Developmental Cell

CellPress

Review

Next-Generation Lineage Tracing and Fate Mapping to Interrogate Development

Sadie VanHorn[1,2,3] and Samantha A. Morris[1,2,3,*]

[1]Department of Developmental Biology Washington, University School of Medicine in St. Louis, 660 S. Euclid Avenue, Campus Box 8103, St. Louis, MO 63110, USA

[2]Department of Genetics Washington, University School of Medicine in St. Louis, 660 S. Euclid Avenue, Campus Box 8103, St. Louis, MO 63110, USA

[3]Center of Regenerative Medicine, Washington University School of Medicine in St. Louis, 660 S. Euclid Avenue, Campus Box 8103, St. Louis, MO 63110, USA

*Correspondence: s.morris@wustl.edu

https://doi.org/10.1016/j.devcel.2020.10.021

SUMMARY

Lineage tracing and fate mapping, overlapping yet distinct disciplines to follow cells and their progeny, have evolved rapidly over the last century. Lineage tracing aims to identify all progeny arising from an individual cell, placing them within a lineage hierarchy. The recent emergence of genomic technologies, such as single-cell and spatial transcriptomics, has fostered sophisticated new methods to reconstruct lineage relationships at high resolution. In contrast, fate maps, schematics showing which parts of the embryo will develop into which tissue, have remained relatively static since the 1970s. However, fate maps provide spatial information, often lost in lineage reconstruction, that can offer fundamental mechanistic insight into development. Here, we broadly review the origins of fate mapping and lineage tracing approaches. We focus on the most recent developments in lineage tracing, permitted by advances in single-cell genomics. Finally, we explore the current potential to leverage these new technologies to synthesize high-resolution fate maps and discuss their potential for interrogating development at new depths.

Fate mapping and lineage tracing are closely related, yet distinct tools that form a central pillar of developmental biology. Fate maps are schematics where eventual cell fate is projected onto an embryo at a specific stage of development, depicting which cell or region gives rise to a particular tissue. Lineage tracing, stemming from century-old fate mapping experiments, aims to construct a hierarchy of all progeny arising from an individual cell but does not necessarily capture positional information. Early fate mapping efforts visually tracked individual cells or regions within an embryo, linking initial position to future fate (Conklin, 1905; Vogt, 1929). Such approaches quickly proved to be powerful tools in developmental biology, providing a picture of how initial position within an embryo can influence final fate. This rich spatial information offered fundamental mechanistic insight, such as whether the development of a specific organism is mosaic, i.e., highly dependent on lineage, or regulative, i.e., flexible and able to adapt to developmental perturbation (Lawrence and Levine, 2006). Although, in some instances, individual cells could be tracked, fate mapping was commonly deployed at the tissue level, offering broad, sometimes sweeping overviews of developmental patterning. While stemming from fate mapping, lineage tracing is distinct in its aim to reconstruct lineage hierarchy back to an individual cell (Figure 1). Indeed, in this respect, lineage provides valuable information on cell potential, where branching hierarchies can pinpoint the timing of critical cell fate decisions during development.

Lineage tracing strategies are rapidly advancing due to progress in genomics, enabling the construction of increasingly complex lineage hierarchies across diverse developmental systems. However, these emerging technologies typically do not capture spatial information, a crucial component of the fate map. Indeed, fate maps have generally remained static over the past few decades, yet they offer tremendous insight into the relationships between cells across time and position within the embryo. Without this spatial understanding, we lose knowledge of tissue borders, identities of adjacent cells, and intercellular communication—all essential facets in our understanding of cell fate specification. However, at present, fate maps typically only capture cell position and identity as determined from low-dimensional measurements, limiting the resolution of these approaches. Here, we broadly review the origins of fate mapping and lineage tracing approaches. We focus on recent advances in single-cell genomics to overcome several fundamental limitations in the construction of lineage hierarchies. We also discuss nascent technologies to integrate spatial information, highlighting the discoveries enabled by these methods. Finally, we explore the concept of next-generation, dynamic fate maps and the technical advances required for their construction.

FOUNDATIONS OF FATE MAPPING AND LINEAGE TRACING: DIRECT OBSERVATION AND CELL LABELING

Fate mapping relies on tracking cells within a developing organism, in a non-destructive manner. In the late 1800s, advances in light microscopy and dye injection techniques enabled such cell tracking via direct observation (Figure 2A) (Kretzschmar and

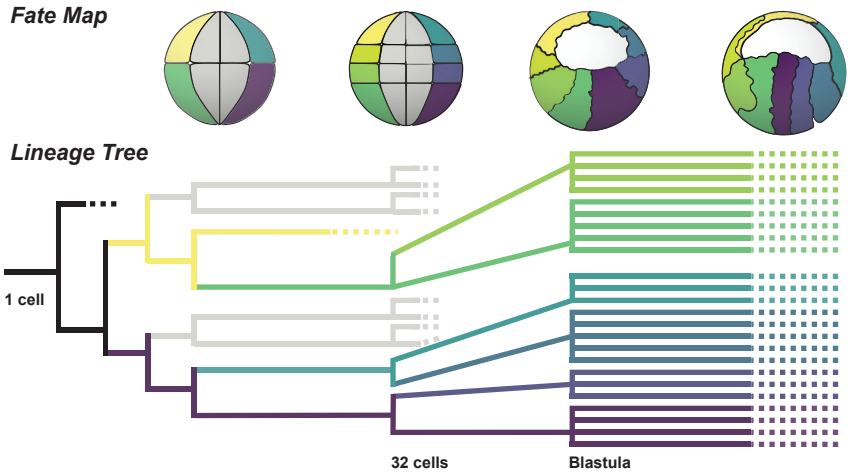

Figure 1. The Distinction between Fate Mapping and Lineage Tracing
Fate maps are schematics representing the developmental potential of specific cells or regions of cells within an embryo at a defined stage. Upper panels: early *Xenopus* development fate map, retaining cell-cell relationships and position. Lower panels: lineage tracing, the identification of all progeny arising from an individual cell. Each color in the branching tree relates to a region of the corresponding fate map. Dashed lines depict lineages not shown for simplicity.

Watt, 2012; Stent, 1998; Stern and Fraser, 2001). Notably, Whitman used direct observation to trace early cleavage and fate in leech development, demonstrating that stereotypical, invariant cell divisions determine eventual cell fate at early cleavage-stages (Whitman, 1878, 1887). Subsequently, in 1905, Conklin generated the first comprehensive fate map by tracking progenitor cells of the tunicate, *Styela partita*. The cell tracing in these experiments was powered by simple observation of cell color changes during differentiation, enabling fate map assembly (Conklin, 1905). With further developments in time-lapse cinematography, cells and their progeny within developing embryos could be traced (Wetzel, 1929). Perhaps the most famous example of this approach is represented by the microscopy-based direct observation of live animals to construct a complete lineage of *C. elegans* development (Sulston et al., 1983). Indeed, this example serves to demonstrate how lineage tracing and fate mapping overlap with the incorporation of spatial information into cell tracking (Figure 1).

The embryos used in these early studies were typically transparent, contained limited cell numbers, and had invariant cell lineages in most cases, supporting the straightforward observation of individual cells and their progeny. Fate mapping and lineage tracing in developmental systems with opaque embryos required cell labeling strategies, such as dyeing or radiolabeling (Kretzschmar and Watt, 2012; Stern and Fraser, 2001). Vogt was the first to develop and apply these techniques in the 1920s, fate mapping a variety of embryos, concluding that a stereotypic lineage does not define the development of some species (Vogt, 1929). Altogether, these techniques were invaluable in the construction of fate maps of many developing organisms, from the stereotyped development of leeches to the variable development of mice (Lawrence and Levine, 2006). For example, the intracellular injection of tracer dye allowed zebrafish fate map construction, which was instructive for understanding that clonal restrictions occur at the onset of gastrulation (Kimmel et al., 1990). However, these approaches still faced several fundamental limitations, such as the extended periods required to observe development *in vitro*, or the limited scalability of dye injection. The optics of light microscopy also restricted the types of embryos that were amenable to observation as they could not become too dense or pigmented during their development.

As described so far, the cell labeling used to build these early fate maps represents the first rudimentary lineage tracing. Indeed, although fate mapping significantly overlaps with lineage tracing, they cannot be considered equivalent since lineage refers to the hierarchy of cellular relationships and not spatial position. The systems in which a complete lineage tree can be captured simultaneously with a fate map, such as the *C. elegans* lineage, are outliers in this respect. The reality is that although lineage tracing has evolved significantly since the first fate maps emerged, our spatial understanding of lineage, the fate map, has been playing catch-up. Indeed, the black-and-white *Xenopus laevis* fate map remains unchanged from the original dye injection experiments in the 1970s (Nakamura and Kishiyama, 1971). Later in this review, we return to the current potential to construct a new generation of fate maps, based on emerging lineage tracing techniques and spatial transcriptomic technologies. Next, we review the evolution of cell and lineage tracing methods with no loss of signal or dye diffusion in progeny, and the ability to discriminate lineage on a single-cell basis, enabling the assembly of lineage hierarchies for increasingly complex tissues and organisms (Figure 2).

