UCSF
UC San Francisco Previously Published Works

Title
Tools and Concepts for Interrogating and Defining Cellular Identity

Permalink
https://escholarship.org/uc/item/1mw930jk

Journal
Cell Stem Cell, 26(5)

ISSN
1934-5909

Authors
McKinley, Kara L
Castillo-Azofeifa, David
Klein, Ophir D

Publication Date
2020-05-01

DOI
10.1016/j.stem.2020.03.015

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Author Manuscript

HHS Public Access

Author manuscript

Cell Stem Cell. Author manuscript; available in PMC 2021 May 07.

Published in final edited form as:

Cell Stem Cell. 2020 May 07; 26(5): 632–656. doi:10.1016/j.stem.2020.03.015.

Tools and concepts for interrogating and defining cellular identity

Kara L. McKinley[*,1], **David Castillo-Azofeifa**[*,2,3], **Ophir D. Klein**[^,2,3,4,5]

[1]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA, USA

[2]Department of Orofacial Sciences, University of California, San Francisco, San Francisco, CA, USA

[3]Program in Craniofacial Biology, University of California, San Francisco, San Francisco, CA, USA

[4]Department of Pediatrics, University of California, San Francisco, San Francisco, CA, USA

[5]Institute for Human Genetics, University of California, San Francisco, San Francisco, CA, USA

Summary

Defining the mechanisms that generate specialized cell types and coordinate their functions is critical for understanding organ development and renewal. New tools and discoveries are challenging and refining our definitions of a cell type. A rapidly growing toolkit for single-cell analyses has expanded the number of markers that can be assigned to a cell simultaneously, revealing heterogeneity within cell types that were previously regarded as homogeneous populations. Additionally, cell types defined by specific molecular markers can exhibit distinct, context-dependent functions, for example between tissues in homeostasis and those responding to damage. Here, we review the current technologies used to identify and characterize cells, and we discuss how experimental and pathological perturbations are adding increasing complexity to our definitions of cell identity.

Keywords

stem cell; plasticity; cell competition; organoid; cell type; scRNAseq; smFISH; microscopy

---

^Correspondence to: ophir.klein@ucsf.edu.
*Co-first authors

Author contributions
KLM and DCA conceptualized, researched, and wrote the article, with input from ODK.

Declaration of interests
The authors declare no competing interests.

Cellular identity can be highly plástic and influenced by different physiological contexts and tools used to measure it. Here, McKinley, Castillo-Azofeifa, and Klein review the rapidly expanding toolkit that can be used to identify and probe cellular features and functions, and discuss how those properties change in different contexts.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author Manuscript
Author Manuscript
Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Introduction

Understanding how specialized cells work together to ensure tissue and organ function is a central objective of developmental and stem cell biology, and a critical step towards achieving this goal is to comprehensively catalog the cells that make up a tissue. Cells can be categorized according to their features, such as molecular markers, or according to their function (Figure 1). Feature- and function-based definitions are tightly linked: identifying features associated with a cell type is essential both to assess its function(s), and to define the signaling pathways, regulatory logic, and cellular structures that endow those functions. In addition, shared cellular features can provide hints of shared functions in systems for which direct functional assessments are challenging, particularly humans. In recent years, our capacity to define cellular features has exploded, with rapid advances in single-cell profiling generating a wealth of high-resolution, high-dimensional data that establish transcriptional, epigenetic and proteomic signatures of cells. These analyses are revealing previously unrecognized heterogeneity and reshaping our understanding of cellular identity.

Current studies are seeking to generate integrated definitions for cell types that encompass both features and functions, but an enduring challenge is that the contribution of a cell to a tissue varies according to its context. Thus, a specific feature, such as a molecular marker, may correlate with a particular function in one context, but feature and function can become uncoupled when context is altered. For example, due to a phenomenon termed cell competition, cells with mutations that appear neutral in one context can be selectively eliminated when combined with wild type neighbors, or wild type cells can be outcompeted by “super fit” cells. Even in contexts in which cells are genetically identical, heterogeneity in the surrounding signals can privilege certain cells, such that cells in close proximity to morphogens or niche factors expand at the expense of more distal cells.

Moreover, cell types that are stable during homeostasis in the adult can perform new functions and/or acquire new features under damage or disease conditions, a phenomenon termed plasticity. For example, cells that have made fate commitments can revert to less differentiated states (de-differentiation) or directly convert to a mature cell type of a distinct lineage (transdifferentiation), including epithelial to mesenchymal or mesenchymal to epithelial transitions (EMT or MET). In many cases, cell-type transitions arise during perturbations that alter a cell’s interaction with its microenvironment. Environmental and pathological perturbations can alter cellular microenvironments, as can many experimental strategies for cell type analysis that remove a cell from its native context and expose it to a new microenvironment, including through transplantation or *ex vivo* culture models. Thus, a key goal for establishing a comprehensive understanding of cell identity is to distinguish what a cell type does in steady state from what it is capable of doing in a given environment.

As a result of these recent studies of cellular features and functions, the term “cell type” has acquired multiple meanings and interpretations (Clevers, 2017). Cell type categories that depend on hard-wired functions, or on functions that invariably track with specific features, are being called into question. We are still in the early stages of comprehensively categorizing the cells in a tissue in homeostasis according to any one molecular layer (transcriptome, cell surface markers, chromatin architecture, and so on), and ongoing efforts

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

are seeking to connect these molecular maps of tissues to the underlying mechanisms of tissue function. Particular challenges include defining the functional consequences of unearthed heterogeneities and determining how the categories that these approaches identify correspond to cellular transitions along differentiation trajectòries in real time. We are also facing a need to generate experimental and computational frameworks to integrate cellular profiles generated with different modalities, and we are challenged to reconcile discrepancies between the groupings of cells that they define. New contexts and stimuli – such as injuries, diseases, aging, and environmental factors – will further refine these pictures, or perhaps upend them.

This review aims to present the diversity of frameworks from which to approach the problem of cell categorization, the tools available to pursue them, and concepts and challenges to consider in their interpretation and synthesis, with a predominant focus on mammalian epithelial tissues. We first synthesize and assess the strategies to categorize cells based on their features and functions. We then delve deeper into cellular function, and how cells exhibiting a specific feature or set of features can exhibit different functions in different contexts. We discuss key findings regarding the assessment of stem cell function and cell plasticity, in which cells exhibit expanded or altered functional repertoires following experimental manipulation or damage. Finally, we consider how cellular context can drive the selective elimination or expansion of certain cells through cell competition. Together, this work highlights the complex interplay between intrinsic and extrinsic properties that endow and coordinate cellular functions.

Tools for assessing cellular features and functions

In this section, we review the wide and rapidly expanding toolkit that is increasing the scale and precision with which tissues can be deconstructed into their component cell types. We focus on advances in pursuit of three major goals: 1) Detecting features associated with a cell type from a pre-defined list of candidates, 2) Identifying new features and cell types through unbiased approaches, and 3) Defining cellular relationships.

Goal 1: Detection of features associated with a cell type from a pre-defined list of candidates

Distinguishing cells based on a limited number of pre-selected features, such as morphology or expression of a set of specific genes or proteins, is a longstanding and powerful approach to distinguish and isolate cell types. Approaches to detect candidate features continue to play criticai roles in understanding cell identity even as unbiased –omic profiling approaches expand, in part due to trade-offs such as cost and ease of implementation. Moreover, defining a limited suite of identifying markers plays a critical role in further downstream characterization of a cell type, for example through genetic perturbations. Ongoing efforts are developing tools capable of detecting an increasing number of candidates in a single sample (Figure 2A).

Microscopy has been a powerful tool for the discrimination between cell types for over a century. Early work distinguished cells based on their morphology and dye-staining properties (Ehrlich, 1877; Golgi, 1885). In the mid-1900s, technologies emerged that

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

allowed cells to be detected based on molecular features: proteins detected with antibodies by immunocytochemistry/immunohistochemistry (Coons et al., 1941), and nucleic acids detected with complementary sequence probes by *in situ* hybridization (ISH; (Gall and Pardue, 1969)). These probes can be conjugated to enzymes that produce a colored precipitate for detection by brightfield microscopy, or to fluorophores, which allow them to be detected *in situ* using a fluorescent microscope or in dissociated cells by flow cytometry. Fluorescent proteins further expand this toolkit by allowing for genetic labeling and live imaging of proteins (Rodriguez et al., 2017) and protein-RNA complexes (Bertrand et al., 1998; Nelles et al., 2016) (Figure 2Ai and iv).

Compared to colorimetric approaches, fluorescent approaches increase the number of features that can be detected in a cell simultaneously, by labeling each detection reagent with distinct fluorophores. However, although the number of available dyes and fluorescent proteins for labeling is large, spectral overlap between fluorophores frequently limits the number of features that can be distinguished in a single experiment to around 3–5. Efforts are ongoing to increase the number of fluorescent labels that can be assessed in a single sample by either microscopy or flow cytometry, particularly using spectral approaches coupled with subsequent linear unmixing algorithms, which distinguish fluorophores according to their signature emission patterns across the spectrum instead of isolating specific wavelength ranges (Valm et al., 2017; Zimmermann, 2005). Alternatively, repeated cycles of antibody staining, signal removal, and re-staining with new antibodies can further increase the number of proteins that can be detected (Gerdes et al., 2013; Lin et al., 2015; Pirici et al., 2009). An expanded suite of proteins can also be detected by conjugating antibodies to DNA barcodes, which are then iteratively revealed by the addition of corresponding fluorescent (oligo)nucleotides (Goltsev et al., 2018; Saka et al., 2019), or by photocleavage of the oligo spot-by-spot and subsequent analysis (commercialized by Nanostring as Digital Spatial Profiling (Merritt et al., 2019)).

Highly multiplexed protein profiling can also be achieved by conjugating antibodies to non-biological metal isotopes (commonly lanthanides) instead of fluorophores, which are then detected by mass spectrometry (Figure 2Aii). This allows for the detection of, in principle, one hundred or more targets simultaneously (Bandura et al., 2009; Bendall et al., 2011). This approach, termed mass cytometry (commercialized as CyTOF), can detect features on dissociated cells in a manner analogous to fluorescence flow cytometry. For example, in recent work the expression of 73 proteins was evaluated in 26 million tumor and non-tumor cells to profile human breast cancer (Wagner et al., 2019). This approach has since been extended beyond cell surface antigens, allowing immune cells to be distinguished based on their global histone modification profiles (EpiTOF) (Cheung 2018), as well as to assess features of cellular metabolism (single cell metabolic profiling; scMEP (Hartmann et al., 2020)). Mass cytometry not only allows for profiling of dissociated cells, but can also be applied to detect proteins in intact tissue sections in a manner analogous to immunofluorescence. In imaging mass cytometry ((Giesen et al., 2014), commercialized as Hyperion) fixed tissue is labeled with lanthanide-conjugated antibodies and ablated with a high-resolution laser spot-by-spot. Each ablated spot is then transferred to the CyTOF for analysis, allowing for the detection of >30 epitopes with spatial resolution. Recent work has extended this approach to allow for simultaneous detection of both proteins and mRNA

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(Schulz et al., 2018). In Multiplexed Ion Beam Imaging (MIBI) (Angelo et al., 2014) (commercialized as IONpath), a focused ion beam is used to scan across the sample, liberating secondary ions from the lanthanides for detection by the mass spectrometer. The relative advantages of these approaches are reviewed elsewhere (Bodenmiller, 2016).

Whereas multiplexed protein detection strategies can report on tens to hundreds of features per sample, multiplexed RNA detection is approaching tens of thousands. A subset of these approaches, such as microarrays and the Nanostring nCounter (Geiss et al., 2008), multiplex the quantification of mRNA levels from homogenized populations of cells, increasing the number of targets that can be detected compared to quantitative reverse-transcriptase polymerase chain reaction (qRT-PCR) (Figure 2Av). Alternatively, individual mRNAs can be detected directly in fixed cells, providing single-cell resolution. These approaches generate sufficient signal for the detection of single RNA molecules either by tiling numerous labeled probes along the target sequence (Femino et al., 1998; Raj et al., 2008) (Figure 2Aviii), utilizing a small number of probes and a signal amplification step through scaffolding (Player et al., 2001; Wang et al., 2012; (Kishi et al., 2019) (Figure 2Avii), or rolling circle amplification (Larsson et al., 2010) (Figure 2Aix).

