell transcriptomic analysis defines heterogeneity and transcriptional dynamics in the adult neural stem cell lineage. Cell Rep. *18*, 777–790.

Eirew, P., Steif, A., Khattra, J., Ha, G., Yap, D., Farahani, H., Gelmon, K., Chia, S., Mar, C., Wan, A., et al. (2015). Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. Nature *518*, 422–426.

Evrony, G.D., Lee, E., Mehta, B.K., Benjamini, Y., Johnson, R.M., Cai, X., Yang, L., Haseley, P., Lehmann, H.S., Park, P.J., and Walsh, C.A. (2015). Cell lineage analysis in human brain using endogenous retroelements. Neuron *85*, 49–59.

Fletcher, R.B., Das, D., Gadye, L., Street, K.N., Baudhuin, A., Wagner, A., Cole, M.B., Flores, Q., Choi, Y.G., Yosef, N., et al. (2017). Deconstructing olfactory stem cell trajectories at single-cell resolution. Cell Stem Cell *20*, 817–830.

Frank, E., and Sanes, J.R. (1991). Lineage of neurons and glia in chick dorsal root ganglia: Analysis in vivo with a recombinant retrovirus. Development *111*, 895–908.

Frieda, K.L., Linton, J.M., Hormoz, S., Choi, J., Chow, K.K., Singer, Z.S., Budde, M.W., Elowitz, M.B., and Cai, L. (2017). Synthetic recording and in situ readout of lineage information in single cells. Nature *541*, 107–111.

Frumkin, D., Wasserstrom, A., Kaplan, S., Feige, U., and Shapiro, E. (2005). Genomic variability within an organism exposes its cell lineage tree. PLoS Comput. Biol. *1*, 50.

Furchtgott, L.A., Melton, S., Menon, V., and Ramanathan, S. (2017). Discovering sparse transcription factor codes for cell states and state transitions during development. eLife *6*. Published online March 15, 2017. https://doi.org/10.7554/eLife.20488.

Gao, R., Davis, A., McDonald, T.O., Sei, E., Shi, X., Wang, Y., Tsai, P.C., Casasent, A., Waters, J., Zhang, H., et al. (2016). Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. Nat. Genet. *48*, 1119–1130.

Gawad, C., Koh, W., and Quake, S.R. (2014). Dissecting the clonal origins of childhood acute lymphoblastic leukemia by single-cell genomics. Proc. Natl. Acad. Sci. USA *111*, 17947–17952.

Gawad, C., Koh, W., and Quake, S.R. (2016). Single-cell genome sequencing: Current state of the science. Nat. Rev. Genet. *17*, 175–188.

Gebara, E., Bonaguidi, M.A., Beckervordersandforth, R., Sultan, S., Udry, F., Gijs, P.J., Lie, D.C., Ming, G.L., Song, H., and Toni, N. (2016). Heterogeneity of radial glia-like cells in the adult hippocampus. Stem Cells *34*, 997–1010.

Gerlach, C., Rohr, J.C., Perié, L., van Rooij, N., van Heijst, J.W., Velds, A., Urbanus, J., Naik, S.H., Jacobs, H., Beltman, J.B., et al. (2013). Heterogeneous differentiation patterns of individual CD8+ T cells. Science *340*, 635–639.

Gerrits, A., Dykstra, B., Kalmykowa, O.J., Klauke, K., Verovskaya, E., Broekhuis, M.J., de Haan, G., and Bystrykh, L.V. (2010). Cellular barcoding tool for clonal analysis in the hematopoietic system. Blood *115*, 2610–2618.

Grün, D., and van Oudenaarden, A. (2015). Design and analysis of single-cell sequencing experiments. Cell *163*, 799–810.

Grün, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., Clevers, H., and van Oudenaarden, A. (2015). Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature *525*, 251–255.

Grün, D., Muraro, M.J., Boisset, J.C., Wiebrands, K., Lyubimova, A., Dharmadhikari, G., van den Born, M., van Es, J., Jansen, E., Clevers, H., et al. (2016). De novo prediction of stem cell identity using single-cell transcriptome data. Cell Stem Cell *19*, 266–277.