LEVERAGING NATURAL PHENOTYPIC VARIATION AS HERITABLE MARKS

At the turn of the 20th century, as embryology became more sophisticated, surgical and genetic manipulations shed light on cell lineage, though typically only to tissue-level resolution. For example, interspecies grafts exploited differences in graft coloration to support the direct observation of organizer potential before transgenic markers such as green fluorescent protein (GFP) were available (Figure 2B) (Spemann and Mangold, 1924; Wetzel, 1929). Genetic mosaicism studies also emerged around this time. *Drosophila simulans* gynandromorphs, genetic mosaics of male and female cells, yielded preliminary fate maps showing that different landmarks in the fly derive from separate cleavage nuclei (Sturtevant, 1929). Forty years later, these datasets were re-analyzed to generate comprehensive fate maps (Garcia-Bellido and Merriam, 1969). As an alternative to gynandromorphs, *Drosophila* mosaics could be generated by mitotic recombination (Stern, 1936). Analysis of the wing disc in these mosaics led to the discovery of compartments—the lineage segregation of proliferating cells into non-intermingling groups of cells, leading to functional subdivisions that shape the

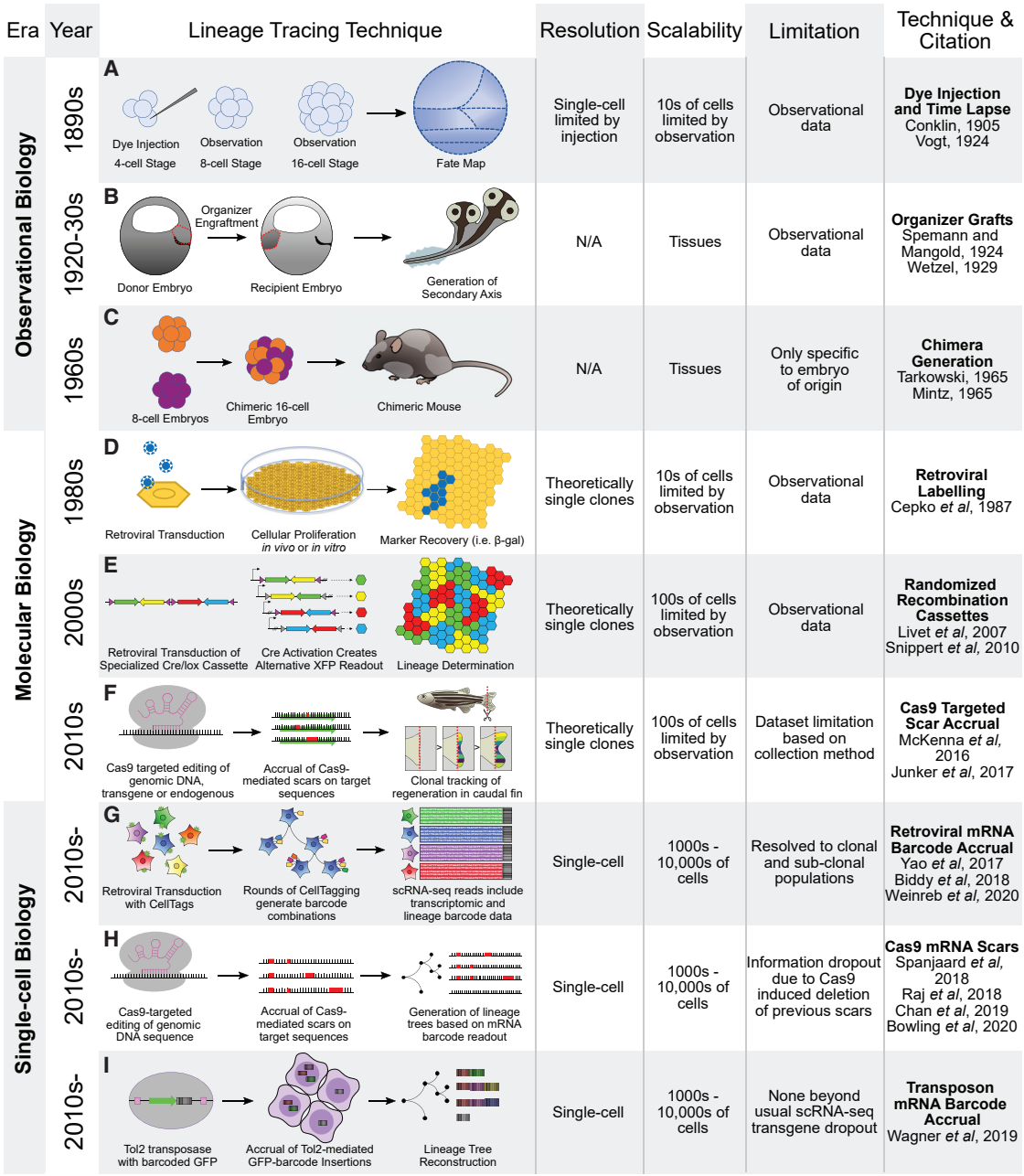

Figure 2. Evolution of Lineage Tracing Techniques

(A–I) Graphical depictions of (A) fate mapping via dye injection in the 1890s, (B) axis development and grafting, (C) generation of chimeric mouse embryos, (D) retroviral labeling of cells, (E) specialized Cre-*loxP* cassettes for clonal analysis, (F) Cas9 scar accrual in organisms, (G) viral barcoding approaches for clonal and lineage analysis, (H) Cas9 scar accrual compatible with scRNA-seq, and (I) transposon-mediated barcode accrual.

organism—a core concept in developmental biology (Garcia-Bellido et al., 1973).

In the 1960s, Tarkowski and Mintz leveraged the regulative nature of mouse development to generate the first experimental mosaic mammals (Figure 2C) (Mintz, 1965; Tarkowski, 1961). These chimeric embryos were created by aggregating cleavage-stage mouse embryos, enabling evaluation of the contributions of each original embryo to the adult mouse. These classic experiments used embryos with different genetically heritable characteristics for chimera generation, revealing the clonal origins of melanocytes (Tarkowski, 1964a), effects of mosaic genetic hermaphroditism (Tarkowski, 1964b), and developmental repercussions of known lethal mutations (Mintz, 1964). Beyond intraspecies embryo aggregations, Le Douarin pioneered

chimerism between different species (Le Douarin and Barq, 1969). Phenotypic differences between nuclei of the two species (chick and quail) enabled the distinction between transplanted cells and host cells, to track the migration, plasticity, and fate of neural crest cells (Le Douarin, 1980). Uniting all these approaches, in contrast to prior cell labeling experiments, these phenotypic ''labels'' were indelible, supporting long-term cell tracking, dispensing with the requirement for continuous observation. However, while these naturally occurring phenotypic differences helped establish a great deal of knowledge, their resolution was generally coarse, restricted to groups of cells and tissue-level observations. Furthermore, most of these techniques do not allow the direct monitoring of cells and their progeny, rendering them ineffective for lineage hierarchy construction.

LINEAGE TRACING USING GENETIC RECOMBINATION TOOLS

We have so far discussed techniques to identify lineage relationships based on cell morphology, prospectively marking cells with dyes, or by exploiting natural phenotypic variation between cells. However, lineage tracing using these approaches was restricted to specific organisms, inefficiently and non-specifically targeting cells. The advent of gene cloning in the 1970s (Chang and Cohen, 1974; Cohen et al., 1973) set the stage for transgenic lineage tracing. In 1987, Cepko used retroviral transduction to introduce β-galactosidase as a marker into the developing rat retina (Figure 2D) (Turner and Cepko, 1987). Careful titering of the virus to obtain as few marked cells as possible allowed the inference of lineage from individual cells, revealing that a single retinal progenitor can generate multiple cell types. These experiments can be defined as a clonal analysis since they identify the progeny derived from an individual cell, but the resolution is insufficient to construct a lineage hierarchy. Indeed, many of the early lineage tracing methods we discuss are able to identify all progeny arising from an individual cell, but they are not able to meet a stricter definition of lineage tracing, which requires that ancestor-progeny relationships are resolved to assemble a lineage tree.

Lineage tracing evolved quickly with the discovery and application of recombinase enzymes, Cre (Sauer and Henderson, 1988; Sternberg and Hamilton, 1981) and FLP (Golic and Lindquist, 1989). Cre and FLP induce recombination between their DNA target sites, *loxP*, and FRT, respectively, where target-site directionality can generate either inversions or deletions of the intervening DNA in a predictable manner (Nagy, 2000). For lineage tracing, Cre or FLP recombinase is constitutively expressed under the control of a tissue-specific promoter. Within these targeted cells, *loxP*/FRT sites are recombined, resulting in the excision of stop codons placed upstream of a ubiquitously expressed reporter, such as GFP. Consequently, the reporter is expressed and becomes a permanent, heritable mark of the targeted cell population. In flies, FLP-FRT-mediated recombination proved to be highly efficient and less toxic than Cre-*loxP* (Golic and Lindquist, 1989; Siegal and Hartl, 1996) and was first used to label clones in the *Drosophila* embryo and ovary (Harrison and Perrimon, 1993).

In mammals, the first Cre reporter mice incorporated the ubiquitous expression of β-galactosidase (Soriano, 1999) and EGFP (Mao et al., 2001). Recombinase-mediated lineage tracing advanced rapidly, where inducible recombination introduced spatial and temporal control, enabling the labeling of different cell subpopulations (Metzger et al., 1995). The identification of intestinal stem cells represents a benchmark example of the power of lineage tracing using recombination (Barker et al., 2007). Hypothesizing that the Wnt target gene *Lgr5* marked stem cells of the small intestine, Barker et al. (2007) developed a mouse line by knocking both CreERT2 and EGFP into the *Lgr5* locus. The fusion of Cre to the estrogen receptor (ER) allows drug-inducible control of recombination, via administration of tamoxifen in this case. These experiments demonstrated that *Lgr5*-expressing cells give rise to all intestinal lineages, supporting long-term maintenance of the intestinal epithelium.

Although genetic recombination-based lineage tracing was quickly adopted, several limitations emerged. For instance, tamoxifen administration can induce apoptosis, perturbing normal tissue homeostasis (Zhu et al., 2013). In terms of labeling fidelity, the expression of a single gene may not specifically mark the cell population of interest, although improved labeling specificity can be achieved by coupling recombination to the expression of multiple genes. For example, Split-Cre comprises two cleaved, inactive Cre fragments, each driven by a different promoter. Activation of both promoters within the same cell results in Cre enzyme reconstitution, driving recombination (Hirrlinger et al., 2009).

INCREASING LINEAGE TRACING RESOLUTION AND PRECISION

Another vital consideration for recombination-based lineage tracing is the frequency of cell labeling. Within a densely labeled population, it is not possible to distinguish whether two marked cells in close proximity are derived from a common ancestor or independent labeling events. While sparse labeling can provide theoretically clonal resolution, this approach may capture subpopulations that are not representative of the overall population behavior (Blanpain and Simons, 2013). ''Mosaic'' labeling captures a larger proportion of the population, while multicolor lineage tracing, such as ''Brainbow,'' allows identification of clones within the many labeled cells. Brainbow uses stochastic Cre-mediated recombination to induce combinatorial expression of fluorescent reporter transgenes, creating up to 90 discernable fluorescent signatures (Livet et al., 2007). The Brainbow system powers clonal analysis from mosaic labeling of cell populations as cells derived from a common ancestor can be identified based on their shared color. For example, the original Brainbow report enabled observation of neural circuits in mossy fiber neurons of the cerebellum and interdigitation of Bergmann glia within Purkinje cell glial sheaths (Livet et al., 2007). Confetti shares similarities with Brainbow but generates a single random fluorescent protein output, creating four discernable fluorescent signatures (Snippert et al., 2010) (Figure 2E).