To multiplex these approaches, each mRNA sequence is assigned an identifying barcode, which can be read out either by *in situ* sequencing (Figure 2Aix) or fluorescent *in situ* hybridization (FISH) (Figure 2Aviii). For *in situ* sequencing approaches, the barcode is a short nucleotide sequence that is introduced into the cDNA amplicon during rolling circle amplification and read out by sequencing-by-ligation (Figure 2Aix). The earliest iteration of this approach allowed for the detection of 39 transcripts in tissue sections (Ke et al., 2013). The recently developed STARmap technique built on this approach using a modified amplification process and sequencing-by-ligation approach, as well as crosslinking of the amplicons within a tissue-hydrogel. Together, these developments increased the detection efficiency over earlier in situ sequencing approaches, allowing STARmap to detect transcripts from over one thousand genes, and in tissue sections up to 150 μm thick (Wang et al., 2018b). In addition to introducing the barcode during rolling circle amplification, cells can first be barcoded with virus, and the barcode RNA read-out by *in situ* sequencing (BAR-seq), an approach recently used to identify the projection patterns of individual neurons (Chen et al., 2019).

A number of techniques have also been developed for detecting individual RNA molecules by FISH (Codeluppi et al., 2018; Levsky et al., 2002; Lubeck and Cai, 2012). For sequential FISH (seqFISH) (Lubeck et al., 2014) and multiplexed error-robust FISH (MERFISH) (Chen et al., 2015b), each mRNA is assigned an ordered sequence of fluorophores as a barcode (Figure 2Aviii). This fluorophore sequence is read out over sequential rounds of hybridization, imaging, and either disruption of the hybrid or photobleaching. However, a challenge exists for these sequential imaging approaches. During each round of imaging, occasionally a spot that should fluoresce fails to be detected, or, conversely, stray probes or autofluorescence cause a spot that should not fluoresce during that imaging round to be mis-identified as positive. As a result, the sequence that is ultimately read out may be incorrect, potentially in such a way that one barcode is misidentified as another. To address this, MERFISH and subsequently an iteration of SeqFISH termed SeqFISH+ (described below)

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

incorporate an RNA encoding scheme based on Hamming distance, the number of errors that would convert one barcode into another (Chen et al., 2015). For example, a barcode library with Hamming distance of 2 requires that 2 errors occur for one barcode to be mis-identified as another. These approaches were initially applied in cultured cells and subsequently extended to tissues (Moffitt et al., 2016; Shah et al., 2016). The major limitation for increasing the detection of mRNAs to the transcriptome level is the high density of RNAs within a cell, leading to overlap between detected spots. Both SeqFISH and MERFISH have recently addressed this challenge, either by hybridizing only a subset of the RNAs in the cell with a given color at a given time (SeqFISH+; (Eng et al., 2019)) or physically expanding the specimen with expansion microscopy (MERFISH) (Xia et al., 2019). These approaches allow the detection of probes targeting 10,000 genes, a dramatic step towards unbiased profiling of gene expression *in situ.*

Broadly, these diverse techniques for analysis of pre-selected candidate features are subject to a variety of tradeoffs. For example, preserving spatial context, which may reveal new cellular relationships shaping cell identity, may come at the cost of throughput compared to techniques using dissociated cells. In addition, in many cases, as the number of features that can be detected increases (Figure 2A), so, too, does the challenge of executing the experiment in terms of reagent costs, equipment, and requirements for technical expertise. For example, lower feature numbers are often detectable with equipment commonly found in core facilities, while the execution of some highly multiplexed approaches remains confined to a relatively small number of labs. The commercialization of many key assays is making their implementation more straightforward, although in some cases high consumables costs may remain an important consideration, and optimization for a particular tissue of interest remains a critical step. Despite these considerations, the yield of highly multiplexed profiling techniques is immense, allowing us to generate increasingly comprehensive pictures of tissue composition, and reshaping our understanding of what distinguishes cells from one another and the biological networks that control these distinctions. The rapid growth of new and improved technologies in this arena continues, pushing towards higher throughput, more accurate, and more accessible tools that detect greater and greater numbers of features.

Goal 2: Identification of new features or cell types

The majority of strategies described above delineate cells according to a limited set of candidate features, necessarily introducing the investigator's preconceptions about which features might be important to examine. Exciting developments over the past few years, particularly in single cell sequencing technology, now allow for profiling cells from diverse tissues in an unbiased manner. These strategies are being used to identify new cell types and to annotate known cell types with new constellations of markers.

mRNA sequencing from single cells (scRNA-seq) has become the leading technology for molecular profiling of the cellular composition of organs and organisms (Cao et al., 2017; Fincher et al., 2018; Karaiskos et al., 2017; Plass et al., 2018; Regev et al., 2017; Tabula Muris et al., 2018) (Figure 2Avi). In brief, the scRNAseq workflow involves isolating single cells and assigning each a unique barcode, such that the mRNA from many cells can be

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

pooled for sequencing and subsequently re-assigned to their cell of origin (Hashimshony et al., 2012; Jaitin et al., 2014; Klein et al., 2015; Macosko et al., 2015; Ramskold et al., 2012; Tang et al., 2009) reviewed in (Ziegenhain et al., 2018)). Cells can then be grouped based on the similarity of their transcriptomes through unsupervised clustering (reviewed in (Kiselev et al., 2019)). These analyses are revealing heterogeneity within populations previously assumed to be homogenous, for example the spatial variation of diverse cell types in the intestinal epithelium, including tuft cells, enteroendocrine cells, and enterocytes (Beumer et al., 2018; Glass et al., 2017; Haber et al., 2017; Herring et al., 2018; Moor et al., 2018). They are also uncovering new and rare cell types (Grun et al., 2015; Jindal et al., 2018), such as the pulmonary ionocyte in the mammalian airway (Montoro et al., 2018; Plasschaert et al., 2018). Efforts are ongoing to reduce cost of scRNAseq approaches through multiplexed barcodes, thereby increasing the accessibility, throughput, and potential applications of these approaches for mechanistic studies (Cao et al., 2017; Datlinger et al., 2019; Gehring et al., 2019; Kang et al., 2018; McGinnis et al., 2019; Rosenberg et al., 2018; Stoeckius et al., 2018).

Several important considerations exist for transcriptome-centric approaches. First, these analyses report not only on stable cell types, but also on the transitions of cell types through states, for example transiting through the cell cycle, or maturation or activation states of immune cells (Jaitin et al., 2014; Shalek et al., 2013). While in some cases capturing these state transitions may be desirable, in other cases their effects may mask biological signals of interest. For example, heterogeneity between cells in different cell cycle stages can confound inference of developmental trajectories, such that in some cases it is preferable to regress out the effect of the cell cycle from the dataset ((Buettner et al., 2015; Vento-Tormo et al., 2018), discussed further in (Luecken and Theis, 2019)). These analyses also require dissociation of cells from the tissue, which can alter transcriptional profiles (van den Brink et al., 2017), although this effect can be mitigated by treatment with the transcriptional inhibitor actinomycin D (Act-seq; (Wu et al., 2017)). In addition, some cells are more sensitive to the dissociation process than others, which can introduce biases in the cells that are recovered. This dissociation bias can be reduced by sequencing RNA from individual nuclei rather than cells (single nucleus RNA-seq, snRNAseq, sNuc-seq) (Grindberg et al., 2013; Habib et al., 2016; Koenitzer et al., 2020; Lake et al., 2016).

Importantly, the cellular dissociation required for many scRNAseq approaches results in the loss of valuable information regarding the spatial context of the cells. To overcome this problem, numerous strategies have been developed to combine transcriptional profiling with spatial information. One possibility is to infer the position of scRNAseq profiles based on their expression of key landmark genes for which spatial position is known from *in situ* hybridization atlases (Achim et al., 2015; Karaiskos et al., 2017; Satija et al., 2015). Alternatively, cells from defined positions can be isolated by laser capture microdissection before sequencing (Baccin et al., 2019; Moor et al., 2018; Zechel et al., 2014), by photoactivation and cell sorting (NICHE-seq) (Medaglia et al., 2017), or by using a photo-uncaging system to hyrbridize DNA oligonucleotides to cells in illuminated regions (ZipSeq) (Hu et al., 2020). Recent technologies have also employed a more gentle tissue dissociation that preserves cell conjugates to identify interacting cells (Boisset et al., 2018; Giladi et al., 2020; Halpern et al., 2018).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

A subset of approaches combines transcriptional profiles with spatial information by maintaining tissue architecture during profiling. For example, short sequences can be sequenced directly in cells fixed on a microscope slide. This approach, termed Fluorescent *In Situ* Sequencing (FISSEQ), uses the same *in situ* cDNA synthesis and rolling circle amplification principles described above for multiplexed mRNA detection. However, instead of sequencing user-defined barcodes, this approach sequences approximately 30 bp of the transcript itself, allowing for unbiased determination of the identity of each amplicon (Lee et al., 2014) (Figure 2Aix'''). An alternative suite of approaches introduces spatial barcodes during the scRNAseq sample process, so that RNAs can be pooled for sequencing but subsequently mapped back to their coordinates. For example, INSTA-seq (Fürth et al., 2019) uses a sequencing-by-ligation approach similar to FISSEQ but with new developments to significantly reduce the number of imaging cycles to detect 12 base pair barcodes from each amplicon in situ before next-generation sequencing of the amplicon. The barcode then allows the reads to be mapped back to the amplicon's position in the cell. Alternative approaches introduce a positional barcode into each cDNA by arraying RNA-capture oligonucleotides with a discrete barcode at each position (Figure 2Avi). The first application of this approach allowed for discrimination between RNAs with ~100 μm resolution (Stahl et al., 2016). Recent developments have improved the resolution of such approaches, first to 10 μm (Slide-Seq; (Rodriques et al., 2019)) and subsequently to 2 μm (High-Density Spatial Transcriptomics, HDST; (Vickovic et al., 2019)). Finally, the development of multiplexed single molecule FISH approaches to detect ~ 10,000 genes described above (MERFISH coupled with expansion microscopy (Xia et al., 2019) and SeqFISH+ (Eng et al., 2019)) open up the possibility that multiplexed single molecule FISH, which previously required upfront selection of candidate genes, can be used for unbiased transcriptome-wide profiling with spatial resolution.

Approaches for spatial transcriptomics are evolving rapidly, and, although tradeoffs exist between approaches, new technologies are rapidly overcoming limitations. Positional barcoding approaches have recently been commercialized (as Spatial Transcriptomics, subsequently acquired and marketed as Visium) rendering these approaches particularly accessible. However, they have lower detection efficiencies compared with FISH-based detection approaches such as MERFISH and SeqFISH. An added benefit of optical approaches (FISH- and *in situ* sequencing) is that they provide information regarding subcellular localization of mRNAs, which plays important roles in diverse cellular functions (Jung et al., 2014; Lecuyer et al., 2007; Moor et al., 2017). Importantly, FISH-based approaches still require pre-designed oligonucleotides, and therefore do not facilitate the identification of unexpected transcript variants including single-nucleotide variants that can be detected by sequencing-based approaches.

Although transcriptome-centric strategies currently dominate unbiased cell categorization efforts, other –omic level profiling can facilitate the discrimination between cells and add more layers to cellular definitions. Recent work has reported label-free proteomic profiling from single cells by ultrasensitive mass spectrometry (Virant-Klun et al., 2016; Zhu et al., 2018) (Figure 2Aiii). This allows for clustering of cells, assignment of new proteins associated with specific cell types, identification of heterogeneity within populations, and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

ordering of cells along a developmental trajectory (Specht et al., 2019; Zhu et al., 2019), although these approaches are still in their infancy.

A particular focus of recent work has been the profiling of chromatin structure and composition (Figure 2B). A subset of bulk approaches for genome-wide profiling of DNA modifications, histone modifications, protein-DNA interactions, and chromatin accessibility have been modified for use in single cells ((Figure 2B), reviewed in (Ludwig and Bintu, 2019; Shema et al., 2019), facilitating their use for cell type classification. For example, DNA methylation signatures can distinguish cell types in the mammalian cortex (Luo et al., 2017; Mulqueen et al., 2018), as well as identify sister cells in the four-cell mouse embryo (Mooijman et al., 2016). Similarly, single-cell profiling of chromatin modifications or protein-DNA interactions (Figure 2B) can discriminate cellular subpopulations in culture (Kaya-Okur et al., 2019; Rotem et al., 2015) and in tumors (Grosselin et al., 2019). A particularly powerful approach for profiling cell types is genome-wide profiling of chromatin accessibility, for example nucleosome positioning through micrococcal nuclease digestion (MNase-seq) (Lai et al., 2018), or exposed DNA based on the preferential integration of transposons (single cell ATAC-seq and single cell combinatorial indexing ATAC-seq) (Buenrostro et al., 2018; Chen et al., 2018; Cusanovich et al., 2018a; Cusanovich et al., 2018b). These studies highlight a particular strength of epigenomic profiling for cell type characterization, which is the capacity to identify epigenetic changes that precede changes in gene expression (Inoue et al., 2019; Ziffra et al., 2019), which may for example be suggestive of priming for differentiation towards distinct lineages (Buenrostro et al., 2018; Lai et al., 2018).