Guo, M., Bao, E.L., Wagner, M., Whitsett, J.A., and Xu, Y. (2017). SLICE: Determining cell differentiation and lineage based on single cell entropy. Nucleic Acids Res. *45*, 54.

Haber, A.L., Biton, M., Rogel, N., Herbst, R.H., Shekhar, K., Smillie, C., Burgin, G., Delorey, T.M., Howitt, M.R., Katz, Y., et al. (2017). A single-cell survey of the small intestinal epithelium. Nature *551*, 333–339.

Hanchate, N.K., Kondoh, K., Lu, Z., Kuang, D., Ye, X., Qiu, X., Pachter, L., Trapnell, C., and Buck, L.B. (2015). Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. Science *350*, 1251–1255.

Haque, A., Engel, J., Teichmann, S.A., and Lönnberg, T. (2017). A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. Genome Med. *9*, 75.

Hashimshony, T., Wagner, F., Sher, N., and Yanai, I. (2012). CEL-seq: Single-cell RNA-seq by multiplexed linear amplification. Cell Rep. *2*, 666–673.

Hashimshony, T., Senderovich, N., Avital, G., Klochendler, A., de Leeuw, Y., Anavy, L., Gennert, D., Li, S., Livak, K.J., Rozenblatt-Rosen, O., et al. (2016). CEL-Seq2: Sensitive highly-multiplexed single-cell RNA-Seq. Genome Biol. *17*, 77.

Heinäniemi, M., Nykter, M., Kramer, R., Wienecke-Baldacchino, A., Sinkkonen, L., Zhou, J.X., Kreisberg, R., Kauffman, S.A., Huang, S., and Shmulevich, I. (2013). Gene-pair expression signatures reveal lineage control. Nat. Methods *10*, 577–583.

Hormoz, S., Singer, Z.S., Linton, J.M., Antebi, Y.E., Shraiman, B.I., and Elowitz, M.B. (2016). Inferring cell-state transition dynamics from lineage trees and endpoint single-cell measurements. Cell Syst. *3*, 419–433.

Hou, Y., Song, L., Zhu, P., Zhang, B., Tao, Y., Xu, X., Li, F., Wu, K., Liang, J., Shao, D., et al. (2012). Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. Cell *148*, 873–885.

Insel, T.R., Landis, S.C., and Collins, F.S.; The NIH BRAIN Initiative (2013). Research priorities. Science *340*, 687–688.

Islam, S., Zeisel, A., Joost, S., La Manno, G., Zajac, P., Kasper, M., Lönnerberg, P., and Linnarsson, S. (2014). Quantitative single-cell RNA-seq with unique molecular identifiers. Nat. Methods *11*, 163–166.

Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., Mildner, A., Cohen, N., Jung, S., Tanay, A., and Amit, I. (2014). Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science *343*, 776–779.

Jao, L.E., Wente, S.R., and Chen, W. (2013). Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proc. Natl. Acad. Sci. USA *110*, 13904–13909.

Ji, Z., and Ji, H. (2016). TSCAN: Pseudo-time reconstruction and evaluation in single-cell RNA-seq analysis. Nucleic Acids Res. *44*, 117.

Jones, M.J., Goodman, S.J., and Kobor, M.S. (2015). DNA methylation and healthy human aging. Aging Cell *14*, 924–932.

Joost, S., Zeisel, A., Jacob, T., Sun, X., La Manno, G., Lönnerberg, P., Linnarsson, S., and Kasper, M. (2016). Single-cell transcriptomics reveals that differentiation and spatial signatures shape epidermal and hair follicle heterogeneity. Cell Syst. *3*, 221–237.

Ju, Y.S., Martincorena, I., Gerstung, M., Petljak, M., Alexandrov, L.B., Rahbari, R., Wedge, D.C., Davies, H.R., Ramakrishna, M., Fullam, A., et al. (2017). Somatic mutations reveal asymmetric cellular dynamics in the early human embryo. Nature *543*, 714–718.