Overall, these approaches bring recombinase-mediated lineage tracing increasingly closer to clonal resolution as the number of unique labels markedly expands. Reproducibility, pattern analysis, and statistics all support the assumption that

theoretically single clones can be traced using the Brainbow (Livet et al., 2007; Weissman and Pan, 2015) and Confetti reporters (Snippert et al., 2010). However, the restricted color palettes using these optical techniques limit the construction of comprehensive lineages and fate maps. Specifically, unrelated cells in close proximity may stochastically acquire the same color combination, representing a "label collision" that erroneously merges two unrelated lineages. In contrast with these restricted fluorophore combinations, genetic DNA "barcoding," called Polylox, uses random Cre-*LoxP*-mediated recombination events, creating up to ~1.9 million unique genetic barcodes to label cells (Pei et al., 2017). This technique identified hematopoietic stem cell (HSC) clones *in vivo*, where skewed barcode distribution suggested the independence of erythroid-myeloid differentiation from common myeloid progenitors and confirmed previous tree-like lineage structures of hematopoiesis (Pei et al., 2017). Relative to the 4, and ~90 possible labels afforded by Confetti and Brainbow, respectively, Polylox offers a much more statistically robust model for lineage tracing, decreasing the likelihood of label collision.

Although an increase in label diversity can improve lineage tracing resolution, there are many confounders of interpreting lineage results. For instance, the reporter expression used to label cells may be silenced in specific cell populations, as has been demonstrated with retrovirus-mediated cell labeling (Walsh and Cepko, 1992). This selective loss of labels from specific populations can mask genuine lineage relationships. Such silencing events can be overcome by the integration of the genetic labeling components into "safe-harbors" of the genome (Rinkevich et al., 2011; Snippert et al., 2010). In addition to the loss of lineage information from silencing, differential labeling of cell populations, inefficient cell and label capture, and loss of cells due to apoptosis can result in lineage termination, confounding interpretation of results. Next, we further explore the development and application of genomic technologies for lineage tracing, which provide unique power but require careful consideration of these confounding factors.

NEW GENOMIC TECHNOLOGIES TO ENABLE HIGH-RESOLUTION LINEAGE TRACING

The accurate assessment of cell identity is fundamental to fate mapping and lineage tracing. However, the lineage tracing techniques we have discussed so far use relatively few features, such as phenotype, position, or select marker expression, to identify cells. While this approach can be powerful in well-characterized systems, feature selection to assess cell identity is typically driven by prior biological knowledge, potentially biasing lineage reconstruction. At present, single-cell genomic technologies support a more objective assessment of cell identity, enabling the capture of many thousands of gene expression measurements while maintaining the cellular resolution required for accurate lineage reconstruction (Kester and van Oudenaarden, 2018). Since its relatively recent emergence in 2009 (Tang et al., 2009), single-cell RNA sequencing (scRNA-seq) has seen wide adoption and rapid development. Beyond gene expression measurements, multi-omic single-cell assays integrate additional techniques to quantify proteomes, genomes, and epigenomes (Stuart and Satija, 2019).

In concert with this technical progress, sophisticated computational tools have emerged to visualize and interpret the resulting complex datasets, leading to the identification of novel cell types and new mechanistic insights into biological processes. With particular relevance for developmental biologists, single-cell data have supported the inference of cell differentiation trajectories. A single "snapshot" of a developing cell population allows the ordering of individual cells according to gradual changes in their transcriptomes or chromatin accessibility profiles. This "pseudo-temporal ordering" yields plots that resemble branching lineages, although differentiation origins cannot be inferred from cells that are not captured or insufficiently sampled (Kester and van Oudenaarden, 2018). These approaches are valuable for identifying regulatory factors, such as transcription factors associated with specific branches of differentiation. However, it is crucial to acknowledge that such computational inference from single-cell data does not provide ground-truth lineage information, contrasting with the techniques discussed in previous sections. Indeed, to map a true lineage, cell ancestor-progeny relationships must be mapped, and in this respect, genomic technologies offer powerful tools to track and reconstruct lineage relationships. Next, we review the emergence of innovative techniques to encode heritable lineage information within the genome, captured in parallel with other cellular features, at high resolution, and with little limitation on scale.

Labeling and Tracking Cells Using Heritable Genetic Barcodes

Genome-based lineage tracing exploits sequence information to define lineage relationships between cells. Some techniques leverage naturally occurring somatic mutations that can be detected using scRNA-seq to retrospectively identify cells derived from a common ancestor (Leung et al., 2017; Lodato et al., 2015; Ludwig et al., 2019). Although this approach supports the investigation of systems that are not amenable to experimental manipulation, the relative infrequency of somatic mutation produces phylogenies of limited resolution. However, somatic mutation in mitochondrial DNA (mtDNA) has recently been shown to support clonal tracking at 1,000-fold the resolution of previous approaches and can allow the parallel capture of gene expression and chromatin accessibility (Ludwig et al., 2019). This increase in resolution is afforded by the higher somatic mutation rates of mtDNA, relative to nuclear DNA, and the high copy number of mitochondrial genomes in the cell. Nevertheless, this retrospective lineage tracing approach cannot support the reconstruction of complete lineage trees. More commonly, single-cell lineage tracing (scLT) is achieved by prospectively introducing a heritable genetic sequence into a cell, then determining clonal inheritance and constructing cell lineage retrospectively from sequencing data (Kester and van Oudenaarden, 2018).

The first scLT techniques are reminiscent of Cepko's original clonal analysis in the 1980s, utilizing retroviral transgene integration (Lu et al., 2011; Naik et al., 2013; Porter et al., 2014) or transposable elements (Rodriguez-Fraticelli et al., 2018; Sun et al., 2014) to incorporate DNA sequences into cells and to distinguish them from each other via sequencing. These discriminating DNA sequences are more commonly referred to as "barcodes" to

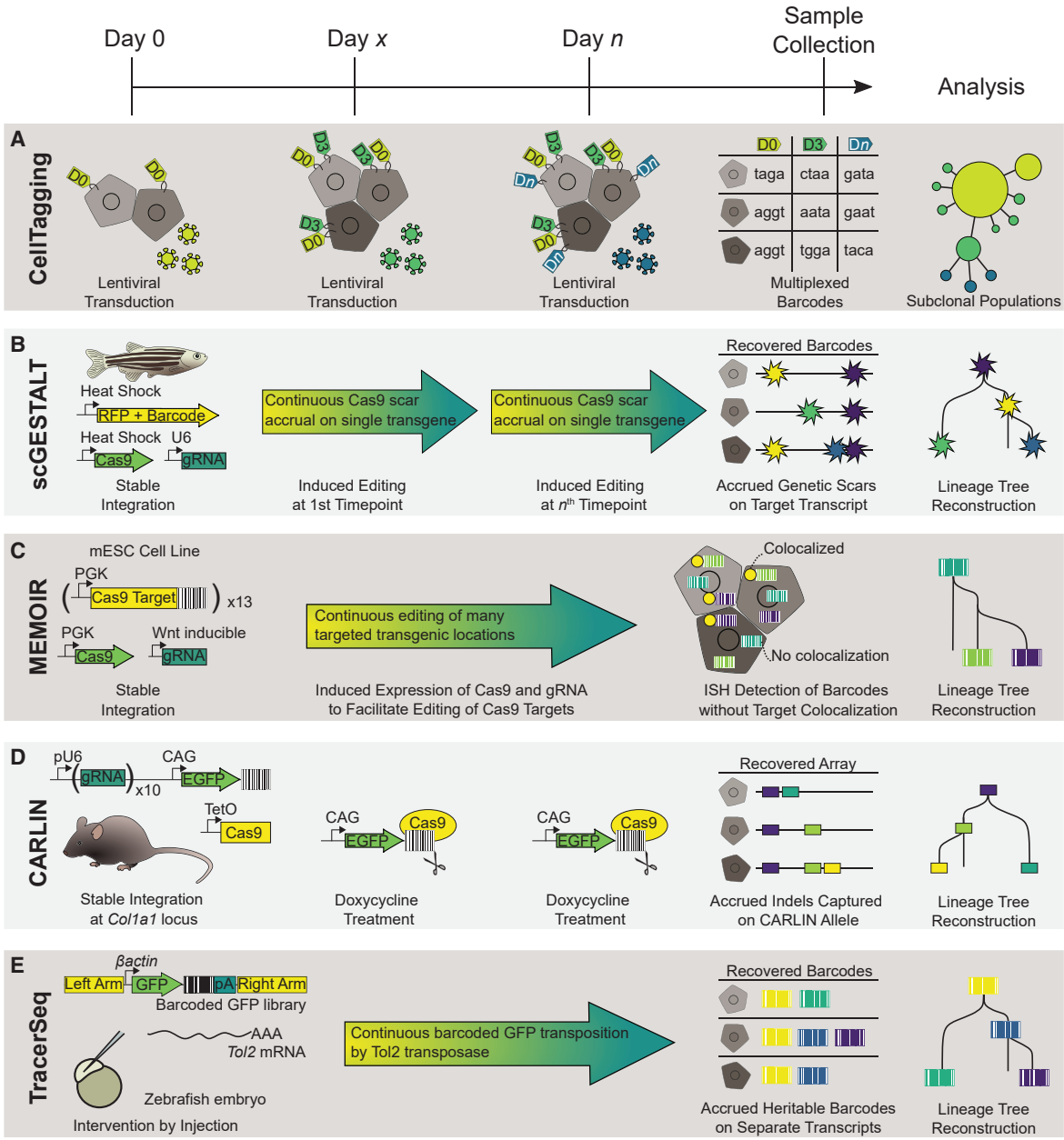

Figure 3. Inheritance of Barcodes for Single-Cell Lineage Tracing

(A) Accrual of heritable barcodes (CellTagging) for lineage construction.

(B) scGESTALT (genome editing of synthetic target arrays for lineage tracing), a Cas9-CRIPSR-based approach for continuous scar accrual on a single transgene.

(C) MEMOIR (memory by engineered mutagenesis with optical *in situ* readout), tracks CRISPR-Cas9-targeted scars that accrue over time in visually distinct cells. This technique employs barcoded "scratchpads," which are randomly edited over time via Cas9 targeting and read by a specialized set of smFISH probes. (D) CARLIN (CRISPR array repair lineage tracing) comprises a stably integrated Cas9 target allele, containing 10 locations for double-strand breaks and subsequent repair to occur. Editing can be induced, and barcodes created from DNA repair are captured via scRNA-seq for lineage reconstruction.