Combined profiling of genomic, epigenomic, transcriptomic, and proteomic features can further refine cellular delineations and can reveal underlying regulatory relationships (Figure 2C). For example, numerous approaches layer additional molecular measurements on top of scRNA-seq data, combining transcriptomic profiling with genome-wide profiling of copy-number variants, DNA methylation, and chromatin accessibility, as well as measurements of candidate proteins (Figure 2C, reviewed in (Zhu et al., 2020)). These approaches can map additional molecular layers onto scRNA-seq-defined cell types, and, in some cases, distinguish cellular subpopulations beyond those identified from transcriptomic profiling alone (Stoeckius et al., 2017; Ziffra et al., 2019).. In addition to simultaneous measurements of distinct molecular features from the same cell, or the same pool of cells processed in parallel, new algorithms are facilitating the integration of discrete transcriptomic, epigenomic, and targeted proteomic datasets (Stuart et al., 2019; Welch et al., 2019). Together, these approaches move towards a more complete picture of cell identity and its underlying regulation, while further amplifying the challenge of determining the extent to which additional heterogeneity identified at each layer connects to variability at the functional level.

Goal 3: Define cellular relationships

3a) Live microscopy—A central component of a cell's identity is its position in the lineage hierarchy, meaning the identities of its mother and/or daughter cells. Live microscopy (Figure 3A) can reveal the ground truth of these mother-daughter relationships

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

by direct observation, in contrast to approaches that infer cellular relationships from snapshots of cells at discrete time points. Live imaging approaches can capture cell divisions, cell movements, cell death, and changes in morphology (Figure 1), and allow for continuous observations of specific cells over time to identify heterogeneities in behavior within a population, such as rates of differentiation. Advances in in vitro cell culture systems ranging from two-dimensional cultures to three-dimensional organotypic culture systems including slice cultures and organoids, as well as in imaging technologies and data analysis pipelines, are rapidly increasing the resolution, timeframes, and throughput of the assessment of the dynamics of cellular relationships.

The power of direct, continuous observation for defining cellular hierarchies is exemplified by the pioneering work of Charles Whitman, E.B. Wilson, Edward Conklin and others in the early embryos of marine invertebrates (Conklin, 1897; Whitman, 1887; Wilson, 1892), and by John Sulston's studies that defined the complete lineage tree of Caenorhabditis elegans (Sulston et al., 1983). In the stem cell field, live imaging approaches for defining cell lineage have been empowered by the development of *in vitro* stem cell culture systems that recapitulate aspects of *in vivo* division and differentiation patterns, both in two dimensions as well as in three dimensional culture systems such as organoids or explants. For example, mammalian neural stem/progenitor cell cultures derived from rodent embryos and adults, as well as human fetal brain, can be imaged by phase contrast microscopy every few minutes over the course of one or two weeks, and progeny fates defined by morphology and post-imaging immunostaining. This approach has allowed for direct assessment of asymmetric versus symmetric divisions and the construction of lineage trees, as well as identifying the relative timing of differentiation events (for example, the generation of neurons and glial cells) (Costa et al., 2011; Piltti et al., 2018; Qian et al., 1998; Qian et al., 2000; Ravin et al., 2008; Winter et al., 2015). In the hematopoietic system, live imaging and tracking of embryonic stem cell-derived cells and primary hematopoietic progenitors have clarified the generation of blood cells from embryonic endothelial cells (Eilken et al., 2009) and the instructive role of cytokines in generating monocytic or granulocytic cells from bipotent cells (Rieger et al., 2009). Similar approaches have also revealed heterogeneities in the differentiation of embryonic stem cells exposed to bulk signals (Brown et al., 2017a), as well as the ability of local signals to instruct cell division orientation and gene expression (Habib et al., 2013). These studies highlight the power of *in vitro* systems to identify changes in cellular properties over time and dissect mechanisms underlying cell fate choices by facilitating direct and rapid perturbations. Moreover, understanding cellular behavior in culture provides critical information for therapeutic applications that may require expansión of cells in culture before transplantation.

Live imaging can also be used for dynamic assessment of cellular behaviors and relationships in living organisms using a variety of approaches to render the tissue of interest optically accessible. In some cases, internal organs can be surgically exposed (Ewald et al., 2011), for example revealing immune cell migration patterns in the liver and spleen (Egen et al., 2008; Swirski et al., 2009), but these experiments are generally terminal. Implantation of a transparent window can facilitate repeated imaging sessions (Sandison, 1924). Dorsal skinfold chambers (Algire and Legallais, 1949) allow for the visualization of xenograft cell behaviors (Brown et al., 2001), such as cell division and therapeutic response (Orth et al.,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

2011). Alternatively, windows can be implanted over the organ of interest (Alieva et al., 2014). This approach has been applied particularly with mammary imaging windows and abdominal imaging windows, for example to track cell behaviors and stem cell dynamics in healthy tissues (Ritsma et al., 2014; Scheele et al., 2017), as well as the cellular dynamics of tumor growth and metastasis (Kedrin et al., 2008; Ritsma et al., 2012; Sobolik et al., 2016). In other cases, structures can be imaged directly without surgical interventions, and sites of interest can be revisited over multiple imaging sessions using anatomical markers or tattoos as references. For example, the calvarium (skull bones) of the mouse is sufficiently transparent to visualize behaviors of hematopoietic stem/progenitor cells or leukemic cells (Adams et al., 2009; Christodoulou et al., 2020; Colmone et al., 2008; Lo Celso et al., 2009; Mazo et al., 1998; Sipkins et al., 2005). The accessibility of the skin has rendered this organ particularly powerful to uncover cellular dynamics through optical approaches, both live imaging and targeted laser ablation to disrupt specific cells (Rompolas et al., 2012; Rompolas et al., 2013). Together, intravital imaging studies have generated a wealth of information of cellular behaviors required for tissue function (Marsh et al., 2018), as well as revealing heterogeneous behaviors within cell populations based on differences in their local environment, as we discuss further in the section on *Assessing cell function* (Mesa et al., 2018; Ritsma et al., 2014; Rompolas et al., 2013; Rompolas et al., 2016)

These live imaging analyses are facilitated by diverse microscopy techniques, coupled with technological advances for automated cell segmentation, tracking, and lineage reconstruction (Amat et al., 2014; Bao et al., 2006; Du et al., 2014; Faure et al., 2016; Mace et al., 2013; McDole et al., 2018; Ulman et al., 2017; Wan et al., 2019; Wolff et al., 2018). Broadly, a tradeoff exists between the physiological complexity of the system and technical complexity of the imaging experiment, with considerations including resolution, speed, signal-to-noise, cost, phototoxicity, and, increasingly, the computational challenges of the data analysis (reviewed in (Combs and Shroff, 2017; Thorn, 2016)). Conventional widefield microscopy has proved useful for imaging sparsely labeled cells in transparent organisms, such as developing zebrafish (Kimmel et al., 1990; Woo and Fraser, 1995). However, confocal microscopy (Minsky, 1961), which rejects out-of-focus light through the use of a pinhole in front of the detector, has been particularly valuable for tissue imaging, as it allows thick samples to be imaged in discrete optical sections.

A major challenge for tracing cellular hierarchies in living, three-dimensional tissues and organisms is that imaging illumination can damage the system, such that the act of observing cellular behavior can alter it (Magidson and Khodjakov, 2013). Conventional widefield and confocal microscopes illuminate fluorophores outside of the imaging focal plane, causing photodamage to regions that do not participate in generating the final image. In contrast, lightsheet fluorescence microscopy illuminates a single plane within the object at a given time (Huisken et al., 2004), allowing for low-photodamage optical sectioning. Broadly, lightsheet fluorescence microscopes use a cylindrical lens (Huisken et al., 2004), digital scanning laser beam (Keller et al., 2008) or Bessel beam (Planchon et al., 2011) to form a sheet that illuminates only a thin volume of the sample. Moving the specimen through the lightsheet, or scanning the sheet over the sample, allows the full sample volume to be imaged. These approaches allow for visualization of cell divisions and movements with high temporal and spatial resolution in developing organisms (Huisken et al., 2004; Keller et al.,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

2008; Krzic et al., 2012; McDole et al., 2018; Tomer et al., 2012; Udan et al., 2014; Wolff et al., 2018; Wu et al., 2013) and organoids (McKinley et al., 2018; Serra et al., 2019). Finally, efforts are underway to further reduce illumination requirements on conventional microscopes using deep learning-based approaches to improve the signal-to-noise ratio of images collected under low light conditions (Fang et al., 2019; Weigert et al., 2018).

When imaging tissue, variations in refractive index between the sample and its medium and between different objects in the sample lead to optical aberrations that distort the image. Although efforts are underway to resolve medium-sample refractive index mismatches (Boothe et al., 2017), light scattering and absorbance from cellular components within the tissue remain major obstacles. For fixed tissue, clearing approaches that seek to equilibrate refractive index throughout the sample can reduce scattering and allow for high-resolution volumetric imaging (reviewed in (Richardson and Lichtman, 2015)). For living tissues, one possibility to circumvent this issue is to acquire images of the specimen from multiple angles (multiview imaging). This can be achieved by rotating the sample or by imaging from multiple positions simultaneously using additional objectives (Chhetri et al., 2015; Krzic et al., 2012; Royer et al., 2016; Tomer et al., 2012; Wu et al., 2013). These different views can then be registered computationally (Preibisch et al., 2010). Alternatively, adaptive imaging or adaptive optics approaches can detect optical aberrations and apply corrections to compensate for them (Ji, 2017; Liu et al., 2018; Royer et al., 2016; Wilding et al., 2016). This allows for the long-term imaging of systems that change their optical properties over time, for example imaging of the development of whole mouse embryos from E6.5 to E8.5 (McDole et al., 2018). Greater tissue penetration can be achieved with the use of longer wavelengths, which, broadly, scatter less. For example, use of a far-red fluorescent reporter allowed for light-sheet imaging of the developing mouse heart, 600 μm deep within the embryo (McDole et al., 2018). The low-scattering nature of long-wavelength light also contributes to the capacity of two-photon microscopy to improve tissue penetration (Denk et al., 1990). In two-photon microscopy, a fluorophore is excited by absorbing two low-energy photons essentially simultaneously. Since the likelihood of two photons hitting the fluorophore rapidly falls off away from the focal point, two-photon microscopy minimizes out-of-focus fluorescence and generates extremely low background. The improved tissue penetration of two-photon imaging has made it particularly well suited for intravital imaging approaches. Together, these approaches provide a rapidly expanding toolkit for the assessment of cellular relationships through direct observation.

3b) Lineage tracing—Lineage tracing – using a heritable mark to track progeny of cells of interest – can facilitate the identificaron of cellular relationships through live imaging or in fixed or dissociated tissues. Early lineage tracing studies took advantage of natural variations in pigmentation ((Conklin, 1905; Rawles, 1948)), gross chromosomal markers (Ford et al., 1956; Wu et al., 1968) or features such as heterochromatin distribution (Le Douarin, 1980). Alternatively, dyes can be applied or injected (Serbedzija et al., 1989; Vogt, 1929) or marker transgenes can be integrated into the genome by viral transduction (Price et al., 1987; (Dick et al., 1985; Keller et al., 1985). Current applications focus particularly on tracing cellular progeny, using endogenous or induced genetic variants (Figure 3B).