Junker, J.P., Spanjaard, B., Peterson-Maduro, J., Alemany, A., Hu, B., Florescu, M., and van Oudenaarden, A. (2017). Massively parallel clonal analysis using CRISPR/Cas9 induced genetic scars. bioRxiv. Published online January 4, 2017. https://doi.org/10.1101/056499.

Kalhor, R., Mali, P., and Church, G.M. (2017). Rapidly evolving homing CRISPR barcodes. Nat. Methods *14*, 195–200.

Kalhor, R., Kalhor, K., Leeper, K., Graveline, A., Mali, P., and Church, G.M. (2018). A homing CRISPR mouse resource for barcoding and lineage tracing. bioRxiv. Published online March 12, 2018. https://doi.org/10.1101/280289.

Kim, T.H., Saadatpour, A., Guo, G., Saxena, M., Cavazza, A., Desai, N., Jadhav, U., Jiang, L., Rivera, M.N., Orkin, S.H., et al. (2016). Single-cell transcript profiles reveal multilineage priming in early progenitors derived from Lgr5(+) intestinal stem cells. Cell Rep. *16*, 2053–2060.

Kimmerling, R.J., Lee Szeto, G., Li, J.W., Genshaft, A.S., Kazer, S.W., Payer, K.R., de Riba Borrajo, J., Blainey, P.C., Irvine, D.J., Shalek, A.K., and Manalis, S.R. (2016). A microfluidic platform enabling single-cell RNA-seq of multigenerational lineages. Nat. Commun. *7*, 10220.

Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell *161*, 1187–1201.

Knouse, K.A., Wu, J., Whittaker, C.A., and Amon, A. (2014). Single cell sequencing reveals low levels of aneuploidy across mammalian tissues. Proc. Natl. Acad. Sci. USA *111*, 13409–13414.

Knouse, K.A., Wu, J., and Amon, A. (2016). Assessment of megabase-scale somatic copy number variation using single-cell sequencing. Genome Res. *26*, 376–384.

La Manno, G., Soldatov, R., Hochgerner, H., Zeisel, A., Petukhov, V., Kastriti, M., Lonnerberg, P., Furlan, A., Fan, J., Liu, Z., et al. (2017). RNA velocity in single cells. bioRxiv. Published online October 19, 2017. https://doi.org/10.1101/206052.

Lacar, B., Linker, S.B., Jaeger, B.N., Krishnaswami, S., Barron, J., Kelder, M., Parylak, S., Paquola, A., Venepally, P., Novotny, M., et al. (2016). Nuclear RNA-seq of single neurons reveals molecular signatures of activation. Nat. Commun. *7*, 11022.

Laurenti, E., and Göttgens, B. (2018). From haematopoietic stem cells to complex differentiation landscapes. Nature *553*, 418–426.

Lawson, K.A., Meneses, J.J., and Pedersen, R.A. (1986). Cell fate and cell lineage in the endoderm of the presomite mouse embryo, studied with an intracellular tracer. Dev. Biol. *115*, 325–339.

Le Douarin, N.M., and Teillet, M.A. (1973). The migration of neural crest cells to the wall of the digestive tract in avian embryo. J. Embryol. Exp. Morphol. *30*, 31–48.

Lee, T., and Luo, L. (1999). Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. Neuron *22*, 451–461.

Leung, M.L., Davis, A., Gao, R., Casasent, A., Wang, Y., Sei, E., Vilar, E., Maru, D., Kopetz, S., and Navin, N.E. (2017). Single-cell DNA sequencing reveals a late-dissemination model in metastatic colorectal cancer. Genome Res. *27*, 1287–1299.

Lisi, V., Singh, B., Giroux, M., Guzman, E., Painter, M.W., Cheng, Y.C., Huebner, E., Coppola, G., Costigan, M., Woolf, C.J., and Kosik, K.S. (2017). Enhanced neuronal regeneration in the CAST/Ei mouse strain is linked to expression of differentiation markers after injury. Cell Rep. *20*, 1136–1147.

Lodato, M.A., Woodworth, M.B., Lee, S., Evrony, G.D., Mehta, B.K., Karger, A., Lee, S., Chittenden, T.W., D'Gama, A.M., Cai, X., et al. (2015). Somatic mutation in single human neurons tracks developmental and transcriptional history. Science *350*, 94–98.