(E) TracerSeq: barcoded GFP reporter insertion into the genome, using the Tol2 transposase. Accrual of barcodes creates unique lineage identity labels for each cell.

denote the unique labeling of each ancestor and their descendants (Figure 3). Here, we use the terms "barcode" and "label" interchangeably. Some of the initial DNA barcoding scLT methods used an alternate approach by tracking CRISPR-Cas9-mediated genetic scarring on a stable transgene (Figure 2F) (Junker et al., 2016; McKenna et al., 2016). However, these early techniques typically required DNA-based barcode sequencing, which could not capture the transcriptome, forfeiting detailed information on cell identity and state.

To render cell barcoding approaches compatible with scRNA-seq, techniques evolved to express DNA barcodes as RNA transcripts (Biddy et al., 2018; Weinreb et al., 2020; Yao et al., 2017).

These methods commonly insert a barcode within the 3′UTR of a transgene, where a constitutive promoter drives its high expression, enabling reliable capture of the barcode in parallel with the transcriptome (Figure 2G). The barcodes, unique due to their randomized short sequences of nucleotides, are typically introduced via lentiviral transduction. Thus, these genetic identifiers are integrated into the genome of a cell and inherited by its progeny. It is important to note here that lineage tracing in this manner is strictly clonal analysis because it can identify populations arising from an individual cell, but a single round of cell labeling cannot yield a resolved lineage tree. To overcome this constraint, modified barcodes called "CellTags" were developed to label cells in successive rounds, providing a relative timescale to their accumulation, enabling lineage tree construction (Figure 3A) (Biddy et al., 2018; Kong et al., 2020).

These lentiviral barcoding approaches are particularly suited to accessible *in vitro* cell culture or regenerative systems where cells can be harvested, labeled, and transplanted. These systems are powerful for "clonal resampling," where a portion of the clone is sampled early, leaving the rest of the clone to differentiate (Wagner and Klein, 2020). This experimental design allows early cell state to be linked to eventual fate, providing clues into the early regulatory factors that set cells onto a defined lineage branch. For example, CellTagging revealed that distinct lineage conversion trajectories are determined at reprogramming initiation, and uncovered gene regulatory hallmarks of cells destined to reprogram (Biddy et al., 2018; Kamimoto et al., 2020). A similar cell barcoding strategy has mapped state-fate relationships in hematopoiesis, demonstrating that the gene expression of cell ancestors does not reliably predict the fate of their descendants. These findings suggest the existence of heritable properties guiding fate determination that scRNA-seq fails to capture (Weinreb et al., 2020). Additional information, such as chromatin accessibility, may uncover these hidden heritable properties, delivering further mechanistic insight into development and reprogramming.

Mutable Barcodes for the Reconstruction of Lineage Hierarchies

Elegant strategies have been devised based on the gradual mutation of barcodes that are "built-in" to the genome, enabling lineage tracing in whole organisms. Throughout a biological process, these barcodes slowly morph, allowing identification of related cells and phylogenetic tree construction. CRISPR-Cas9 has been fundamental in the development of these technologies, where the introduction of random mutations into genomic DNA uniquely labels cells. A typical feature of these approaches incorporates genetic barcodes within a multi-copy transgenic reporter, targeted by a sgRNA. The expression of Cas9 induces the cumulation of edits within these barcodes over time, allowing lineage reconstruction. In the first iteration of this approach, genomic DNA sequencing captured information within these barcodes (Figure 2F) (Junker et al., 2016; McKenna et al., 2016). However, by targeting edits to a synthetic construct that produces a transcript captured via scRNA-seq, cell identity and lineage can be measured in parallel (Figure 2H) (Alemany et al., 2018; Raj et al., 2018; Spanjaard et al., 2018). By targeting the same synthetic construct multiple times, genetic scar accrual places individual cells within a lineage hierarchy, akin to how phylogenetic relationships between species are defined based on shared and unique characteristics (Figure 3B).

To track cells within an intact, developing organism, the GESTALT (genome editing of synthetic target arrays for lineage tracing) zebrafish models have CRISPR-Cas9 editable cassettes engineered within their genomes (McKenna et al., 2016) and are compatible with scRNA-seq-based lineage tracing (Raj et al., 2018). The first experiments with the GESTALT zebrafish sampled labeled cells from different tissues of the adult. Lineage reconstruction revealed that most cells in the adult fish arise from few embryonic progenitors (McKenna et al., 2016). Further, when applied to study neural development in zebrafish, some neural progenitor populations encompassed a substantial spatial spread, suggesting progenitor populations were more migratory than previously thought. Furthermore, a high diversity of cell types within cell lineages indicated that progenitor populations maintain a higher degree of potency than expected (Raj et al., 2018). However, large deletions that erase lineage records limit the broader application of the GESTALT methods. Other approaches emerging at the same time, ScarTrace and LINNAEUS, mitigated these effects by distributing Cas9 targets across the genome, each integration harboring fewer targets (Alemany et al., 2018; Baron and van Oudenaarden, 2019; Junker et al., 2016; Spanjaard et al., 2018).

Indeed, Cas9 editing saturation limited the early CRISPR-Cas9 schemes, where identical edits may be introduced into independent cells, resulting in possible false-positive lineage relationships. Furthermore, editing periods were generally short, restricting their application to fast-developing organisms. These limitations presented a particular challenge for lineage tracing in mammals, where increased label diversity is required to label large cell populations over long periods. In response, self-targeting approaches use a "homing" guide RNA (hgRNA) to direct CRISPR-Cas9 to its own DNA locus, diversifying its sequence to act as a genetic barcode (Kalhor et al., 2017, 2018; Perli et al., 2016). However, this method does not support single-cell resolution analyses. An alternate approach combines CRISPR-Cas9 editing with PiggyBac transposase to integrate target-site cassettes across the genome. Using this system to study early mouse development revealed convergent differentiation from extraembryonic and embryonic endoderm to a common endoderm state (Chan et al., 2019). These results support previous observations that a small portion of the hindgut arises from extraembryonic progenitors (Kwon et al., 2008).

These initial mouse lineage tracing studies required the manipulation of embryos for each new experiment. Also, the high number of random transgene insertions precluded the establishment of breeding lines, limiting analysis of adult tissues. To overcome these restrictions, the CARLIN (CRISPR array repair lineage tracing) mouse line couples inducible Cas9 expression with a single, stably integrated mutable locus to record lineage data (Figure 3D). With this system, Bowling et al. report a biased distribution of HSC clones across long bones, suggesting that the niche into which HSCs home influences their expansion potential. Additionally, the authors demonstrated that after myeloablation, most blood cells are replenished by a small group of progenitors, shedding light on the clonal dynamics of hematopoietic regeneration (Bowling et al., 2020). Together, these

examples demonstrate the value of CRISPR-Cas9-based lineage tracing methods to provide new insights into mammalian development.

Lineage Reconstruction via Transposon-Mediated Barcode Accrual

The suite of CRISPR-Cas9 strategies discussed so far are limited by lineage recording dropout due to deletions or creation of the same edit from common repair. In contrast, barcode accrual can generate a greater diversity of unique heritable sequences through its combinatorial power. However, lentivirus-based methods are restricted by the number of times cells can be transduced, yielding rudimentary lineage trees. One method of barcode accrual overcomes these challenges to increase the labeling frequency: TracerSeq (Wagner et al., 2018). TracerSeq uses the Tol2 transposase to insert barcoded GFP reporters into the genome, which are transcribed and captured by scRNA-seq to obtain lineage data. Over the course of a biological process, barcode accrual creates unique label signatures for each cell, enabling lineage tree construction (Figure 2I). With this method, it is also possible to control transposition rates, enabling control of barcode insertion frequency (Figure 3E).

TracerSeq provided two important insights into early zebrafish development. First, within the 24-h post-fertilization embryo, cells from disparate embryonic fields can converge in their differentiation to produce transcriptionally similar cells. On the contrary, closely related cells can yield vastly divergent cell types. Indeed, Sulston's original fate mapping experiments in *C. elegans* concluded that, during development, similar neuron types could arise from discrete lineages (Sulston et al., 1983). This parallel supports the observations of convergent and divergent differentiation in zebrafish, further demonstrating our need to understand further how cellular diversity arises during development and differentiation as a whole.

Limitations of scLT Methods

The cell barcoding systems discussed so far face various technical challenges, most frequently associated with barcode capture inefficiency. Loss of lineage information can arise from barcode dropout, particularly when scRNA-seq is used to detect transcribed cell labels. This partial detection of barcodes is a particular issue when multiple, independent barcodes comprise a complete lineage label. Likewise, it is paramount to computationally correct errors arising from PCR amplification and sequencing to capture intact lineage information. Beyond these technical errors, barcodes are lost from the deletion of previously generated scars that held recorded lineage information, as is the case with the GESTALT-based methods (McKenna et al., 2016; Raj et al., 2018). It is also vital to consider silencing, which is an issue when relying on the capture of expressed barcodes. The detection of barcodes via genomic DNA can overcome this limitation, as is the case with ScarTrace (Alemany et al., 2018) and some image-based methods discussed later.

In addition to the failure to collect sufficient and accurate barcode information to reconstruct cell relationships, the labeling and capture rate of cells within a population is also an important consideration. For example, small yet biologically interesting clones may be lost due to inefficient cell capture. Such failures in cell capture can arise from cell death or inadequate cell dissociation. Furthermore, barcoding only a fraction of the cell population under study could severely limit accurate lineage reconstruction, confounding any biological interpretations made. Finally, barcode homoplasy, the labeling of unrelated cells with identical barcodes, results in the collapse of independent clones and lineages. These various error modes have been reviewed in detail (Wagner and Klein, 2020).

A recent computational stimulation of cell division and accumulation of CRISPR-Cas9-induced mutations has supported a quantitative exploration of the limitations discussed so far. This approach enabled lineage accuracy to be estimated under different scenarios, including mutation nature and complexity, mutation frequency, cell lineage depth, and target dropout (Salvador-Martínez et al., 2019). From these simulations, the authors make a series of recommendations in terms of target number and mutation rate and also suggest that uneven cell division rates require higher mutation frequencies to maintain lineage reconstruction accuracy. These simulations demonstrate how computational approaches may assist in the design of new experimental scLT strategies.