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

To identify the progeny of cells expressing a particular gene or small subset of genes, site-specific genetic recombination can be used to drive expression of a reporter gene, such as a fluorescent protein or enzyme, in the cells of interest and their progeny (reviewed in (Hsu, 2015; Kretzschmar and Watt, 2012)). These experiments commonly use Cre recombinase expressed under control of a cell type-specific promoter, which catalyzes recombination at DNA recognition motifs termed loxP sites to drive reporter expression. Recombination can be induced at a specific time point using doxycycline-inducible or tamoxifen-inducible Cre systems, with the caveat that high doses of tamoxifen have been found to affect the mammary gland (Rios et al., 2014; Shehata et al., 2014), pancreas (Ahn et al., 2019), intestine (Zhu et al., 2013) and stomach (Huh et al., 2010). In addition to Cre/lox, alternative pairs of recombinases and recombination sites include Flp/FRT and Dre/Rox. This variety of recombination approaches can be used to track multiple cell types and their progeny simultaneously. In addition, they allow for intersectional methods to mark a cell type identified by a unique combination of genes, rather than a single gene by creating a logical AND gate, in which two genes must be expressed to drive marker expression (Hermann et al., 2014; Madisen et al., 2015). An AND gate can also be created by splitting Cre into two parts under control of different promoters, so that they can only form a functional Cre and drive recombination in cells in which both promoters are expressed (Casanova et al., 2003; Xu et al., 2007). Split Cre approaches can also be combined with domains that dimerize in the presence of defined wavelengths of light, allowing for activation of Cre recombination in cells selected microscopically (Meador et al., 2019; Taslimi et al., 2016).

A variety of reporters are available that can mark all cells in the population uniformly upon recombination; alternatively, reporters that generate different marks within a population of cells allow progeny from different cells within the population to be distinguished (Figure 3B). For example, the Mosaic Analysis with Double Markers (MADM) system can mark sibling cells with distinct fluorescent proteins (Zong et al., 2005). In Brainbow or confetti reporters, recombination assigns one of many fluorescent proteins to each cell at random (Livet et al., 2007; Snippert et al., 2010b), The capacity to generate distinct fluorescent marks within the population allows many different cells expressing a common gene to be readily distinguished from one another, for example to identify, track, and morphologically characterize them microscopically in complex environments (Currie et al., 2016; Livet et al., 2007), as well as to determine how cells derived from a common progenitor are distributed within a tissue (McKinley et al., 2018; Pan et al., 2013). Crucially, the capacity to track multiple independent clones can reveal whether progenitor cells within a population exhibit different potential (Ghigo et al., 2013; Rinkevich et al., 2011; Snippert et al., 2010). The resolution of such clonal lineage analysis improves with an increasing number of potential marks; the potential diversity of labels can be increased by replacing fluorescent reporters with DNA sequences, as in the *Polylox* reporter, which can generate a maximum of 1.8 million unique DNA barcodes in Cre-expressing cells (Pei et al., 2017). Such DNA-based tools facilitate high-resolution lineage tracing of a large number of clones compared to fluorescent proteins, albeit with the loss of spatial and morphological information due to the dissociation required to recover the barcode.

Numerous strategies are available to generate highly variable DNA sequences in cells of interest and their progeny. One approach uses a library of DNA barcodes to mark cells,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

which can be introduced into the cells by viral transduction (Walsh and Cepko, 1992; (Lu et al., 2011; Schepers et al., 2008), zinc-finger mediated homologous recombination (Porter et al., 2014), or transposition (TracerSeq, (Wagner et al., 2018)). If the barcode is transcribed, the clonal relationships can be integrated with transcriptomic profiling by scRNA-seq (Biddy et al., 2018; Wagner et al., 2018; Weinreb et al., 2020; Yao et al., 2017). Alternatively, as the virus integrates quasi-randomly into the genome, the insertion site can function as the barcode. This approach has been powerful for studies of the human hematopoietic system, for which cells can be barcoded ex vivo and transplanted to trace clonal dynamics during repopulation *in vivo* in mice (Guenechea et al., 2001; Lu et al., 2011; McKenzie et al., 2006; Nolta et al., 1996) as well as non-human primates (Schmidt et al., 2002; Yu et al., 2018). Moreover, the use of viral vectors for gene therapy has allowed for longitudinal profiling of insertion sites following transplantation of gene-corrected hematopoietic stem/progenitor cells into human patients, particularly as a test for clonal dominance events induced by the insertion (Aiuti et al., 2013; Biffi et al., 2013). Importantly, the transplantation process to introduce viral-barcoded cells *in vivo* may affect cellular functions. For example, a mouse model that mobilizes a transposon to generate insertion-site barcodes allowed for profiling of steady-state hematopoiesis (Sun et al., 2014) and revealed significant differences from the transplantation hematopoeitic hierarchy (Rodriguez-Fraticelli et al., 2018) (reviewed in (Baron and van Oudenaarden, 2019). We discuss modulation of cell function by transplantation further in the *Assessing cell function* section.

Dynamic editing of a target locus is increasingly used to generate diverse DNA-based barcodes for lineage tracing. The majority of these systems use CRISPR/Cas9 mutagenesis, in which Cas9 is directed to cut at a specific sequence in the genome based on complementarity with a short, user-supplied RNA sequence termed a single guide RNA (sgRNA). This results in genetic lesions in the selected regions through error-prone repair of double-strand breaks (reviewed in (McKenna and Gagnon, 2019)). These approaches target Cas9 to cut within synthetic arrays, as in GESTALT (McKenna et al., 2016), scGESTALT (Raj et al., 2018), MEMOIR (Frieda et al., 2017), CARLIN (Bowling et al., 2019), Zombie (Askary et al., 2019) and others (Chan et al., 2019). Alternatively, Cas9 can cut within fluorescent proteins, as in the ScarTrace method (Alemany et al., 2018; Junker et al., 2017), LINNAEUS (Spanjaard et al., 2018) and others (Schmidt et al., 2017). Finally, a subset of approaches direct cuts within the sequence encoding the sgRNA itself, in the case of the MARC1 mouse (Kalhor et al., 2018) and mSCRIBE (Perli et al., 2016). The mutations can then be read out by sequencing the target locus (Alemany et al., 2018; Junker et al., 2017; McKenna et al., 2016; Schmidt et al., 2017), or, if the barcode is transcribed, by scRNA sequencing (Alemany et al., 2018; Bowling et al., 2019; Chan et al., 2019; Raj et al., 2018; Spanjaard et al., 2018) or by smFISH (Frieda et al., 2017; Askary et al., 2019). Recent work has also used phage integrases as an alternative approach to CRISPR to generate high-diversity dynamic barcodes that can be read out by smFISH (intMEMOIR) (Chow et al., 2020). The high diversity of DNA-based lineage tracing approaches makes them amenable for simultaneous lineage tracing from a wide variety of different cell types. This largely unbiased approach not only generates lineage hierarchies with much higher throughput, but is also revealing intriguing circumstances in which cells from different lineages converge on a given transcriptional signature, suggestive of convergent differentiation (Chan et al., 2019;

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Wagner et al., 2018). Thus, these unbiased approaches are revealing that the synthesis of cell lineage and cell features can refine models for cell identity and differentiation trajectories.

Finally, in addition to introduced DNA edits, lineage can also be traced retrospectively using spontaneous somatic mutations (Behjati et al., 2014; Frumkin et al., 2005) such as long interspersed nuclear element 1 (LINE-1) retrotransposition events, copy-number variants, single-nucleotide variants and microsatellite growth/shrinkage. In addition to mutations in nuclear DNA, it is possible to use mutations in mitochondrial DNA, which are highly enriched when preparing ATAC-seq libraries, for simultaneous profiling of lineage relationships and chromatin accessibility (Ludwig et al., 2019; Xu et al., 2019) (Figure 3B). These retrospective lineage tracing approaches allow for the identification of cellular hierarchies in systems not amenable to directed genetic modification, and have therefore provided significant insights into lineage hierarchies for human tissues such as the brain, blood, and embryo (Biezuner et al., 2016; Cai et al., 2014; Evrony et al., 2015; Ju et al., 2017; Lee-Six et al., 2018; Lodato et al., 2015; Osorio et al., 2018).

3c) Single-cell transcriptomics—Even without a lineage-tracing component, scRNAseq technologies can provide information on the relationships between cells (Figure 3C) (reviewed in (Lederer and La Manno, 2020)). Because existing scRNAseq pipelines lyse or fix the cells to define transcriptomes, it is not possible to track changes over time on the per-cell level. However, many cell types at different stages of differentiation are present within a tissue at a given time, such that, when the population is considered as a whole, a static snapshot can encapsulate all of the steps along the differentiation trajectories of a tissue (for example, (Haber et al., 2017; Halpern et al., 2017)) or whole organism (Plass et al., 2018; Siebert et al., 2019). Samples can also be taken from multiple different timepoints, allowing for the characterization of differentiation trajectories across development of entire embryos (Briggs et al., 2018; Farrell et al., 2018; Wagner et al., 2018). Differentiation trajectories can be inferred from these snapshot data by pseudotime analysis, also known as trajectory inference, which orders cells based on si milarity in their gene expression (Haghverdi et al., 2016; Trapnell et al., 2014). Over seventy trajectory inference tools have been developed, which are reviewed and evaluated elsewhere (Saelens et al., 2019; Tritschler et al., 2019; Weinreb et al., 2018b). Importantly, pseudotime analysis infers an ordering of cells according to their relative progress through a biological process, but does not provide information about the actual duration of events. Recent efforts have sought to couple real-time information with transcriptional ordering. For example, a recent study used a bifluorescent reporter comprised of two fluorophores with different maturation kinetics and half-lives as a cellular timer that positions transcriptional profiles relative to real time (Gehart et al., 2019).

New technologies and approaches are moving beyond inferred differentiation trajectories to direct measurements of dynamics from single cell transcriptomic data. Pioneering work revealed that scRNAseq data not only report on a cell’s gene expression levels at the time point of the experiment, but also contain information on a cell’s immediate future by incidentally capturing newly transcribed, unspliced precursor mRNAs (La Manno et al., 2018). The relative amounts of nascent RNA and mature mRNA are reflective of how gene expression is changing in the cell. When a gene has been recently activated, there are more

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

unspliced transcripts than spliced; conversely, when a gene has been recently repressed, spliced transcripts persist but unspliced transcripts decrease. Thus, measurements of spliced and unspliced transcripts reveal the rate of change of mRNA (RNA velocity) and therefore predict future mRNA abundance and forthcoming transcriptional states. Ongoing work is extending this framework, for example encompassing multi-omics data (Gorin et al., 2019) and reconstructing trajectories from sparse velocity information (Qiu et al., 2019). Alternatively, since mRNAs are transcribed in the nucleus and exported to the cytoplasm upon maturation, relative nuclear and cytoplasmic mRNA abundance, measured by FISH, can also predict future transcriptional states *in situ* (Xia et al., 2019). Finally, whereas RNA velocity takes advantage of serendipitously captured nascent RNA, metabòlic labeling approaches in which nascent RNA is labeled with 4-thiouridine (4sU) or 5-ethynyl-uridine (EU) have recently been coupled with scRNA seq to facilitate direct assessment of future transcriptional profiles in single cells (Battich et al., 2020; Cao et al., 2019; Erhard et al., 2019; Hendriks et al., 2019; Qiu et al., 2019).

The toolkits to detect, identify, and analyze diverse cellular features and functions are expanding at breakneck speed. These approaches are continuing to push us towards a more precise mechanistic understanding of organ function during homeostasis.

Assessing cell function: modulation of cellular contributions by tissue context

As the tools and strategies described above are applied to organs experiencing mutational, damage, and disease burdens, they are revealing that the connections between cellular features and functions are highly dependent on a cell's context. In this section, we discuss recent studies assessing cellular functions across diverse contexts. We first discuss how experimental or pathological alterations are reshaping our understanding of stem cell function. We then examine how cellular context can determine whether a cell expands within the tissue or is eliminated through competitive interactions. These studies are underscoring the limitations of viewing cell identity as a hardwired, intrinsic property, and expanding our understanding of the intricate relationships between cells and their microenvironment.

Functional definitions of epithelial stem cells

Many efforts to define cellular functions are focused on the identification and characterization of stem cells, which hold significant potential for therapeutic applications due to their capacity to expand and generate diverse cell types. Broadly, stem cells are defined by their capacity to both maintain the stem cell population through self-renewal and to generate many distinct differentiated cell types. Initial models proposed that epithelial stem cells execute both functions with every division, generating one stem cell and one differentiating cell through asymmetric divisions (Potten, 1974). However, it has become clear that this paradigm does not hold in many epithelia. Pioneering work in the skin demonstrated that, in addition to asymmetric divisions yielding one proliferating cell and one committed cell, progenitors could also undergo symmetric divisions that generate two progenitors or two committed progeny (Clayton et al., 2007). Subsequent work in the male germline and intestine of both mice and flies showed that stem cells in these tissues can be

Author Manuscript

stochastically lost and replaced by new stem cells arising from symmetric division of their neighbors (de Navascues et al., 2012; Klein et al., 2010; Lopez-Garcia et al., 2010; Sheng and Matunis, 2011; Snippert et al., 2010). Later work has revealed similar principles at work in other solid tissues, including the esophagus (Doupe et al., 2012), oral mucosa (Jones et al., 2019), and epidermis (Rompolas et al., 2016). These experiments and others have focused the definition of epithelial stem cells on their ability to renew the stem cell population and generate differentiating cells of diverse lineages at the population level, rather than at each division (Post and Clevers, 2019).