Lu, R., Neff, N.F., Quake, S.R., and Weissman, I.L. (2011). Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. Nat. Biotechnol. *29*, 928–933.

Lubeck, E., Coskun, A.F., Zhiyentayev, T., Ahmad, M., and Cai, L. (2014). Single-cell in situ RNA profiling by sequential hybridization. Nat. Methods *11*, 360–361.

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell *161*, 1202–1214.

Marco, E., Karp, R.L., Guo, G., Robson, P., Hart, A.H., Trippa, L., and Yuan, G.C. (2014). Bifurcation analysis of single-cell gene expression data reveals epigenetic landscape. Proc. Natl. Acad. Sci. USA *111*, 5643–5650.

Marques, S., Zeisel, A., Codeluppi, S., van Bruggen, D., Mendanha Falcão, A., Xiao, L., Li, H., Häring, M., Hochgerner, H., Romanov, R.A., et al. (2016). Oligodendrocyte heterogeneity in the mouse juvenile and adult central nervous system. Science *352*, 1326–1329.

Martincorena, I., Roshan, A., Gerstung, M., Ellis, P., Van Loo, P., McLaren, S., Wedge, D.C., Fullam, A., Alexandrov, L.B., Tubio, J.M., et al. (2015). Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. Science *348*, 880–886.

Mass, E., Ballesteros, I., Farlik, M., Halbritter, F., Günther, P., Crozet, L., Jacome-Galarza, C.E., Händler, K., Klughammer, J., Kobayashi, Y., et al. (2016). Specification of tissue-resident macrophages during organogenesis. Science *353*, 353.

McConnell, M.J., Lindberg, M.R., Brennand, K.J., Piper, J.C., Voet, T., Cowing-Zitron, C., Shumilina, S., Lasken, R.S., Vermeesch, J.R., Hall, I.M., and Gage, F.H. (2013). Mosaic copy number variation in human neurons. Science *342*, 632–637.

McKenna, A., Findlay, G.M., Gagnon, J.A., Horwitz, M.S., Schier, A.F., and Shendure, J. (2016). Whole-organism lineage tracing by combinatorial and cumulative genome editing. Science *353*, aaf7907.

Michlits, G., Hubmann, M., Wu, S.H., Vainorius, G., Budusan, E., Zhuk, S., Burkard, T.R., Novatchkova, M., Aichinger, M., Lu, Y., et al. (2017). CRISPR-UMI: Single-cell lineage tracing of pooled CRISPR-Cas9 screens. Nat. Methods *14*, 1191–1197.

Mintz, B. (1967). Gene control of mammalian pigmentary differentiation. I. Clonal origin of melanocytes. Proc. Natl. Acad. Sci. USA *58*, 344–351.

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.

Muotri, A.R., Chu, V.T., Marchetto, M.C., Deng, W., Moran, J.V., and Gage, F.H. (2005). Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature *435*, 903–910.

Muraro, M.J., Dharmadhikari, G., Grün, D., Groen, N., Dielen, T., Jansen, E., van Gurp, L., Engelse, M.A., Carlotti, F., de Koning, E.J., and van Oudenaarden, A. (2016). A single-cell transcriptome atlas of the human pancreas. Cell Syst. *3*, 385–394.

Naik, S.H., Perié, L., Swart, E., Gerlach, C., van Rooij, N., de Boer, R.J., and Schumacher, T.N. (2013). Diverse and heritable lineage imprinting of early haematopoietic progenitors. Nature *496*, 229–232.

Navin, N., Krasnitz, A., Rodgers, L., Cook, K., Meth, J., Kendall, J., Riggs, M., Eberling, Y., Troge, J., Grubor, V., et al. (2010). Inferring tumor progression from genomic heterogeneity. Genome Res. *20*, 68–80.

Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D., et al. (2011). Tumour evolution inferred by single-cell sequencing. Nature *472*, 90–94.