A new generation of scLT strategies would comprise editable genomic barcodes that slowly "evolve," to accrue at least one mutation per cell division, without the continued requirement for experimental intervention or problematic sequence deletions and insertions. The above simulations suggested that 0.05 to 0.25 mutations per cell division can yield accurate trees of around 65,000 cells (Salvador-Martínez et al., 2019). One potential new method to achieve this is called CHYRON (cell history recording by ordered insertion), which uses terminal deoxynucleotide transferase (TdT) to insert random nucleotides at a single locus (Loveless et al., 2019). Alternatively, fusion of dCas to alternative types of base editors could provide the unparalleled targeting of CRISPR-Cas systems, while offering more predictability. These Cas fusion proteins, lacking endonuclease activity, would also offer the opportunity to preserve the CRISPR-Cas target site, supporting continual evolution via base editing without losing targetability at any given site. Nucleic acid editors, such as the cytidine deaminase APOBEC1, could be candidates if utilized with a CRISPR-dCas9 system to guide base editing activity to predetermined regions (Koblan et al., 2018; Liu et al., 2019). Together, these approaches based on continuous barcode evolution have the potential to enable lineage tracing in more complex systems, over longer periods.

Computational Challenges for Lineage Reconstruction

We expect to see the development of many types of molecular techniques for lineage recording, but these approaches rely on the rigorous interpretation of the complex data they generate. One of the first computational approaches to reconstruct phylogenies from CRISPR-Cas9 scLT data utilized a Camin-Sokal maximum parsimony method. This approach is commonly used to create phylogenetic trees, resolving evolutionary relationships into the least complex tree to explain the different outcomes (Raj et al., 2018). The initial drawback of this approach was that it generally went unvalidated until other technologies and techniques were able to probe this kind of Cas9-induced scarring data further. Indeed, Spanjaard et al. (2018) note that in their system, Camin-Sokal maximum parsimony fails to reconstruct a correct lineage tree. In response, the authors developed

a custom algorithm for accurate tree reconstruction, fulfilling the maximum parsimony criterion, in the absence of complete scar detection. Lineage trees are reconstructed iteratively, with spurious connections, caused by cell doublets, removed. The method performed well on simulated data and enabled the inference of missing scar information, directly addressing some of the technical limitations discussed above.

An important aspect to note about these pipelines is that they are so nuanced that there is no standard training dataset or benchmarking approach to assess their accuracy. It is also difficult to directly compare the molecular techniques we have described here, so their comparative efficacy is generally unknown. Furthermore, it is incredibly challenging to assess the performance of these methods in the absence of known cell lineage. In most cases, these "ground truth" data are currently unknown. Aiming to address these challenges, Cassiopeia, a suite of scalable maximum parsimony approaches for phylogenetic tree reconstruction, incorporates a simulation framework to evaluate algorithms and experimental design, in addition to a ground truth dataset to benchmark different lineage reconstruction methods (Jones et al., 2020). Cassiopeia incorporates several algorithms to accurately infer phylogenies while being robust to experimental nuances such as mutation rate and number of target sites, as well as being scalable while resilient to data loss. Although Cassiopeia is tailored for current CRISPR-Cas9 scLT methods, the approach is also poised to adapt to evolving barcode strategies, as discussed above. Finally, new pipelines are emerging to reconstruct lineages beyond the window of active barcoding, or from a single round of cell barcoding (Weinreb and Klein, 2020).

Together, these experimental and computational developments in genomics bring exciting new possibilities for lineage tracing at high resolution. However, one fundamental limitation of these techniques remains: single-cell capture via the genomic methods detailed so far typically requires cell dissociation, abandoning spatial information, an essential component of fate mapping. For the remainder of this review, we focus on new techniques to retain spatial information during transcriptome capture and explore the integration of methods to construct high-resolution fate maps (Figure 4).

COUPLING CELL POSITION, IDENTITY, AND LINEAGE WITH SPATIAL TRANSCRIPTOMICS

To overcome the loss of spatial information resulting from cell dissociation, several computational methods to infer spatial position have emerged. The popular scRNA-seq software, Seurat, was initially developed to infer cell location within an embryo by integrating scRNA-seq data with *in situ* RNA expression information (Satija et al., 2015). In a similar approach, single-cell mRNA expression profiles have been matched to positional gene expression data to assign cells to specific positions within tissues (Achim et al., 2015). More recently, these computational methods have dispensed with the upfront requirement for spatial information of marker gene expression, taking scRNA-seq datasets and inferring spatial positioning based on known characteristics of the tissue of origin and transcriptomic similarity of cells (Nitzan et al., 2019). This method assumes that cells in close proximity are transcriptionally similar, which may not always be accurate, however. New technologies to capture spatial gene expression, termed "spatial transcriptomics," have matured rapidly in recent years (Moffitt et al., 2018), offering high-resolution templates with which to integrate scRNA-seq data.

Capture of Cell Location and Identity with Spatial Transcriptomics

Spatial transcriptomics techniques collect multimodal data by measuring both single-cell transcriptome and position. Coupling cell identity to location via these methods offers the missing link to integrate single-cell datasets with live imaging to create a comprehensive, high-dimensional single-cell atlas of the developing embryo, which we explore shortly. Following the advent of scRNA-seq, the drive to preserve or incorporate cellular location and phenotype into these datasets grew to further our understanding of how tissue structure is related to cellular heterogeneity. Next, we review progress in this nascent field and how live imaging is poised to create integrative fate maps to interrogate development.

Similar to scRNA-seq, spatial transcriptomics has recently seen tremendous gains in scale. Single-molecule fluorescence in situ hybridization (smFISH) was one of the first demonstrations of spatial transcriptomics with sufficient resolution to evaluate differences between cells based on transcript numbers (Figure 4A) (Femino et al., 1998). In 2015, Chen et al. introduced multiplexed error-robust fluorescence in situ hybridization (MERFISH) as a massively scaled-up application of smFISH. MERFISH deploys elegant multiplexing and encoding to overcome the limited number of fluorophores that can be imaged by previous smFISH approaches. Using a combinatorial labeling approach, MERFISH associates unique barcodes with individual transcripts, where these barcodes are built up one bit at a time and read via sequential rounds of hybridization and imaging (Chen et al., 2015) (Figure 4B). This approach overcomes the need to resolve a broad color palette and, combined with automation of the readout and sequencing process, exponentially increases the number of unique transcripts identified (Chen et al., 2015; Moffitt et al., 2018). SeqFISH (Lubeck et al., 2014) is a similar technique that was adapted using microfluidics to support multiple rounds of hybridization (Shah et al., 2016). Recently, this method has been developed further as SeqFISH+, where sequential hybridization effectively expands the barcode base palette to 60 "pseudocolors," increasing the scale and resolution of transcript detection (Figure 4C) (Eng et al., 2019). It is noteworthy that dropout of transcripts, i.e., expressed genes that are not detected, a limitation of scRNA-seq, is much less likely to occur in these smFISH-based protocols (Torre et al., 2018).

In situ sequencing offers another variety of spatial transcriptomics that does not rely on smFISH technology. For example, FISSEQ (fluorescent *in situ* RNA sequencing) couples RNA amplification with SOLiD (sequencing by oligonucleotide ligation and detection) sequencing within intact tissue to read transcript identity and position (Lee et al., 2014). Beyond FISSEQ, multiple other methods have evolved that use libraries of targeted probes, which allow improved coverage of the transcripts of interest (Chen et al., 2018; Iyer et al., 2018; Wang et al., 2018). While *in situ* sequencing offers subcellular specificity and the ability to sequence transcripts within tissue samples,

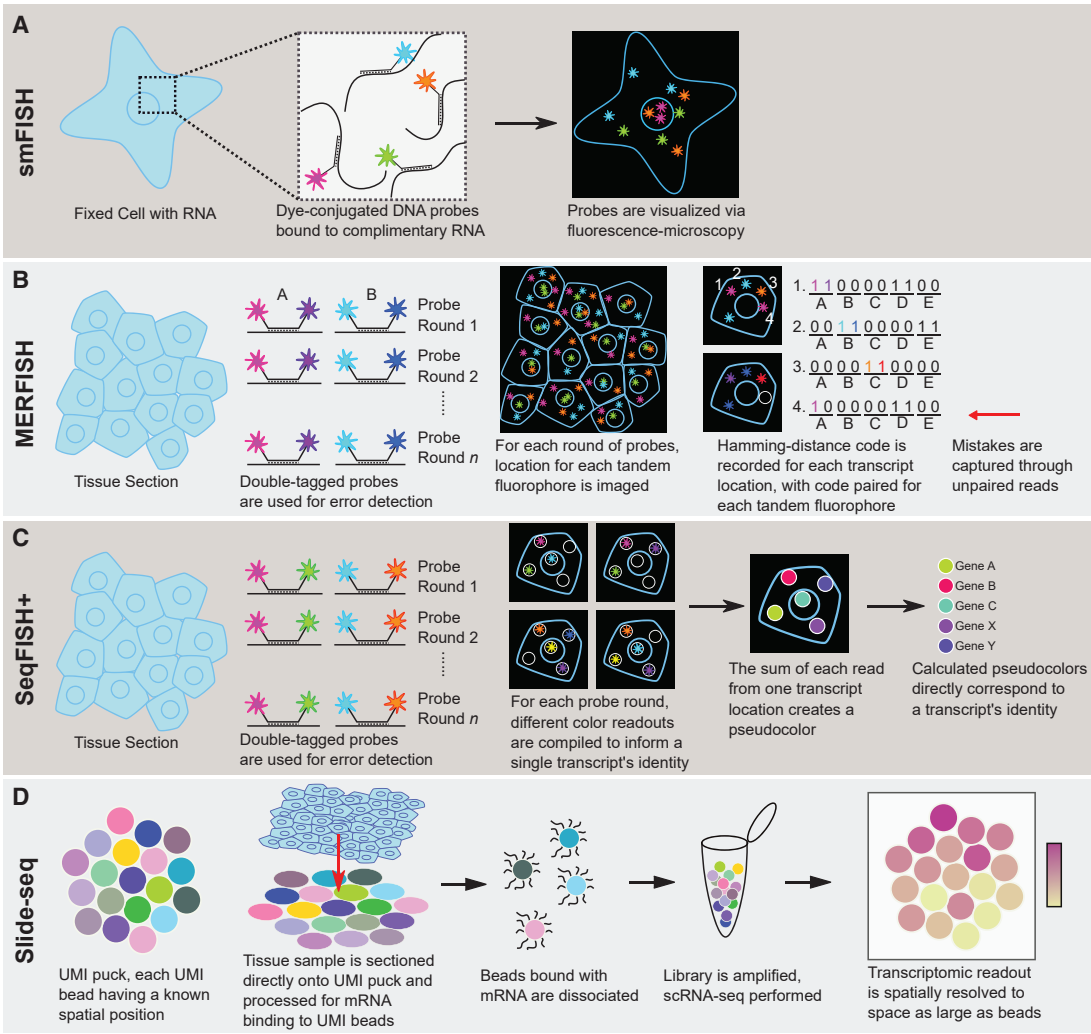

Figure 4. Spatial Transcriptomics Technologies

Visual comparison of the predominant techniques for measuring transcriptional expression in cells while maintaining spatial context.