Although epithelial stem cells are broadly capable of generating progeny that adopt either differentiating or stem cell fates, these decisions can be biased by local variations in the microenvironment, such as proximity to niche signals. For example, in the mouse intestine, the tissue is compartmentalized into invaginations called crypts, each containing numerous stem cells. Over time, these crypts become clonal as progeny from one stem cell eventually take over the entire unit (Griffiths et al., 1988; Potten and Loeffler, 1990; Winton et al., 1988). Intravital imaging revealed that stem cells at the bases of intestinal crypts are approximately three times more likely to colonize the crypt than stem cells at more peripheral positions (Ritsma et al., 2014). Similarly, live imaging analysis in the hair follicle demonstrated that the position of a stem cell within the hair follicle niche correlates with stem cell fate (Rompolas et al., 2013). Intriguingly, recent work has demonstrated that adult stem cells in the intestine arise from cells that receive a positional advantage during reorganization of the tissue during development (Guiu et al., 2019). These data suggest that proximity to niche signals can allow a subset of cells to expand preferentially (Figure 4A). In addition, recent work suggests that neighboring cell behavior can also modulate stem cell expansion. For example, recent live-imaging analysis of epidermal stem cell behavior revealed that epidermal stem cell division is triggered by neighboring stem cell differentiation (Mesa et al., 2018). Together, these data suggest that microenvironmental signals can distinguish between apparently homogenous stem cells, privileging the expansion of subsets of cells and generating functional heterogeneity within epithelial stem cell populations.

Functional tests of stemness

Extensive work has sought to characterize stem cells in diverse epithelia by identifying specific cellular features and associated markers that enrich for cells with the long-term potential to populate a tissue. As ongoing studies test stemness with new technologies and in new contexts, they are expanding our understanding of stem cell function and revealing limitations of previous definitions and experimental strategies. In particular, these studies have shown that repopulation potential can segregate with different cellular features depending on the physiological context or the technique used to assess it (Figure 4B–D).

Transplantation is a longstanding approach to identify stem cell populations (Figure 4C). Trailblazing work in the 1950s demonstrated that transplanted material could repopulate the hematopoietic system following irradiation (Ford et al., 1956; Lorenz et al., 1951) and the mammary gland following fat pad clearing (Deome et al., 1959). Moreover, the engrafted material could be re-transplanted into secondary recipients, indicating long-term self-

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

renewal capacity (Barnes et al., 1959; Daniel et al., 1968; Deome et al., 1959; Hoshino and Gardner, 1967). In the hematopoietic system, this approach demonstrated the existence of multipotent, self-renewing stem cells (Becker et al., 1963; Siminovitch et al., 1963; Till and Ea, 1961; Wu et al., 1968) that were subsequently prospectively isolated by their complement of cell surface markers (Spangrude et al., 1988). These early basic science discoveries have led to enormous clinical advances, as hematopoietic stem cell transplantation has revolutionized the treatment of hematological disorders and malignancies (Appelbaum, 2007).

However, transplantation studies have also given rise to controversy over the potential of stem cell populations. In the mammary gland, transplantation studies demonstrated that increasingly refined cell populations and, eventually, single cells, could give rise to both of the major epithelial lineages of the mammary gland, luminal cells and basal myoepithelial cells (Kordon and Smith, 1998; Shackleton et al., 2006; Stingl et al., 2006). However, subsequent work suggested that the mammary gland is repopulated under physiological conditions by lineage-restricted progenitors (Van Keymeulen et al., 2011), suggesting that the broader multipotency observed in the transplant studies resulted from plasticity induced by the new cellular microenvironment. Indeed, the extent to which the mammary gland is maintained by bi-potent stem cells or unipotent progenitors is still the subject of debate (reviewed in (Lloyd-Lewis et al., 2017)). Similarly, in the skin, *Lrig1*-positive cells can give rise to all epidermal lineages following grafting, but they exhibit more limited potential and contribute only to the interfollicular epidermis and sebaceous gland under steady state conditions (Jensen et al., 2009). In addition to expanded potential when transplanted orthotopically into a damaged site, transplantation to an ectopic site can also expand cellular potential. One dramatic example of environment shaping function is that epithelial cells of the thymus can contribute to hair follicle lineages following grafting (Bonfanti et al., 2010).

These studies reveal that the interpretation of transplantation experiments aiming to characterize stem cells can be affected both by introducing cells into a new microenvironment, as well as through possible injuries and regenerative programs induced by the transplantation process. In contrast, the advancement of lineage tracing methods, described above, has provided a minimally invasive solution to assess stem cell potential *in situ.* Nonetheless, transplantation approaches continue to facilitate new discoveries in stem cell function in health and disease, particularly for genetically intractable systems such as humans (Shimokawa et al., 2017), and also provide important information about the potential of cells to contribute therapeutically to repair.

An alternative approach to assess stem cell function is to determine the capacity of a cell type to expand and generate diverse cell types in culture. For example, in some cases, single cells plated *in vitro* in extracellular matrices can give rise to organoids—stable, complex tissues with diverse cellular composition—thereby indicating both the self-renewal capability and multilineage potential of the cells (Figure 4D). Based on early work showing that 3D culture can generate structures with functional properties and cell composition resembling the *in vivo* mammary gland (Barcellos-Hoff et al., 1989; Lee et al., 1985), the formation of mammary gland cultures capable of propagation in culture, termed mammospheres (Dontu et al., 2003), was adopted as a strategy to isolate mammary gland

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

stem cells (Liao et al., 2007). The subsequent identificaron of organoid-forming capacity in single intestinal stem cells (Sato et al., 2009) has led to the now widely used intestinal organoid model. This approach has been subsequently expanded to tissues including the stomach (Barker et al., 2010; Stange et al., 2013), liver (Hu et al., 2018; Huch et al., 2013), airway (Rock et al., 2009), and tongue (Hisha et al., 2013; Ren et al., 2014). Importantly, as organoids are minimal systems removed from their native context, they require supplementation with growth factors to support their growth. The combination of additives present in the medium has a significant impact on organoid forming capacity. For example, depending on the culture conditions, intestinal organoid formation is either restricted to *Lgr5*-positive cells, reflective of their unique multilineage potential under homeostasis, or expanded to non-*Lgr5*-positive cells (Castillo-Azofeifa et al., 2019; Serra et al., 2019; van Es et al., 2012). Thus, the process of generating organoids can also uncover non-homeostatic potential, and when assessing stemness through organoid-forming potential, it is important to consider the extent to which the culture conditions reflect the *in vivo* microenvironment.

Modulation of stem cell behavior by damage

In addition to experimental manipulations, tissue damage can also cause stem cell populations to expand their potential, or cause new cell populations to acquire stem-like properties (Figure 4B). For example, in the skin, stem cells from distinct compartments can mobilize towards wounds to repair damaged tissue, in some cases occupying new niches and adopting the stem cell function associated with their new position (Hoeck et al., 2017; Ito et al., 2005; Levy et al., 2007; Page et al., 2013; Rompolas et al., 2013; Snippert et al., 2010a). Additional work has revealed that cells can also cross lineage boundaries following damage in tissues of the intestine, stomach, tooth and lung ((Ayyaz et al., 2019; Castillo-Azofeifa et al., 2019; Jadhav et al., 2017; Sharir et al., 2019; Tata et al., 2013; Tian et al., 2011; van Es et al., 2012; Yui et al., 2018) and reviewed in (Burclaff and Mills, 2018; de Sousa e Melo and de Sauvage, 2019; Tata and Rajagopal, 2017)). These results call to mind classical experiments in *C. elegans that* demonstrated that cells could compensate for cell types lost by laser ablation (Kimble, 1981; Sulston and White, 1980).

In both the intestinal epithelium and the trachea, cells that have already made fate commitments can revert into stem cells, essentially reversing the traditional lineage hierarchy in homeostasis (Tata et al., 2013; Tian et al., 2011). For example in the intestine, diverse cell types of both the secretory and absorptive lineages can undergo reversion to stemness in response to assaults including irradiation, chemotherapy, helminth infection and DSS-induced colitis (reviewed in (de Sousa and de Sauvage, 2019)). Intriguingly, this response includes a re-activation of a number of genes associated with fetal development (Nusse et al., 2018; Yui et al., 2018), a phenomenon also observed in the stomach (Fernandez Vallone et al., 2016). This capacity to restore the stem cell pool from non-stem cells is observed in tumors as well, but it exhibits striking context-dependence. In particular, adenomas in the intestine arise from *LGR5+* stem cells (Schepers et al., 2012) but can be maintained by non-stem (*LGR5–*) cells if the stem cell pool is compromised (de Sousa e Melo et al., 2017). In contrast, *LGR5*+ stem cells are critical for maintenance of intestinal-derived liver metastases (de Sousa e Melo et al., 2017; Fumagalli et al., 2020) (Figure 4E, F).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Although many of the cell types that have been shown thus far to revert to stemness are not terminally differentiated, recent work has revealed that cells exhibiting highly specialized functions, which under homeostasis contribute at very low levels to tissue repopulation, are able to re-enter the cell cycle and repopulate the tissue following damage. In the intestine, post-mitotic Paneth cells can re-enter the cell cycle and produce ribbons of progeny in response to inflammation (Schmitt et al., 2018). Similarly, Chief cells in the stomach, and hepatocytes and cholangiocytes in the liver, can drive dramatic expansion of tissue in response to damage (Font-Burgada et al., 2015; Leushacke et al., 2017; Stange et al., 2013). Finally, compensation for lost functions is not restricted to recovery of stem cell functions, as, when Paneth cells are ablated in the intestine, enteroendocrine and tuft cells can adopt the Paneth cell position and provide some of this cell type's stem cell-supporting functions (van Es et al., 2019). Together, these studies have demonstrated that many cell types have significant potential to adopt features and functions beyond those observed in homeostasis, which can be unleashed in response to both experimental perturbations and damage.

Modulation of cellular contributions through competitive interactions with neighboring cells

An additional example of context-dependent cell function is cell competition, in which the identity of neighboring cells determines whether a cell expands in the tissue or is eliminated. In particular, cells that contribute to tissue function when surrounded by genetically identical cells may be actively eliminated when brought into contact with cells of increased fitness (Figure 4J–K). As a result, fitter cells become 'winners' and expand and colonize the tissue at the expense of weaker cells called 'losers', which are eliminated through engulfment, apoptosis, extrusion, delamination and differentiation (Figure 4K). Thus, the contribution of a cell to a tissue can be dramatically modulated by its fitness relative to its neighbors.

Cell competition was first described in the *Drosophila* wing imaginal disc (Morata and Ripoll, 1975b). Flies heterozygous for mutations in ribosomal genes (termed *minutes)* are viable and fertile, but have slower growth rates and minor structural abnormalities. Cells heterozygous for *minute* survive as long as the entire tissue is composed of heterozygous mutant cells. However, in the presence of wild-type cells, mutant cells are eliminated by apoptosis (Morata and Ripoll, 1975b). Similarly, mice heterozygous for a ribosomal mutation (*bst)* are viable with minor defects, but *bst* heterozygous cells are eliminated when combined with wild-type cells in chimeras (Oliver et al., 2004). Thus, mutations that are seemingly inert when introduced throughout the tissue can confer a selective disadvantage when combined with cells of a distinct genetic complement.

A wide variety of mutations can enable a cell of a given genetic background to make substantially different contributions to the tissue depending on its fitness relative to its neighbors. Many of these were first identified in *Drosophila,* as summarized in Figure 5 and reviewed in detail elsewhere (Baker, 2017; Bowling et al., 2019; Johnston, 2014; Nagata and Igaki, 2018). Subsequent work has identified cell competition in a variety of mammalian systems, including the hematopoietic system (Bondar and Medzhitov, 2010), epiblast (Clavería et al., 2013; Díaz-Díaz et al., 2017; Sancho et al., 2013), and embryonic skin (Ellis et al., 2019). Many studies have focused on cells harboring heterozygous loss-of-function

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

alleles, which are eliminated when apposed to wild type neighbors, as in the case of the key growth regulator *myc* and its related isoforms (Clavería et al., 2013; Ellis et al., 2019). Conversely, gain-of-function mutations, such as those that upregulate *myc*, can increase a cell's fitness relative to its neighbors, generating "super-competitors" that expand at the expense of wild type neighbors (Clavería et al., 2013; Moreno and Basler, 2004). Similar interactions may also be at play among stem cell populations. For example, in the mouse and human intestine, stem cells harboring oncogenic mutations preferentially replace their wild type neighbors (Nicholson et al., 2018; Snippert et al., 2014; Vermeulen et al., 2013). However, in tissues such as the hair follicle, cell competition can also function as a tumor suppressive mechanism, with wild type cells suppressing the outgrowth of clones harboring pre-oncogenic mutations (Brown et al., 2017b; Pineda et al., 2019). Broadly, these studies have revealed a wide variety of tissues and mutational insults in which cellular contribution depends on the relative fitness of other cells within the tissue.