Nguyen, L.V., Cox, C.L., Eirew, P., Knapp, D.J., Pellacani, D., Kannan, N., Carles, A., Moksa, M., Balani, S., Shah, S., et al. (2014). DNA barcoding reveals diverse growth kinetics of human breast tumour subclones in serially passaged xenografts. Nat. Commun. *5*, 5871.

Nik-Zainal, S., Van Loo, P., Wedge, D.C., Alexandrov, L.B., Greenman, C.D., Lau, K.W., Raine, K., Jones, D., Marshall, J., Ramakrishna, M., et al.; Breast Cancer Working Group of the International Cancer Genome Consortium (2012). The life history of 21 breast cancers. Cell *149*, 994–1007.

Osawa, M., Hanada, K., Hamada, H., and Nakauchi, H. (1996). Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. Science *273*, 242–245.

Ostertag, E.M., and Kazazian, H.H., Jr. (2001). Biology of mammalian L1 retrotransposons. Annu. Rev. Genet. *35*, 501–538.

Palpant, N.J., Wang, Y., Hadland, B., Zaunbrecher, R.J., Redd, M., Jones, D., Pabon, L., Jain, R., Epstein, J., Ruzzo, W.L., et al. (2017). Chromatin and transcriptional analysis of mesoderm progenitor cells identifies HOPX as a regulator of primitive hematopoiesis. Cell Rep. *20*, 1597–1608.

Papalexi, E., and Satija, R. (2018). Single-cell RNA sequencing to explore immune cell heterogeneity. Nat. Rev. Immunol. *18*, 35–45.

Patel, A.P., Tirosh, I., Trombetta, J.J., Shalek, A.K., Gillespie, S.M., Wakimoto, H., Cahill, D.P., Nahed, B.V., Curry, W.T., Martuza, R.L., et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science *344*, 1396–1401.

Paul, F., Arkin, Y., Giladi, A., Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Winter, D., Lara-Astiaso, D., Gury, M., Weiner, A., et al. (2015). Transcriptional heterogeneity and lineage commitment in myeloid progenitors. Cell *163*, 1663–1677.

Pedersen, R.A., Wu, K., and Bałakier, H. (1986). Origin of the inner cell mass in mouse embryos: Cell lineage analysis by microinjection. Dev. Biol. *117*, 581–595.

Pei, W., Feyerabend, T.B., Rössler, J., Wang, X., Postrach, D., Busch, K., Rode, I., Klapproth, K., Dietlein, N., Quedenau, C., et al. (2017). Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. Nature *548*, 456–460.

Perli, S.D., Cui, C.H., and Lu, T.K. (2016). Continuous genetic recording with self-targeting CRISPR-Cas in human cells. Science *353*, 353.

Picelli, S., Björklund, A.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.

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. Nat. Methods *14*, 979–982.

Quintana, E., Shackleton, M., Sabel, M.S., Fullen, D.R., Johnson, T.M., and Morrison, S.J. (2008). Efficient tumour formation by single human melanoma cells. Nature *456*, 593–598.

Raj, B., Wagner, D.E., McKenna, A., Pandey, S., Klein, A.M., Shendure, J., Gagnon, J.A., and Schier, A.F. (2018). Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. Nat. Biotechnol.

Reizel, Y., Chapal-Ilani, N., Adar, R., Itzkovitz, S., Elbaz, J., Maruvka, Y.E., Segev, E., Shlush, L.I., Dekel, N., and Shapiro, E. (2011). Colon stem cell and crypt dynamics exposed by cell lineage reconstruction. PLoS Genet. *7*, e1002192.

Reizel, Y., Itzkovitz, S., Adar, R., Elbaz, J., Jinich, A., Chapal-Ilani, N., Maruvka, Y.E., Nevo, N., Marx, Z., Horovitz, I., et al. (2012). Cell lineage analysis of the mammalian female germline. PLoS Genet. *8*, e1002477.

Rodriguez-Fraticelli, A.E., Wolock, S.L., Weinreb, C.S., Panero, R., Patel, S.H., Jankovic, M., Sun, J., Calogero, R.A., Klein, A.M., and Camargo, F.D. (2018). Clonal analysis of lineage fate in native haematopoiesis. Nature *553*, 212–216.