(A) smFISH is used by hybridizing fluorophore-conjugated DNA probes to RNA within fixed cells and then imaging to capture the fluorescent readout.

(B) MERFISH uses a double-tagged probe for error detection with many rounds of hybridization. Through rounds of hybridization, the multiple reads that are generated are indicative of sequences present in the transcript and can identify them at sub-cellular resolution and is only limited by diffraction.

(C) SeqFISH+ is a technique conceptually similar to MERFISH but instead utilizes calculated "pseudocolors" that are assigned to each transcript that can be detected with the probe library.

(D) Slide-seq utilizes a UMI puck, directly adapted from UMI beads used for scRNA-seq, by arranging these beads into a single layer "puck" and directly slicing tissues onto the beads, transcripts can be hybridized to the UMI-containing oligos on the beads and sequenced similarly to scRNA-seq. Spatial maps of transcript expression can be generated based on the original UMI bead position.

probe-based approaches such as MERFISH and SeqFISH+ are perhaps more scalable, albeit generally constricted by targeted libraries of probes.

In an alternate approach, Slide-seq uses positionally mapped DNA barcodes (unique molecular identifiers, or "UMIs") to capture transcriptomes across a section of tissue, retaining spatial information at near single-cell resolution (Rodriques et al., 2019). Slide-seq takes the earlier UMI beads used for cell capture by microfluidics (Macosko et al., 2015) but attaches them to a coverslip to form a "puck" of barcoded beads whose sequence is associated with specific two-dimensional coordinates. Following the placement of a thin tissue section onto the puck, transcripts from the cells hybridize to the barcoded beads. Subsequent library preparation and sequencing enable the association of transcripts with specific barcodes, which in turn allows the deduction of their original location to within 20 μm (Figure 4D). The main drawback of Slide-seq is its limited resolution and low transcript detection sensitivity. For example, around one-third of sequenced beads capture mRNA from two different cell types, suggesting that single-cell resolution is not achieved. Additionally, less than 100 transcripts are captured per bead, although recent technical improvements to the reverse transcription step have increased this to over 500 transcripts per bead (Stickels et al., 2020). Another limitation is the tissue destruction that accompanies this approach, precluding the

collection of valuable phenotypic data via imaging, which is an essential component for constructing fate maps.

Genomic Techniques to Capture Cell Location and Lineage

While spatial transcriptomics serves to survey broad gene expression patterns within intact tissue, targeted genomic technologies exist to couple cell position with retrospective lineage tracing. Memory by engineered mutagenesis with optical *in situ* readout (MEMOIR) uses CRISPR-Cas9-based targeted mutagenesis to randomly edit barcoded recorder elements called "scratchpads" (Frieda et al., 2017). Read by a specialized set of smFISH probes, these targeted sequences accrue mutations over time to record lineage while retaining spatial information (Figure 3C). Advancing on MEMOIR, Zombie dispenses with reliance on barcode transcription by using T7 phage RNA polymerases to transcribe scratchpads after tissue fixation. This approach eliminates any detrimental effects that may occur due to the constitutive expression of barcodes within living cells and sidesteps the issue of silencing, as discussed earlier (Askary et al., 2019). In more recent work, IntMEMOIR leverages serine integrase-induced barcode editing, using smFISH to read edits in parallel with fluorescence reporter expression and a GAL4-UAS system to restrict labeling to specific cell types (Chow et al., 2020). Because IntMEMOIR incorporates a three-state memory element, tens of thousands of unique labels can be generated—an order of magnitude greater than MEMOIR. However, these methods will likely be underutilized until broader transcriptome analysis is incorporated. Indeed, if coupled with spatial transcriptomics, these integrated approaches can act as a connection to tether datasets together by cell lineage.

TOWARD AN INTEGRATIVE FATE MAP TO INTERROGATE DEVELOPMENT

The genomic technologies we have discussed so far enable the measurement of cell identity and lineage relationships, at single-cell resolution. Furthermore, the most recent of these methods preserve spatial and phenotypic information to yield high-dimensional insights into cell identity. Together, these approaches have supported significant advances in lineage tracing, offering new insight into embryonic development and beyond. Next, we address the strengths and limitations of these techniques for constructing fate maps—schematics representing the developmental potential of specific cells or regions of cells within an embryo at a defined stage, as outlined at the beginning of this review. For the construction of a fate map, cells must be traced throughout development, in a non-destructive manner. However, current fate mapping approaches yield relatively limited "low-dimensional" information on cell identity and state. In contrast, the range of genomic technologies we have explored in this review has the potential to generate high-dimensional fate maps, but these methods universally rely on cell destruction or fixation. An integrative approach will be essential to overcome these limitations, to generate high-resolution fate maps that are transcriptionally, spatially, and temporally defined. Below, we outline the critical components of this integrative approach.

Advances in Cell Imaging for High-Resolution, Non-invasive Cell Tracking

The ability to directly visualize and track cells has played a central role in fate mapping and lineage tracing since their conception. Indeed, time-lapse microscopy has supported the non-destructive, direct observation of embryonic development over time (Stern and Fraser, 2001). Light-sheet microscopy has proven fundamental to increasing our ability to probe deeper into tissue because it can image sections of tissue without inducing photodamage (Huisken et al., 2004; Pantazis and Supatto, 2014). The first light-sheet microscope allowed optical sectioning within living fish embryos, with greater tissue depth and resolution than confocal laser scanning microscopy, demonstrating the promise of this imaging technique (Huisken et al., 2004). Modifications to light-sheet microscopy have further advanced imaging depth capabilities, overcoming the limitations that initially hindered large or pigmented embryos from being studied. By utilizing these advanced techniques, zebrafish (Keller et al., 2008), fly (Krzic et al., 2012), and even mouse embryos (McDole et al., 2018) have been viewed *in toto*, throughout their morphogenesis. These studies have offered dynamic single-cell atlases of embryonic development, preserving cell lineage relationships, and will be central to any integrative fate map.

However, the utility of live imaging goes far beyond lineage and positioning, leading to a branch of study called phenomics. Phenomics encompasses any phenotypic observation made via imaging, such as cell migration patterns and speed, cell shape, and fluorescently labeled proteins (Grys et al., 2017; Nketia et al., 2017; Pantazis and Supatto, 2014). As a result, the dimensionality of the data captured by live imaging increases. Indeed, there is a wealth of phenotypic information that can be captured via non-destructive label-free imaging approaches aside from visually tracking cell lineage and position. Differential interference contrast (DIC), phase contrast, and dark-field microscopy are all able to capture distinct morphological characteristics that can be identified and analyzed via deep learning (Grys et al., 2017; Nketia et al., 2017). Characteristics such as cell size, density, and more are collected, increasing the dimensionality of the dataset. Ultimately, this has supported machine learning approaches to predict protein expression and localization in cells based on light microscopy images alone (Christiansen et al., 2018).

Data Integration to Synthesize High-Dimensional, Dynamic Fate Maps

Despite being able to add dimensionality to fate maps acquired by the above label-free imaging approaches, it would remain challenging to reach a similar richness of dimensionality compared with genomics-based lineage tracing approaches. These latter tools, such as scRNA-seq and spatial transcriptomics, typically capture the expression of thousands of transcripts within individual cells. This increase in dimensionality supports much more nuanced descriptions of cell identity and state, and the molecular mechanisms underpinning their regulation. By coupling live imaging to single-cell transcriptomics, positional and temporal data can be added to the multidimensional transcriptomic readouts. Since cellular location and developmental context impact

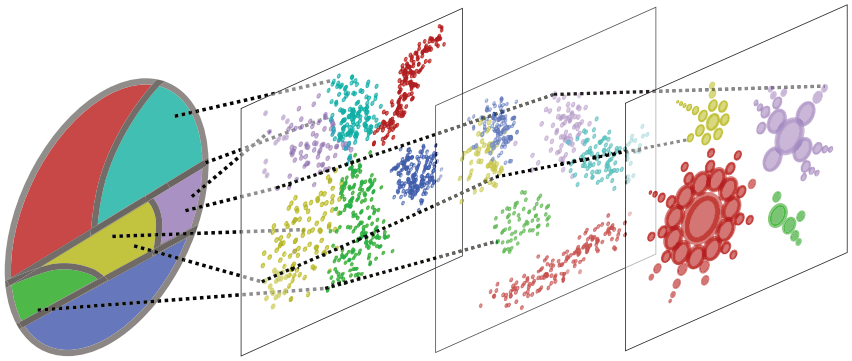

Figure 5. Data Integration to Construct High-Dimensional, Dynamic Fate Maps

Integration of multimodal and multi-omic data to create high-dimensional fate maps, where cell identity, state, lineage, and behavior can be probed at any stage. This highly quantitative dataset would enable *in silico* experimentation, allowing the impact of molecular and cellular perturbations on development to be modeled.

Fate Mapping	***Spatial Transcriptomics***	***Multiomics***	***Lineage Tracing***
Live Imaging	smFISH	scRNA-seq	CellTagging
Deep Learning	*in situ* Sequencing	CITE-seq	scartrace
Light Sheet Microscopy	MERFISH	scATAC-seq	LINNAEUS
	SeqFISH+	sci-MET	scGESTALT
	Slide-seq		TracerSeq
			CARLIN

transcriptional regulation, a greater understanding of these characteristics' interdependence can be gleaned. Furthermore, scLT data can be evaluated in tandem with live imaging-based lineage tracing, allowing rigorous interrogation of cell lineages. However, it is not currently possible to experimentally combine these high-resolution genomic strategies with non-destructive live imaging; thus, computational integration approaches will fulfill an essential role in this respect.

Beyond RNA capture, a suite of genomic technologies to measure the transcriptome, epigenome, and proteome now exist (Stuart and Satija, 2019). While experimental methods to capture these different properties from the same individual cell are rapidly developing, computational approaches to integrate different data modalities, "multi-omics," captured across independent experiments are emerging (Stuart et al., 2019). These tools are similar to the above approaches to infer cell position from scRNA-seq, based on known *in situ* gene expression patterns. A critical aspect of these latest tools is the compilation of "anchors," a core set of transcripts measured across different datasets in order to merge them (Stuart et al., 2019). Indeed, many methods exist to combine scRNA-seq data with different modalities, though they typically use scRNA-seq data as a reference to tie together the multimodal information (Efremova and Teichmann, 2020).