Additional studies have added a further layer of complexity to a cell's functional contribution relative to its neighbors, by demonstrating that the same mutations can impose both winner and loser status, depending on the physiological state of the organism. For example, during development, cells with high activity of the transcriptional regulator YAP are winners, whereas adult cells overexpressing YAP are losers (Chiba et al., 2016; Hashimoto and Sasaki, 2019). Additionally, environmental factors, such as nutritional intake, can also reassign winner and loser status. For example, several epithelia have been shown to eliminate cells harboring ras^{v12} mutations when combined with wild type cells (Kon et al., 2017). However, when mice are fed a high-fat diet, elimination of ras^{v12} cells from the intestine and pancreas is suppressed, resulting in tumor-like masses (Sasaki et al., 2018). Thus, a cell's functional contribution varies according to its fitness relative to its neighbors, but this, in turn, varies in diverse physiological contexts.

Together, these studies are revealing that cell identity is highly sensitive to the context in which it is assessed. In addition to the genetic, disease, and damage contexts highlighted here, extensive work is examining additional factors that modulate cellular features and functions, such as age (reviewed in (Keyes and Fuchs, 2018) (Figure 4G–I), nutrient availability/utilization (reviewed in (Intlekofer and Finley, 2019; Mihaylova et al., 2014)), microbes (reviewed in (Larsen et al., 2020)), sex (as in the fly, (Hudry et al., 2019; Hudry et al., 2016)), and beyond (Figure 4L). Moreover, it is emerging that past experiences, such as inflammation, can be stored as epigenetic memories that influence future cellular functions (Naik et al., 2017). These studies are underscoring key limitations to extrapolating cellular functions across conditions, and opening up a wealth of new biology regarding how cells coordinate and compete to maintain and regenerate organs over the lifespan.

Conclusions and outlook

The rapid advances of diverse technologies are dramatically expanding the dimensions along which cell identity can be defined. These studies are revealing the cellular composition of tissues with increasing resolution, and also pointing toward strategies to coax cells to adopt features and functions of therapeutic value. As the number of molecular fingerprints assigned to cell types grows, so too does the number of different contexts in which cell

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

function is being assessed, across diverse injury and disease models and genetic combinations. Together, our rapidly expanding capability to detect both features and functions are revealing that "cell types" that were perceived as monolithic and stable, in fact represent composites of multiple cells with distinguishable molecular signatures and have the capability to adopt new features and functions in new contexts.

From the perspective of identifying cells based on molecular features, the numerous potential modalities for cell profiling present opportunities and challenges. These approaches are generating more complete perspectives on cellular designations and providing fundamental insights into the causal mechanisms that drive phenotypes. Major challenges for ongoing work are to integrate the molecular profiles defined through these numerous different technologies, and to understand why different modalities may not converge to identify the same delineations between cells. Finally, a key goal is to understand how subdivisions of molecular fingerprints between cells translate to functional consequences.

Functional assessments of cell types in diverse contexts are continuing to reveal that cells have the capability to adopt features and functions beyond those observed in homeostasis, and these can be uncovered by experimental and pathological perturbations. These observations suggest that, in contrast to a model connecting one cell type to one function, cell types might, instead, be more accurately described as a suite of potential functions that can be unleashed in specific contexts. From this perspective, and given the role that plasticity can play in disease progression, an intriguing question is what sets the boundaries on those functions. How is plasticity constrained in healthy tissue? How is plasticity reverted after damage is resolved? Experiments demonstrating how microenvironmental changes can shift cellular features and functions continue to raise important questions on the extent to which cell function is hardwired, calling to mind Sydney Brenner's evocative description of cell specification based on ancestry (the so-called European plan) or based on neighbors (which he called the American plan) (referenced in (Fraser and Harland, 2000). In this regard, the wealth of techniques described here that preserve spatial context while performing high-dimensional profiling lay exceptional groundwork to dissect the interplay between intrinsic and extrinsic features driving cellular functions.

Acknowledgments

We thank Rachel Zwick, Todd Nystul, Taylor Skokan, Chip Ferguson, Bob Horvitz, and members of the Klein and Vale laboratories for helpful feedback and discussions. We apologize to those authors whose work we could not cite due to space constraints. Research in the Klein laboratory is supported by NIH R35-DE026602 and Tobacco-Related Disease Research Program (University of California) 587629. KLM was supported by the Damon Runyon Cancer Research Foundation (DRG-2282-17) and by the Eunice Kennedy Shriver National Institute Of Child Health & Human Development of the National Institutes of Health (K99HD101021).

References

Adams GB, Alley IR, Chung UI, Chabner KT, Jeanson NT, Lo Celso C, Marsters ES, Chen M, Weinstein LS, Lin CP, et al. (2009). Haematopoietic stem cells depend on Galpha(s)-mediated signaNing to engraft bone marrow. Nature 459, 103–107. [PubMed: 19322176]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Aiuti A, Biasco L, Scaramuzza S, Ferrua F, Cicalese MP, Baricordi C, Dionisio F, Calabria A, Giannelli S, Castiello MC, et al. (2013). Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. Science 341, 1233151. [PubMed: 23845947]

Alemany A, Florescu M, Baron CS, Peterson-Maduro J, and van Oudenaarden A (2018). Whole-organism clone tracing using single-cell sequencing. Nature 556, 108–112. [PubMed: 29590089]

Algire GH, and Legallais FY (1949). Recent developments in the transparent-chamber technique as adapted to the mouse. J Natl Cancer Inst 10, 225–253, incl 228 pl. [PubMed: 15393709]

Alieva M, Ritsma L, Giedt RJ, Weissleder R, and van Rheenen J (2014). Imaging windows for long-term intravital imaging: General overview and technical insights. Intravital 3, e29917. [PubMed: 28243510]

Amat F, Lemon W, Mossing DP, McDole K, Wan Y, Branson K, Myers EW, and Keller PJ (2014). Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data. Nat Methods 11, 951–958. [PubMed: 25042785]

Askary A, Sanchez-Guardado L, Linton JM, Chadly DM, Budde MW, Cai L, Lois C, and Elowitz MB (2019). In situ readout of DNA barcodes and single base edits facilitated by in vitro transcription. Nat Biotechnol.

Baccin C, Al-Sabah J, Velten L, Helbling PM, Grunschlager F, Hernandez-Malmierca P, Nombela-Arrieta C, Steinmetz LM, Trumpp A, and Haas S (2019). Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization. Nat Cell Biol.

Bao Z, Murray JI, Boyle T, Ooi SL, Sandel MJ, and Waterston RH (2006). Automated cell lineage tracing in Caenorhabditis elegans. Proc Natl Acad Sci U S A 103, 2707–2712. [PubMed: 16477039]

Baron CS, and van Oudenaarden A (2019). Unravelling cellular relationships during development and regeneration using genetic lineage tracing. Nat Rev Mol Cell Biol 20, 753–765. [PubMed: 31690888]

Battich N, Beumer J, de Barbanson B, Krenning L, Baron CS, Tanenbaum ME, Clevers H, and van Oudenaarden A (2020). Sequencing metabolically labeled transcripts in single cells reveals mRNA turnover strategies. Science 367, 1151–1156. [PubMed: 32139547]

Biddy BA, Kong W, Kamimoto K, Guo C, Waye SE, Sun T, and Morris SA (2018). Single-cell mapping of lineage and identity in direct reprogramming. Nature 564, 219–224. [PubMed: 30518857]

Biffi A, Montini E, Lorioli L, Cesani M, Fumagalli F, Plati T, Baldoli C, Martino S, Calabria A, Canale S, et al. (2013). Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. Science 341, 1233158. [PubMed: 23845948]

Boisset JC, Vivie J, Grun D, Muraro MJ, Lyubimova A, and van Oudenaarden A (2018). Mapping the physical network of cellular interactions. Nat Methods 15, 547–553. [PubMed: 29786092]

Bondar T, and Medzhitov R (2010). p53-mediated hematopoietic stem and progenitor cell competition. Cell Stem Cell 6, 309–322. [PubMed: 20362536]

Boothe T, Hilbert L, Heide M, Berninger L, Huttner WB, Zaburdaev V, Vastenhouw NL, Myers EW, Drechsel DN, and Rink JC (2017). A tunable refractive index matching medium for live imaging cells, tissues and model organisms. Elife 6.

Bowling S, Sritharan D, Osorio FG, Nguyen M, Cheung P, Rodriguez-Fraticelli A, Patel S, Fujiwara Y, Li BE, Orkin SH, et al. (2019). An engineered CRISPR/Cas9 mouse line for simultaneous readout of lineage histories and gene expression profiles in single cells. bioRxiv, 797597.

Brown EB, Campbell RB, Tsuzuki Y, Xu L, Carmeliet P, Fukumura D, and Jain RK (2001). In vivo measurement of gene expression, angiogenesis and physiological function in tumors using multiphoton laser scanning microscopy. Nat Med 7, 864–868. [PubMed: 11433354]

Brown K, Loh KM, and Nusse R (2017a). Live Imaging Reveals that the First Division of Differentiating Human Embryonic Stem Cells Often Yields Asymmetric Fates. Cell Rep 21, 301–307. [PubMed: 29020617]

Brown S, Pineda CM, Xin T, Boucher J, Suozzi KC, Park S, Matte-Martone C, Gonzalez DG, Rytlewski J, Beronja S, et al. (2017b). Correction of aberrant growth preserves tissue homeostasis. Nature 548, 334–337. [PubMed: 28783732]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Buenrostro JD, Corces MR, Lareau CA, Wu B, Schep AN, Aryee MJ, Majeti R, Chang HY, and Greenleaf WJ (2018). Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of Human Hematopoietic Differentiation. Cell 173, 1535–1548 e1516. [PubMed: 29706549]

Cao J, Packer JS, Ramani V, Cusanovich DA, Huynh C, Daza R, Qiu X, Lee C, Furlan SN, Steemers FJ, et al. (2017). Comprehensive single-cell transcriptional profiling of a multicellular organism. Science 357, 661–667. [PubMed: 28818938]

Cao J, Zhou W, Steemers F, Trapnell C, and Shendure J (2019). Characterizing the temporal dynamics of gene expression in single cells with sci-fate. bioRxiv, 666081.

Castillo-Azofeifa D, Fazio EN, Nattiv R, Good HJ, Wald T, Pest MA, de Sauvage FJ, Klein OD, and Asfaha S (2019). Atoh1(+) secretory progenitors possess renewal capacity independent of Lgr5(+) cells during colonic regeneration. EMBO J 38.

Chan MM, Smith ZD, Grosswendt S, Kretzmer H, Norman TM, Adamson B, Jost M, Quinn JJ, Yang D, Jones MG, et al. (2019). Molecular recording of mammalian embryogenesis. Nature 570, 77–82. [PubMed: 31086336]

Chen KH, Boettiger AN, Moffitt JR, Wang S, and Zhuang X (2015). RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. Science 348, aaa6090. [PubMed: 25858977]

Chen X, Miragaia RJ, Natarajan KN, and Teichmann SA (2018). A rapid and robust method for single cell chromatin accessibility profiling. Nat Commun 9, 5345. [PubMed: 30559361]

Chen X, Sun YC, Zhan H, Kebschull JM, Fischer S, Matho K, Huang ZJ, Gillis J, and Zador AM (2019). High-Throughput Mapping of Long-Range Neuronal Projection Using In Situ Sequencing. Cell 179, 772–786 e719. [PubMed: 31626774]

Chow K-HK, Budde MW, Granados AA, Cabrera M, Yoon S, Cho S, Huang T. h., Koulena N, Frieda KL, Cai L, et al. (2020). Imaging cell lineage with a synthetic digital recording system. bioRxiv.