Rosenberg, A.B., Roco, C.M., Muscat, R.A., Kuchina, A., Sample, P., Yao, Z., Gray, L., Peeler, D.J., Mukherjee, S., Chen, W., et al. (2018). Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. Science *360*, 176–182.

Ross, E.M., and Markowetz, F. (2016). OncoNEM: Inferring tumor evolution from single-cell sequencing data. Genome Biol. *17*, 69.

Rozenblatt-Rosen, O., Stubbington, M.J.T., Regev, A., and Teichmann, S.A. (2017). The Human Cell Atlas: From vision to reality. Nature *550*, 451–453.

Salipante, S.J., and Horwitz, M.S. (2006). Phylogenetic fate mapping. Proc. Natl. Acad. Sci. USA *103*, 5448–5453.

Schepers, K., Swart, E., van Heijst, J.W., Gerlach, C., Castrucci, M., Sie, D., Heimerikx, M., Velds, A., Kerkhoven, R.M., Arens, R., and Schumacher, T.N. (2008). Dissecting T cell lineage relationships by cellular barcoding. J. Exp. Med. *205*, 2309–2318.

Setty, M., Tadmor, M.D., Reich-Zeliger, S., Angel, O., Salame, T.M., Kathail, P., Choi, K., Bendall, S., Friedman, N., and Pe'er, D. (2016). Wishbone identifies bifurcating developmental trajectories from single-cell data. Nat. Biotechnol. *34*, 637–645.

Shah, S., Lubeck, E., Zhou, W., and Cai, L. (2016). In situ transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. Neuron *92*, 342–357.

Shin, J., Berg, D.A., Zhu, Y., Shin, J.Y., Song, J., Bonaguidi, M.A., Enikolopov, G., Nauen, D.W., Christian, K.M., Ming, G.L., and Song, H. (2015). Single-cell RNA-seq with waterfall reveals molecular cascades underlying adult neurogenesis. Cell Stem Cell *17*, 360–372.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: Roles in mammalian development. Nat. Rev. Genet. *14*, 204–220.

Spanjaard, B., Hu, B., Mitic, N., Olivares-Chauvet, P., Janjuha, S., Ninov, N., and Junker, J.P. (2018). Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. Nat. Biotechnol. Published online April 9, 2018. https://doi.org/10.1038/nbt/4124.

Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2017). Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. bioRxiv. Published online April 19, 2017. https://doi.org/10.1101/128843.

Stubbington, M.J.T., Lönnberg, T., Proserpio, V., Clare, S., Speak, A.O., Dougan, G., and Teichmann, S.A. (2016). T cell fate and clonality inference from single-cell transcriptomes. Nat. Methods *13*, 329–332.

Sun, J., Ramos, A., Chapman, B., Johnnidis, J.B., Le, L., Ho, Y.J., Klein, A., Hofmann, O., and Camargo, F.D. (2014). Clonal dynamics of native haematopoiesis. Nature *514*, 322–327.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. Nat. Methods *6*, 377–382.

Tang, Q., Iyer, S., Lobbardi, R., Moore, J.C., Chen, H., Lareau, C., Hebert, C., Shaw, M.L., Neftel, C., Suva, M.L., et al. (2017). Dissecting hematopoietic and renal cell heterogeneity in adult zebrafish at single-cell resolution using RNA sequencing. J. Exp. Med. *214*, 2875–2887.

Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., 2nd, Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al. (2016a). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189–196.

Tirosh, I., Venteicher, A.S., Hebert, C., Escalante, L.E., Patel, A.P., Yizhak, K., Fisher, J.M., Rodman, C., Mount, C., Filbin, M.G., et al. (2016b). Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. Nature *539*, 309–313.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. *32*, 381–386.

Treutlein, B., Lee, Q.Y., Camp, J.G., Mall, M., Koh, W., Shariati, S.A., Sim, S., Neff, N.F., Skotheim, J.M., Wernig, M., and Quake, S.R. (2016). Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. Nature *534*, 391–395.