The construction of a high-dimensional, dynamic fate map will require the computational integration of several distinct, complementary datasets (Figure 5). The first component is live imaging for high-resolution capture of cell position, phenotype, and lineage relationships across a specified developmental period. Second, scRNA-seq at defined stages throughout this developmental period should also be collected. The third crucial component to bridge these datasets, enabling their integration, is spatial transcriptomics. By gathering information on the expression of key anchor genes, represented by 100–1,000s of transcripts, preserving cell position and phenotype at defined stages of development, live imaging and transcriptomics can be merged. Thus, for each cell during its development, high-dimensional information on gene expression can be inferred. Furthermore, this approach has the potential for the integration of other modalities. The result will be a high-dimensional fate map, where cell identity, state, lineage, and behavior can be probed at any stage. Together, this would facilitate improved modeling of embryonic development, creating exciting opportunities to interrogate molecular perturbations across multiple levels of regulation, enabling enhanced *in silico* experimentation. Additional modeling at the cellular level would also support sophisticated probabilistic modeling of development. Indeed, such *in silico* strategies would constitute a significant force to guide scientific inquiry because they provide functional and mechanistic hypotheses for further experimental investigation and validation.

CONCLUDING REMARKS

Since the first fate map was recorded at the beginning of the 20th century, the lineage tracing toolbox has evolved from simple observation to recorded live imaging, single-cell resolution lineage tracing, and retrospective computational analysis. Indeed, the technologies required to build high-dimensional fate maps now exist: a visual dynamic fate map for the post-implantation mouse embryo has been captured (McDole et al., 2018); technologies to capture multi-omics data in parallel with single-cell transcriptomics have been developed (Stuart and Satija, 2019); single cells from transcriptomic datasets can be resolved to spatial coordinates (Moffitt et al., 2018; Nitzan et al., 2019); finally, diverse datasets can be computationally integrated based on common characteristics (Efremova and Teichmann, 2020; Stuart et al., 2019). Thus, this next generation of fate mapping is poised to help us more deeply comprehend how cell identity is regulated within the context of a developing organism. Moreover, these high-resolution, dynamic models will aid in building a quantitative understanding of the rules governing embryonic development.

ACKNOWLEDGMENTS

We thank members of the Morris lab and our reviewers for helpful feedback. The Morris Lab is funded by National Institute of General Medical Sciences R01 GM126112, Silicon Valley Community Foundation, Chan Zuckerberg Initiative grant CZF2019-002423, and Children's Discovery Initiative grant CDI-SP-2019-847 to S.A.M.; S.A.M. is supported by an Allen Distinguished Investigator Award, a Paul G. Allen Frontiers Group advised grant of the Paul G. Allen Family Foundation, a Vallee Scholar Award, a Sloan Research

Fellowship, and a New York Stem Cell Foundation Robertson Investigator Award; and S.V. is supported by NIH grant 5T32GM007067-46.

REFERENCES

Achim, K., Pettit, J.B., Saraiva, L.R., Gavriouchkina, D., Larsson, T., Arendt, D., and Marioni, J.C. (2015). High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. Nat. Biotechnol. *33*, 503–509.

Alemany, A., Florescu, M., Baron, C.S., Peterson-Maduro, J., and Oudenaarden, A. van. (2018). Whole-organism clone tracing using single-cell sequencing. Nature *556*, 108–112.

Askary, A., Sanchez-Guardado, L., Linton, J.M., Chadly, D.M., Budde, M.W., Cai, L., Lois, C., and Elowitz, M.B. (2019). In situ readout of DNA barcodes and single base edits facilitated by in vitro transcription. Nat. Biotechnol. *38*, 1–10.

Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J., et al. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature *449*, 1003–1007.

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.

Biddy, B.A., Kong, W., Kamimoto, K., Guo, C., Waye, S.E., Sun, T., and Morris, S.A. (2018). Single-cell mapping of lineage and identity in direct reprogramming. Nature *564*, 219–224.

Blanpain, C., and Simons, B.D. (2013). Unravelling stem cell dynamics by lineage tracing. Nat. Rev. Mol. Cell Biol. *14*, 489–502.

Bowling, S., Sritharan, D., Osorio, F.G., Nguyen, M., Cheung, P., Rodriguez-Fraticelli, A., Patel, S., Yuan, W.-C., Fujiwara, Y., Li, B.E., et al. (2020). An engineered CRISPR-Cas9 mouse line for simultaneous readout of lineage histories and gene expression profiles in single cells. Cell *181*, 1410–1422.e27.

Chan, M.M., Smith, Z.D., Grosswendt, S., Kretzmer, H., Norman, T.M., Adamson, B., Jost, M., Quinn, J.J., Yang, D., Jones, M.G., et al. (2019). Molecular recording of mammalian embryogenesis. Nature *570*, 77–82.

Chang, A.C.Y., and Cohen, S.N. (1974). Genome construction between bacterial species in vitro: replication and expression of Staphylococcus plasmid genes in Escherichia coli. Proc. Natl. Acad. Sci. USA *71*, 1030–1034.

Chen, K.H., Boettiger, A.N., Moffitt, J.R., Wang, S., and Zhuang, X. (2015). RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. Science *348*, aaa6090.

Chen, X., Sun, Y.C., Church, G.M., Lee, J.H., and Zador, A.M. (2018). Efficient in situ barcode sequencing using padlock probe-based BaristaSeq. Nucleic Acids Res *46*, e22.

Chow, K.-H.K., Budde, M.W., Granados, A.A., Cabrera, M., Yoon, S., Cho, S., Huang, T., Koulena, N., Frieda, K.L., Cai, L., et al. (2020). Imaging cell lineage with a synthetic digital recording system. bioRxiv. Published online April 25, 2020. https://doi.org/10.1101/2020.02.21.958678.

Christiansen, E.M., Yang, S.J., Ando, D.M., Javaherian, A., Skibinski, G., Lipnick, S., Mount, E., O'Neil, A., Shah, K., Lee, A.K., et al. (2018). In silico labeling: predicting fluorescent labels in unlabeled images. Cell *173*, 792–803.e19.

Cohen, S.N., Chang, A.C.Y., Boyer, H.W., and Helling, R.B. (1973). Construction of biologically functional bacterial plasmids in vitro. Proc. Natl. Acad. Sci. USA *70*, 3240–3244.

Conklin, E.G. (1905). The organization and cell-lineage of the ascidian egg. Science *23*, 40–344.

Efremova, M., and Teichmann, S.A. (2020). Computational methods for single-cell omics across modalities. Nat. Methods *17*, 14–17.

Eng, C.-H.L., Lawson, M., Zhu, Q., Dries, R., Koulena, N., Takei, Y., Yun, J., Cronin, C., Karp, C., Yuan, G.C., et al. (2019). Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. Nature *568*, 235–239.

Femino, A.M., Fay, F.S., Fogarty, K., and Singer, R.H. (1998). Visualization of single RNA transcripts in situ. Science *280*, 585–590.

Frieda, K.L., Linton, J.M., Hormoz, S., Choi, J., Chow, K.-H.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.

Garcia-Bellido, A., and Merriam, J.R. (1969). Cell lineage of the imaginal discs in Drosophila gynandromorphs. J. Exp. Zool. *170*, 61–75.

Garcia-Bellido, A., Ripoll, P., and Morata, G. (1973). Developmental compartmentalisation of the wing disk of Drosophila. Nat. New Biol. *245*, 251–253.

Golic, K.G., and Lindquist, S. (1989). The FLP recombinase of yeast catalyzes site-specific recombination in the Drosophila genome. Cell *59*, 499–509.

Grys, B.T., Lo, D.S., Sahin, N., Kraus, O.Z., Morris, Q., Boone, C., and Andrews, B.J. (2017). Machine learning and computer vision approaches for phenotypic profiling. J. Cell Biol. *216*, 65–71.

Harrison, D.A., and Perrimon, N. (1993). Simple and efficient generation of marked clones in Drosophila. Curr. Biol. *3*, 424–433.

Hirrlinger, J., Scheller, A., Hirrlinger, P.G., Kellert, B., Tang, W., Wehr, M.C., Goebbels, S., Reichenbach, A., Sprengel, R., Rossner, M.J., et al. (2009). Split-Cre complementation indicates coincident activity of different genes in vivo. PLoS One *4*, e4286.

Huisken, J., Swoger, J., Del Bene, F.D., Wittbrodt, J., and Stelzer, E.H.K. (2004). Optical sectioning deep inside live embryos by selective plane illumination microscopy. Science *305*, 1007–1009.

Iyer, E.P.R., Punthambaker, S., Liu, S., Jindal, K., Farrell, M., Murn, J., Ferrante, T., Rudnicki, S., Kohman, R.E., Wassie, A.T., et al. (2018). Barcoded oligonucleotides ligated on RNA amplified for multiplex and parallel in-situ analyses. bioRxiv. Published online March 20, 2018. https://doi.org/10.1101/281121.

Jones, M.G., Khodaverdian, A., Quinn, J.J., Chan, M.M., Hussmann, J.A., Wang, R., Xu, C., Weissman, J.S., and Yosef, N. (2020). Inference of single-cell phylogenies from lineage tracing data. Genome Biol *21*, 92.

Junker, J.P., Spanjaard, B., Peterson-Maduro, J., Alemany, A., Hu, B., Florescu, M., and van Oudenaarden, A. (2016). Massively parallel clonal analysis using CRISPR/Cas9 induced genetic scars. bioRxiv. 056499. https://doi.org/10.1101/056499.

Kalhor, R., Kalhor, K., Mejia, L., Leeper, K., Graveline, A., Mali, P., and Church, G.M. (2018). Developmental barcoding of whole mouse via homing CRISPR. Science *361*, eaat9804.

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

Kamimoto, K., Hoffmann, C.M., and Morris, S.A. (2020). CellOracle: dissecting cell identity via network inference and in silico gene perturbation. bioRxiv. Published online April 21, 2020. https://doi.org/10.1101/2020.02.17.947416.

Keller, P.J., Schmidt, A.D., Wittbrodt, J., and Stelzer, E.H.K. (2008). Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. Science *322*, 1065–1069.

Kester, L., and van Oudenaarden, A. (2018). Single-cell transcriptomics meets lineage tracing. Cell Stem Cell *23*, 166–179.

Kimmel, C.B., Warga, R.M., and Schilling, T.F. (1990). Origin and organization of the zebrafish fate map. Development *108*, 581–594.

Koblan, L.W., Doman, J.L., Wilson, C., Levy, J.M., Tay, T., Newby, G.A., Maianti, J.P., Raguram, A., and Liu, D.R. (2018). Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nat. Biotechnol. *36*, 843–846.