Christodoulou C, Spencer JA, Yeh SA, Turcotte R, Kokkaliaris KD, Panero R, Ramos A, Guo G, Seyedhassantehrani N, Esipova TV, et al. (2020). Live-animal imaging of native haematopoietic stem and progenitor cells. Nature 578, 278–283. [PubMed: 32025033]

Clavería C, Giovinazzo G, Sierra R, and Torres M (2013). Myc-driven endogenous cell competition in the early mammalian embryo. In Nature, pp. 39–44.

Clevers H, Rafelski S, Elowitz M, Klein A, Shendure J, Trapnell C, Lein E, Lundberg E, Uhlen M, Martinez-Arias A et al. (2017). What Is Your Conceptual Definition of “Cell Type” in the Context of a Mature Organism? Cell Syst 4, 255–259. [PubMed: 28334573]

Colmone A, Amorim M, Pontier AL, Wang S, Jablonski E, and Sipkins DA (2008). Leukemic cells create bone marrow niches that disrupt the behavior of normal hematopoietic progenitor cells. Science 322, 1861–1865. [PubMed: 19095944]

Costa MR, Ortega F, Brill MS, Beckervordersandforth R, Petrone C, Schroeder T, Gotz M, and Berninger B (2011). Continuous live imaging of adult neural stem cell division and lineage progression in vitro. Development 138, 1057–1068. [PubMed: 21343361]

Currie JD, Kawaguchi A, Traspas RM, Schuez M, Chara O, and Tanaka EM (2016). Live Imaging of Axolotl Digit Regeneration Reveals Spatiotemporal Choreography of Diverse Connective Tissue Progenitor Pools. Dev Cell 39, 411–423. [PubMed: 27840105]

Cusanovich DA, Hill AJ, Aghamirzaie D, Daza RM, Pliner HA, Berletch JB, Filippova GN, Huang X, Christiansen L, DeWitt WS, et al. (2018a). A Single-Cell Atlas of In Vivo Mammalian Chromatin Accessibility. Cell 174, 1309–1324 e1318. [PubMed: 30078704]

Cusanovich DA, Reddington JP, Garfield DA, Daza RM, Aghamirzaie D, Marco-Ferreres R, Pliner HA, Christiansen L, Qiu X, Steemers FJ, et al. (2018b). The cis-regulatory dynamics of embryonic development at single-cell resolution. Nature 555, 538–542. [PubMed: 29539636]

Datlinger P, Rendeiro AF, Boenke T, Krausgruber T, Barreca D, and Bock C (2019). Ultra-high throughput single-cell RNA sequencing by combinatorial fluidic indexing. bioRxiv, 2019.2012.2017.879304.

de Navascues J, Perdigoto CN, Bian Y, Schneider MH, Bardin AJ, Martinez-Arias A, and Simons BD (2012). Drosophila midgut homeostasis involves neutral competition between symmetrically dividing intestinal stem cells. EMBO J 31, 2473–2485. [PubMed: 22522699]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

de Sousa e Melo F, Kurtova AV, Harnoss JM, Kljavin N, Hoeck JD, Hung J, Anderson JE, Storm EE, Modrusan Z, Koeppen H, et al. (2017). A distinct role for Lgr5(+) stem cells in primary and metastatic colon cancer. Nature 543, 676–680. [PubMed: 28358093]

Díaz-Díaz C, de Manuel LF, Jimenez-Carretero D, Montoya MC, Clavería C, and Torres M (2017). Pluripotency Surveillance by Myc-Driven Competitive Elimination of Differentiating Cells. In Developmental Cell (Elsevier Inc.), pp. 585–599.e584.

Dick JE, Magli MC, Huszar D, Phillips RA, and Bernstein A (1985). Introduction of a selectable gene into primitive stem cells capable of long-term reconstitution of the hemopoietic system of W/Wv mice. Cell 42, 71–79. [PubMed: 4016956]

Du Z, Santella A, He F, Tiongson M, and Bao Z (2014). De novo inference of systems-level mechanistic models of development from live-imaging-based phenotype analysis. Cell 156, 359–372. [PubMed: 24439388]

Egen JG, Rothfuchs AG, Feng CG, Winter N, Sher A, and Germain RN (2008). Macrophage and T cell dynamics during the development and disintegration of mycobacterial granulomas. Immunity 28, 271–284. [PubMed: 18261937]

Eilken HM, Nishikawa S, and Schroeder T (2009). Continuous single-cell imaging of blood generation from haemogenic endothelium. Nature 457, 896–900. [PubMed: 19212410]

Ellis SJ, Gomez NC, Levorse J, Mertz AF, Ge Y, and Fuchs E (2019). Distinct modes of cell competition shape mammalian tissue morphogenesis. In Nature, pp. 497–502.

Erhard F, Baptista MAP, Krammer T, Hennig T, Lange M, Arampatzi P, Jurges CS, Theis FJ, Saliba AE, and Dolken L (2019). scSLAM-seq reveals core features of transcription dynamics in single cells. Nature 571, 419–423. [PubMed: 31292545]

Ewald AJ, Werb Z, and Egeblad M (2011). Preparation of mice for long-term intravital imaging of the mammary gland. Cold Spring Harb Protoc 2011, pdb prot5562.

Faure E, Savy T, Rizzi B, Melani C, Stasova O, Fabreges D, Spir R, Hammons M, Cunderlik R, Recher G, et al. (2016). A workflow to process 3D+time microscopy images of developing organisms and reconstruct their cell lineage. Nat Commun 7, 8674. [PubMed: 26912388]

Ford CE, Hamerton JL, Barnes DW, and Loutit JF (1956). Cytological identification of radiation-chimaeras. Nature 177, 452–454. [PubMed: 13309336]

Fraser SE, and Harland RM (2000). The molecular metamorphosis of experimental embryology. Cell 100, 41–55. [PubMed: 10647930]

Fumagalli A, Oost KC, Kester L, Morgner J, Bornes L, Bruens L, Spaargaren L, Azkanaz M, Schelfhorst T, Beerling E, et al. (2020). Plasticity of Lgr5-Negative Cancer Cells Drives Metastasis in Colorectal Cancer. Cell Stem Cell.

Gall JG, and Pardue ML (1969). Formation and detection of RNA-DNA hybrid molecules in cytological preparations. Proc Natl Acad Sci U S A 63, 378–383. [PubMed: 4895535]

Gehring J, Hwee Park J, Chen S, Thomson M, and Pachter L (2019). Highly multiplexed single-cell RNA-seq by DNA oligonucleotide tagging of cellular proteins. Nat Biotechnol.

Gerdes MJ, Sevinsky CJ, Sood A, Adak S, Bello MO, Bordwell A, Can A, Corwin A, Dinn S, Filkins RJ, et al. (2013). Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. Proc Natl Acad Sci U S A 110, 11982–11987. [PubMed: 23818604]

Ghigo C, Mondor I, Jorquera A, Nowak J, Wienert S, Zahner SP, Clausen BE, Luche H, Malissen B, Klauschen F, et al. (2013). Multicolor fate mapping of Langerhans cell homeostasis. J Exp Med 210, 1657–1664. [PubMed: 23940255]

Giesen C, Wang HA, Schapiro D, Zivanovic N, Jacobs A, Hattendorf B, Schuffler PJ, Grolimund D, Buhmann JM, Brandt S, et al. (2014). Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. Nat Methods 11, 417–422. [PubMed: 24584193]

Giladi A, Cohen M, Medaglia C, Baran Y, Li B, Zada M, Bost P, Blecher-Gonen R, Salame TM, Mayer JU, et al. (2020). Dissecting cellular crosstalk by sequencing physically interacting cells. Nat Biotechnol.

Goltsev Y, Samusik N, Kennedy-Darling J, Bhate S, Hale M, Vazquez G, Black S, and Nolan GP (2018). Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. Cell 174, 968–981 e915. [PubMed: 30078711]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Grindberg RV, Yee-Greenbaum JL, McConnell MJ, Novotny M, O'Shaughnessy AL, Lambert GM, Arauzo-Bravo MJ, Lee J, Fishman M, Robbins GE, et al. (2013). RNA-sequencing from single nuclei. Proc Natl Acad Sci U S A 110, 19802–19807. [PubMed: 24248345]

Grosselin K, Durand A, Marsolier J, Poitou A, Marangoni E, Nemati F, Dahmani A, Lameiras S, Reyal F, Frenoy O, et al. (2019). High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. Nat Genet 51, 1060–1066. [PubMed: 31152164]

Guenechea G, Gan OI, Dorrell C, and Dick JE (2001). Distinct classes of human stem cells that differ in proliferative and self-renewal potential. Nat Immunol 2, 75–82. [PubMed: 11135582]

Habib N, Li Y, Heidenreich M, Swiech L, Avraham-Davidi I, Trombetta JJ, Hession C, Zhang F, and Regev A (2016). Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. Science 353, 925–928. [PubMed: 27471252]

Habib SJ, Chen BC, Tsai FC, Anastassiadis K, Meyer T, Betzig E, and Nusse R (2013). A localized Wnt signal orients asymmetric stem cell division in vitro. Science 339, 1445–1448. [PubMed: 23520113]

Halpern KB, Shenhav R, Massalha H, Toth B, Egozi A, Massasa EE, Medgalia C, David E, Giladi A, Moor AE, et al. (2018). Paired-cell sequencing enables spatial gene expression mapping of liver endothelial cells. Nat Biotechnol 36, 962–970. [PubMed: 30222169]

Hartmann FJ, Mrdjen D, McCaffrey E, Glass DR, Greenwald NF, Bharadwaj A, Khair Z, Baranski A, Baskar R, Angelo M, et al. (2020). Multiplexed Single-cell Metabolic Profiles Organize the Spectrum of Human Cytotoxic T Cells. bioRxiv, 2020.2001.2017.909796.

Hendriks GJ, Jung LA, Larsson AJM, Lidschreiber M, Andersson Forsman O, Lidschreiber K, Cramer P, and Sandberg R (2019). NASC-seq monitors RNA synthesis in single cells. Nat Commun 10, 3138. [PubMed: 31316066]

Hisha H, Tanaka T, Kanno S, Tokuyama Y, Komai Y, Ohe S, Yanai H, Omachi T, and Ueno H (2013). Establishment of a novel lingual organoid culture system: generation of organoids having mature keratinized epithelium from adult epithelial stem cells. Sci Rep 3, 3224. [PubMed: 24232854]

Hsu YC (2015). Theory and Practice of Lineage Tracing. Stem Cells 33, 3197–3204. [PubMed: 26284340]

Hu KH, Eichorst JP, McGinnis CS, Patterson DM, Chow ED, Kersten K, Jameson SC, Gartner ZJ, Rao AA, and Krummel MF (2020). ZipSeq : Barcoding for Real-time Mapping of Single Cell Transcriptomes. bioRxiv.

Hudry B, de Goeij E, Mineo A, Gaspar P, Hadjieconomou D, Studd C, Mokochinski JB, Kramer HB, Placais PY, Preat T, et al. (2019). Sex Differences in Intestinal Carbohydrate Metabolism Promote Food Intake and Sperm Maturation. Cell 178, 901–918 e916. [PubMed: 31398343]

Hudry B, Khadayate S, and Miguel-Aliaga I (2016). The sexual identity of adult intestinal stem cells controls organ size and plasticity. Nature 530, 344–348. [PubMed: 26887495]

Inoue F, Kreimer A, Ashuach T, Ahituv N, and Yosef N (2019). Identification and Massively Parallel Characterization of Regulatory Elements Driving Neural Induction. Cell Stem Cell 25, 713–727 e710. [PubMed: 31631012]

Intlekofer AM, and Finley LWS (2019). Metabolic signatures of cancer cells and stem cells. Nat Metab 1, 177–188. [PubMed: 31245788]

Kang HM, Subramaniam M, Targ S, Nguyen M, Maliskova L, McCarthy E, Wan E, Wong S, Byrnes L, Lanata CM, et al. (2018). Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. Nat Biotechnol 36, 89–94. [PubMed: 29227470]

Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, and Henikoff S (2019). CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun 10, 1930. [PubMed: 31036827]

Kedrin D, Gligorijevic B, Wyckoff J, Verkhusha VV, Condeelis J, Segall JE, and van Rheenen J (2008). Intravital imaging of metastatic behavior through a mammary imaging window. Nat Methods 5, 1019–1021. [PubMed: 18997781]

Keller G, Paige C, Gilboa E, and Wagner EF (1985). Expression of a foreign gene in myeloid and lymphoid cells derived from multipotent haematopoietic precursors. Nature 318, 149–154. [PubMed: 3903518]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Keyes BE, and Fuchs E (2018). Stem cells: Aging and transcriptional fingerprints. J Cell Biol 217, 79–92. [PubMed: 29070608]

Kimble J (1981). Alterations in cell lineage following laser ablation of cells in the somatic gonad of Caenorhabditis elegans. Dev Biol 87, 286–300. [PubMed: 7286433]

Kishi JY, Lapan SW, Beliveau BJ, West ER, Zhu A, Sasaki HM, Saka SK, Wang Y, Cepko CL, and Yin P (2019). SABER amplifies FISH: enhanced multiplexed imaging of RNA and DNA in cells and tissues. Nat Methods 16, 533–544. [PubMed: 31110282]

Klein AM, Nakagawa T, Ichikawa R, Yoshida S, and Simons BD (2010). Mouse germ line stem cells undergo rapid and stochastic turnover. Cell Stem Cell 7, 214–224. [PubMed: 20682447]

Koenitzer JR, Wu H, Atkinson JJ, Brody SL, and Humphreys BD (2020). Single nucleus RNASeq profiling of mouse lung: reduced dissociation bias and improved detection of rare cell types compared with single cell RNASeq. bioRxiv.