Tsang, J.C.H., Vong, J.S.L., Ji, L., Poon, L.C.Y., Jiang, P., Lui, K.O., Ni, Y.B., To, K.F., Cheng, Y.K.Y., Chiu, R.W.K., and Lo, Y.M.D. (2017). Integrative single-cell and cell-free plasma RNA transcriptomics elucidates placental cellular dynamics. Proc. Natl. Acad. Sci. USA *114*, 7786–7795.

Turner, D.L., and Cepko, C.L. (1987). A common progenitor for neurons and glia persists in rat retina late in development. Nature *328*, 131–136.

van Dijk, D., Nainys, J., Sharma, R., Kathail, P., Carr, A.J., Moon, K.R., Mazutis, L., Wolf, G., Krishnaswamy, S., and Pe'er, D. (2017). MAGIC: A diffusion-based imputation method reveals gene-gene interactions in single-cell RNA-sequencing data. bioRxiv. Published online February 25, 2017. https://doi.org/10.1101/111591.

van Heijst, J.W., Gerlach, C., Swart, E., Sie, D., Nunes-Alves, C., Kerkhoven, R.M., Arens, R., Correia-Neves, M., Schepers, K., and Schumacher, T.N. (2009). Recruitment of antigen-specific CD8+ T cells in response to infection is markedly efficient. Science *325*, 1265–1269.

Varshney, G.K., Pei, W., LaFave, M.C., Idol, J., Xu, L., Gallardo, V., Carrington, B., Bishop, K., Jones, M., Li, M., et al. (2015). High-throughput gene targeting and phenotyping in zebrafish using CRISPR/Cas9. Genome Res. *25*, 1030–1042.

Verovskaya, E., Broekhuis, M.J., Zwart, E., Ritsema, M., van Os, R., de Haan, G., and Bystrykh, L.V. (2013). Heterogeneity of young and aged murine hematopoietic stem cells revealed by quantitative clonal analysis using cellular barcoding. Blood *122*, 523–532.

Wang, Y., Waters, J., Leung, M.L., Unruh, A., Roh, W., Shi, X., Chen, K., Scheet, P., Vattathil, S., Liang, H., et al. (2014). Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature *512*, 155–160.

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. USA *115*, 2467–2476.

Weissman, T.A., and Pan, Y.A. (2015). Brainbow: New resources and emerging biological applications for multicolor genetic labeling and analysis. Genetics *199*, 293–306.

Wolf, F.A., Hamey, F., Plass, M., Solana, J., Dahlin, J.S., Gottgens, B., Rajewsky, N., Simon, L., and Theis, F.J. (2017). Graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. bioRxiv. Published online October 25, 2017. https://doi.org/10.1101/208819.

Woodworth, M.B., Girskis, K.M., and Walsh, C.A. (2017). Building a lineage from single cells: Genetic techniques for cell lineage tracking. Nat. Rev. Genet. *18*, 230–244.

Xu, X., Hou, Y., Yin, X., Bao, L., Tang, A., Song, L., Li, F., Tsang, S., Wu, K., Wu, H., et al. (2012). Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. Cell *148*, 886–895.

Yao, Z., Mich, J.K., Ku, S., Menon, V., Krostag, A.R., Martinez, R.A., Furchtgott, L., Mulholland, H., Bort, S., Fuqua, M.A., et al. (2017). A single-cell roadmap of lineage bifurcation in human esc models of embryonic brain development. Cell Stem Cell *20*, 120–134.

Yuan, G.C., Cai, L., Elowitz, M., Enver, T., Fan, G., Guo, G., Irizarry, R., Kharchenko, P., Kim, J., Orkin, S., et al. (2017). Challenges and emerging directions in single-cell analysis. Genome Biol. *18*, 84.

Zafar, H., Wang, Y., Nakhleh, L., Navin, N., and Chen, K. (2016). Monovar: Single-nucleotide variant detection in single cells. Nat. Methods *13*, 505–507.

Zeisel, A., Muñoz-Manchado, A.B., Codeluppi, S., Lönnerberg, P., La Manno, G., Juréus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., et al. (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science *347*, 1138–1142.

Zong, H., Espinosa, J.S., Su, H.H., Muzumdar, M.D., and Luo, L. (2005). Mosaic analysis with double markers in mice. Cell *121*, 479–492.