Kong, W., Biddy, B.A., Kamimoto, K., Amrute, J.M., Butka, E.G., and Morris, S.A. (2020). CellTagging: combinatorial indexing to simultaneously map lineage and identity at single-cell resolution. Nat. Protoc. *15*, 750–772.

Kretzschmar, K., and Watt, F.M. (2012). Lineage tracing. Cell *148*, 33–45.

Krzic, U., Gunther, S., Saunders, T.E., Streichan, S.J., and Hufnagel, L. (2012). Multiview light-sheet microscope for rapid in toto imaging. Nat. Methods *9*, 730–733.

Kwon, G.S., Viotti, M., and Hadjantonakis, A.K. (2008). The endoderm of the mouse embryo arises by dynamic widespread intercalation of embryonic and extraembryonic lineages. Dev. Cell *15*, 509–520.

Lawrence, P.A., and Levine, M. (2006). Mosaic and regulative development: two faces of one coin. Curr. Biol. *16*, R236–R239.

Le Douarin, N., and Barq, G. (1969). Use of Japanese quail cells as "biological markers" in experimental embryology. C. R. Acad. Hebd. Seances. Acad. Sci. D. *269*, 1543–1546.

Le Douarin, N.M. (1980). The ontogeny of the neural crest in avian embryo chimaeras. Nature *286*, 663–669.

Lee, J.H., Daugharthy, E.R., Scheiman, J., Kalhor, R., Yang, J.L., Ferrante, T.C., Terry, R., Jeanty, S.S.F., Li, C., Amamoto, R., et al. (2014). Highly multiplexed subcellular RNA sequencing in situ. Science *343*, 1360–1363.

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.

Liu, Z., Shan, H., Chen, S., Chen, M., Zhang, Q., Lai, L., and Li, Z. (2019). Improved base editor for efficient editing in GC contexts in rabbits with an optimized AID-Cas9 fusion. FASEB J *33*, 9210–9219.

Livet, J., Weissman, T.A., Kang, H., Draft, R.W., Lu, J., Bennis, R.A., Sanes, J.R., and Lichtman, J.W. (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature *450*, 56–62.

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.

Loveless, T.B., Grotts, J.H., Schechter, M.W., Forouzmand, E., Carlson, C.K., Agahi, B.S., Liang, G., Ficht, M., Liu, B., Xie, X., et al. (2019). DNA writing at a single genomic site enables lineage tracing and analog recording in mammalian cells. bioRxiv. Published online September 29, 2019. https://doi.org/10.1101/639120.

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.

Ludwig, L.S., Lareau, C.A., Ulirsch, J.C., Christian, E., Muus, C., Li, L.H., Pelka, K., Ge, W., Oren, Y., Brack, A., et al. (2019). Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. Cell *176*, 1325–1339.e22.

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

Mao, X., Fujiwara, Y., Chapdelaine, A., Yang, H., and Orkin, S.H. (2001). Activation of EGFP expression by Cre-mediated excision in a new ROSA26 reporter mouse strain. Blood *97*, 324–326.

McDole, K., Guignard, L., Amat, F., Berger, A., Malandain, G., Royer, L.A., Turaga, S.C., Branson, K., and Keller, P.J. (2018). In toto imaging and reconstruction of post-implantation mouse development at the single-cell level. Cell *175*, 859–876.e33.

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.

Metzger, D., Ali, S., Bornert, J.M., and Chambon, P. (1995). Characterization of the amino-terminal transcriptional activation function of the human estrogen receptor in animal and yeast cells. J. Biol. Chem. *270*, 9535–9542.

Mintz, B. (1964). Formation of genetically mosaic mouse embryos, and early development of "lethal (t12/t12)–normal" mosaics. J. Exp. Zool. *157*, 273–292.

Mintz, B. (1965). Genetic mosaicism in adult mice of Quadriparental lineage. Science *148*, 1232–1233.

Moffitt, J.R., Bambah-Mukku, D., Eichhorn, S.W., Vaughn, E., Shekhar, K., Perez, J.D., Rubinstein, N.D., Hao, J., Regev, A., Dulac, C., et al. (2018). Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. Science *362*.

Nagy, A. (2000). Cre recombinase: the universal reagent for genome tailoring. Genesis *26*, 99–109.

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.

Nakamura, O., and Kishiyama, K. (1971). Prospective fates of blastomeres at the 32 cell stage of Xenopus laevis Embryos. Proc. Jpn. Acad. *47*, 407–412.

Nitzan, M., Karaiskos, N., Friedman, N., and Rajewsky, N. (2019). Gene expression cartography. Nature *576*, 132–137.

Nketia, T.A., Sailem, H., Rohde, G., Machiraju, R., and Rittscher, J. (2017). Analysis of live cell images: methods, tools and opportunities. Methods *115*, 65–79.

Pantazis, P., and Supatto, W. (2014). Advances in whole-embryo imaging: a quantitative transition is underway. Nat. Rev. Mol. Cell Biol. *15*, 327–339.

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*.

Porter, S.N., Baker, L.C., Mittelman, D., and Porteus, M.H. (2014). Lentiviral and targeted cellular barcoding reveals ongoing clonal dynamics of cell lines in vitro and in vivo. Genome Biol *15*, R75.

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. *36*, 442–450.

Rinkevich, Y., Lindau, P., Ueno, H., Longaker, M.T., and Weissman, I.L. (2011). Germ-layer and lineage-restricted stem/progenitors regenerate the mouse digit tip. Nature *476*, 409–413.

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.

Rodriques, S.G., Stickels, R.R., Goeva, A., Martin, C.A., Murray, E., Vanderburg, C.R., Welch, J., Chen, L.M., Chen, F., and Macosko, E.Z. (2019). Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. Science *363*, 1463–1467.

Salvador-Martínez, I., Grillo, M., Averof, M., and Telford, M.J. (2019). Is it possible to reconstruct an accurate cell lineage using CRISPR recorders? eLife *8*, e40292.

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. *33*, 495–502.

Sauer, B., and Henderson, N. (1988). Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Proc. Natl. Acad. Sci. USA *85*, 5166–5170.

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.

Siegal, M.L., and Hartl, D.L. (1996). Transgene coplacement and high efficiency site-specific recombination with the Cre/loxP system in Drosophila. Genetics *144*, 715–726.

Snippert, H.J., van der Flier, L.G., Sato, T., van Es, J.H., van den Born, M., Kroon-Veenboer, C., Barker, N., Klein, A.M., van Rheenen, J., Simons, B.D., et al. (2010). Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. Cell *143*, 134–144.

Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. Nat. Genet. *21*, 70–71.

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. *36*, 469–473.

Spemann, H., and Mangold, H. (1924). über induktion von embryonalanlagen durch implantation artfremder organisatoren. Arch. Mikrosk. Anat. Entwicklungsmechanik *100*, 599–638.

Stent, G.S. (1998). Developmental cell lineage. Int. J. Dev. Biol. *42*, 237–241.

Stern, C. (1936). Somatic crossing over and segregation in Drosophila melanogaster. Genetics *21*, 625–730.

Stern, C.D., and Fraser, S.E. (2001). Tracing the lineage of tracing cell lineages. Nat. Cell Biol. *3*, E216–E218.

Sternberg, N., and Hamilton, D. (1981). Bacteriophage P1 site-specific recombination: I. Recombination between loxP sites. J. Mol. Biol. *150*, 467–486.

Stickels, R.R., Murray, E., Kumar, P., Li, J., Marshall, J.L., Bella, D.D., Arlotta, P., Macosko, E.Z., and Chen, F. (2020). Sensitive spatial genome wide expression profiling at cellular resolution. bioRxiv. Published online March 14, 2020. https://doi.org/10.1101/2020.03.12.989806.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell *177*, 1888–1902.e21.

Stuart, T., and Satija, R. (2019). Integrative single-cell analysis. Nat. Rev. Genet. *20*, 257–272.

Sturtevant, A.H. (1929). The claret mutant type of Drosophila simulans: a study of chromosome elimination and of cell-lineage. Z. Wiss. Zool. *135*, 323–356.

Sulston, J.E., Schierenberg, E., White, J.G., and Thomson, J.N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. Dev. Biol. *100*, 64–119.

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.

Tarkowski, A.K. (1961). Mouse chimæras developed from fused eggs. Nature *190*, 857–860.

Tarkowski, A.K. (1964a). Patterns of pigmentation in experimentally produced mouse chimaerae. Development *12*, 575–585.

Tarkowski, A.K. (1964b). True hermaphroditism in chimaeric mice. Development *12*, 735–757.

Torre, E., Dueck, H., Shaffer, S., Gospocic, J., Gupte, R., Bonasio, R., Kim, J., Murray, J., and Raj, A. (2018). Rare cell detection by single-cell RNA sequencing as guided by single-molecule RNA FISH. Cell Syst *6*, 171–179.e5.

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.

Vogt, W. (1929). Gestaltungsanalyse am amphibienkeim mit örtlicher vitalfärbung. W. Roux. Archiv f. Entwicklungsmechanik *120*, 384–706.

Wagner, D.E., and Klein, A.M. (2020). Lineage tracing meets single-cell omics: opportunities and challenges. Nat. Rev. Genet. *21*, 410–427.

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.

Walsh, C., and Cepko, C.L. (1992). Widespread dispersion of neuronal clones across functional regions of the cerebral cortex. Science *255*, 434–440.

Wang, X., Allen, W.E., Wright, M.A., Sylwestrak, E.L., Samusik, N., Vesuna, S., Evans, K., Liu, C., Ramakrishnan, C., Liu, J., et al. (2018). Three-dimensional intact-tissue sequencing of single-cell transcriptional states. Science *361*, eaat5691.

Weinreb, C., and Klein, A.M. (2020). Lineage reconstruction from clonal correlations. Proc. Natl. Acad. Sci. USA *117*, 17041–17048.

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.

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.

Wetzel, R. (1929). Untersuchungen am hühnchen. Die entwicklung des keims während der ersten beiden bruttage. W. Roux' Archiv f. Entwicklungsmechanik *119*, 188–321.

Whitman, C.O. (1878). Memoirs: The Embryology of Clepsine. J. Cell Sci. s2–s18.

Whitman, C.O. (1887). A contribution to the history of the germ-layers in Clepsine. J. Morphol. *1*, 105–182.

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.

Zhu, Y., Huang, Y.F., Kek, C., and Bulavin, D.V. (2013). Apoptosis differently affects lineage tracing of Lgr5 and Bmi1 intestinal stem cell populations. Cell Stem Cell *12*, 298–303.