Kretzschmar K, and Watt FM (2012). Lineage tracing. Cell 148, 33–45. [PubMed: 22265400]

Lai B, Gao W, Cui K, Xie W, Tang Q, Jin W, Hu G, Ni B, and Zhao K (2018). Principles of nucleosome organization revealed by single-cell micrococcal nuclease sequencing. Nature 562, 281–285. [PubMed: 30258225]

Lake BB, Ai R, Kaeser GE, Salathia NS, Yung YC, Liu R, Wildberg A, Gao D, Fung HL, Chen S, et al. (2016). Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. Science 352, 1586–1590. [PubMed: 27339989]

Larsen SB, Cowley CJ, and Fuchs E (2020). Epithelial cells: liaisons of immunity. Current Opinion in Immunology 62, 45–53. [PubMed: 31874430]

Lederer AR, and La Manno G (2020). The emergence and promise of single-cell temporal-omics approaches. Curr Opin Biotechnol 63, 70–78. [PubMed: 31918114]

Lin JR, Fallahi-Sichani M, and Sorger PK (2015). Highly multiplexed imaging of single cells using a high-throughput cyclic immunofluorescence method. Nat Commun 6, 8390. [PubMed: 26399630]

Livet J, Weissman TA, Kang H, Draft RW, Lu J, Bennis RA, Sanes JR, and Lichtman JW (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature 450, 56–62. [PubMed: 17972876]

Lo Celso C, Fleming HE, Wu JW, Zhao CX, Miake-Lye S, Fujisaki J, Cote D, Rowe DW, Lin CP, and Scadden DT (2009). Live-animal tracking of individual haematopoietic stem/progenitor cells in their niche. Nature 457, 92–96. [PubMed: 19052546]

Lopez-Garcia C, Klein AM, Simons BD, and Winton DJ (2010). Intestinal stem cell replacement follows a pattern of neutral drift. Science 330, 822–825. [PubMed: 20929733]

Lu R, Neff NF, Quake SR, and Weissman IL (2011). Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. Nat Biotechnol 29, 928–933. [PubMed: 21964413]

Ludwig CH, and Bintu L (2019). Mapping chromatin modifications at the single cell level. Development 146.

Luo C, Keown CL, Kurihara L, Zhou J, He Y, Li J, Castanon R, Lucero J, Nery JR, Sandoval JP, et al. (2017). Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex. Science 357, 600–604. [PubMed: 28798132]

Mace DL, Weisdepp P, Gevirtzman L, Boyle T, and Waterston RH (2013). A high-fidelity cell lineage tracing method for obtaining systematic spatiotemporal gene expression patterns in Caenorhabditis elegans. G3 (Bethesda) 3, 851–863. [PubMed: 23550142]

Marsh E, Gonzalez DG, Lathrop EA, Boucher J, and Greco V (2018). Positional Stability and Membrane Occupancy Define Skin Fibroblast Homeostasis In Vivo. Cell 175, 1620–1633 e1613. [PubMed: 30415836]

Mazo IB, Gutierrez-Ramos JC, Frenette PS, Hynes RO, Wagner DD, and von Andrian UH (1998). Hematopoietic progenitor cell rolling in bone marrow microvessels: parallel contributions by endothelial selectins and vascular cell adhesion molecule 1. J Exp Med 188, 465–474. [PubMed: 9687524]

McDole K, Guignard L, Amat F, Berger A, Malandain G, Royer LA, Turaga SC, Branson K, and Keller PJ (2018). In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level. Cell 175, 859–876 e833. [PubMed: 30318151]

McGinnis CS, Patterson DM, Winkler J, Conrad DN, Hein MY, Srivastava V, Hu JL, Murrow LM, Weissman JS, Werb Z, et al. (2019). MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. Nat Methods 16, 619–626. [PubMed: 31209384]

McKenzie JL, Gan OI, Doedens M, Wang JC, and Dick JE (2006). Individual stem cells with highly variable proliferation and self-renewal properties comprise the human hematopoietic stem cell compartment. Nat Immunol 7, 1225–1233. [PubMed: 17013390]

McKinley KL, Stuurman N, Royer LA, Schartner C, Castillo-Azofeifa D, Delling M, Klein OD, and Vale RD (2018). Cellular aspect ratio and cell division mechanics underlie the patterning of cell progeny in diverse mammalian epithelia. Elife 7.

Medaglia C, Giladi A, Stoler-Barak L, De Giovanni M, Salame TM, Biram A, David E, Li H, Iannacone M, Shulman Z, et al. (2017). Spatial reconstruction of immune niches by combining photoactivatable reporters and scRNA-seq. Science 358, 1622–1626. [PubMed: 29217582]

Merritt CR, Ong GT, Church S, Barker K, Geiss G, Hoang M, Jung J, Liang Y, McKay-Fleisch J, Nguyen K, et al. (2019). High multiplex, digital spatial profiling of proteins and RNA in fixed tissue using genomic detection methods. bioRxiv.

Mesa KR, Kawaguchi K, Cockburn K, Gonzalez D, Boucher J, Xin T, Klein AM, and Greco V (2018). Homeostatic Epidermal Stem Cell Self-Renewal Is Driven by Local Differentiation. In Cell Stem Cell, pp. 677–686.e674. [PubMed: 30269903]

Mihaylova MM, Sabatini DM, and Yilmaz OH (2014). Dietary and metabolic control of stem cell function in physiology and cancer. Cell Stem Cell 14, 292–305. [PubMed: 24607404]

Mooijman D, Dey SS, Boisset JC, 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. [PubMed: 27347753]

Moor AE, Harnik Y, Ben-Moshe S, Massasa EE, Rozenberg M, Eilam R, Bahar Halpern K, and Itzkovitz S (2018). Spatial Reconstruction of Single Enterocytes Uncovers Broad Zonation along the Intestinal Villus Axis. Cell 175, 1156–1167 e1115. [PubMed: 30270040]

Mulqueen RM, Pokholok D, Norberg SJ, Torkenczy KA, Fields AJ, Sun D, Sinnamon JR, Shendure J, Trapnell C, O'Roak BJ, et al. (2018). Highly scalable generation of DNA methylation profiles in single cells. Nat Biotechnol 36, 428–431. [PubMed: 29644997]

Naik S, Larsen SB, Gomez NC, Alaverdyan K, Sendoel A, Yuan S, Polak L, Kulukian A, Chai S, and Fuchs E (2017). Inflammatory memory sensitizes skin epithelial stem cells to tissue damage. Nature 550, 475–480. [PubMed: 29045388]

Nolta JA, Dao MA, Wells S, Smogorzewska EM, and Kohn DB (1996). Transduction of pluripotent human hematopoietic stem cells demonstrated by clonal analysis after engraftment in immune-deficient mice. Proc Natl Acad Sci U S A 93, 2414–2419. [PubMed: 8637888]

Orth JD, Kohler RH, Foijer F, Sorger PK, Weissleder R, and Mitchison TJ (2011). Analysis of mitosis and antimitotic drug responses in tumors by in vivo microscopy and single-cell pharmacodynamics. Cancer Res 71, 4608–4616. [PubMed: 21712408]

Pan YA, Freundlich T, Weissman TA, Schoppik D, Wang XC, Zimmerman S, Ciruna B, Sanes JR, Lichtman JW, and Schier AF (2013). Zebrabow: multispectral cell labeling for cell tracing and lineage analysis in zebrafish. Development 140, 2835–2846. [PubMed: 23757414]

Piltti KM, Cummings BJ, Carta K, Manughian-Peter A, Worne CL, Singh K, Ong D, Maksymyuk Y, Khine M, and Anderson AJ (2018). Live-cell time-lapse imaging and single-cell tracking of in vitro cultured neural stem cells - Tools for analyzing dynamics of cell cycle, migration, and lineage selection. Methods 133, 81–90. [PubMed: 29050826]

Pineda CM, Gonzalez DG, Matte-Martone C, Boucher J, Lathrop E, Gallini S, Fons NR, Xin T, Tai K, Marsh E, et al. (2019). Hair follicle regeneration suppresses Ras-driven oncogenic growth. The Journal of Cell Biology 218, 3212–3222. [PubMed: 31488583]

Pirici D, Mogoanta L, Kumar-Singh S, Pirici I, Margaritescu C, Simionescu C, and Stanescu R (2009). Antibody elution method for multiple immunohistochemistry on primary antibodies raised in the same species and of the same subtype. J Histochem Cytochem 57, 567–575. [PubMed: 19223296]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Porter SN, Baker LC, Mittelman D, and Porteus MH (2014). Lentiviral and targeted cellular barcoding reveals ongoing clonal dynamics of cell lines in vitro and in vivo. Genome Biol 15, R75. [PubMed: 24886633]

Qian X, Goderie SK, Shen Q, Stern JH, and Temple S (1998). Intrinsic programs of patterned cell lineages in isolated vertebrate CNS ventricular zone cells. Development 125, 3143–3152. [PubMed: 9671587]

Qian X, Shen Q, Goderie SK, He W, Capela A, Davis AA, and Temple S (2000). Timing of CNS cell generation: a programmed sequence of neuron and glial cell production from isolated murine cortical stem cells. Neuron 28, 69–80. [PubMed: 11086984]

Qiu Q, Hu P, Govek KW, Camara PG, and Wu H (2019). Massively parallel, time-resolved single-cell RNA sequencing with scNT-Seq. bioRxiv, 2019.2012.2019.882050.

Raj B, Wagner DE, McKenna A, Pandey S, Klein AM, Shendure J, Gagnon JA, and Schier AF (2018). Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. Nat Biotechnol 36, 442–450. [PubMed: 29608178]

Ravin R, Hoeppner DJ, Munno DM, Carmel L, Sullivan J, Levitt DL, Miller JL, Athaide C, Panchision DM, and McKay RD (2008). Potency and fate specification in CNS stem cell populations in vitro. Cell Stem Cell 3, 670–680. [PubMed: 19041783]

Ren W, Lewandowski BC, Watson J, Aihara E, Iwatsuki K, Bachmanov AA, Margolskee RF, and Jiang P (2014). Single Lgr5- or Lgr6-expressing taste stem/progenitor cells generate taste bud cells ex vivo. Proc Natl Acad Sci U S A 111, 16401–16406. [PubMed: 25368147]

Rieger MA, Hoppe PS, Smejkal BM, Eitelhuber AC, and Schroeder T (2009). Hematopoietic cytokines can instruct lineage choice. Science 325, 217–218. [PubMed: 19590005]

Rinkevich Y, Lindau P, Ueno H, Longaker MT, and Weissman IL (2011). Germ-layer and lineage-restricted stem/progenitors regenerate the mouse digit tip. Nature 476, 409–413. [PubMed: 21866153]

Ritsma L, Ellenbroek SIJ, Zomer A, Snippert HJ, de Sauvage FJ, Simons BD, Clevers H, and van Rheenen J (2014). Intestinal crypt homeostasis revealed at single-stem-cell level by in vivo live imaging. Nature 507, 362–365. [PubMed: 24531760]

Ritsma L, Steller EJ, Beerling E, Loomans CJ, Zomer A, Gerlach C, Vrisekoop N, Seinstra D, van Gurp L, Schafer R, et al. (2012). Intravital microscopy through an abdominal imaging window reveals a pre-micrometastasis stage during liver metastasis. Sci Transl Med 4, 158ra145.

Rodriguez-Fraticelli AE, Wolock SL, Weinreb CS, Panero R, Patel SH, Jankovic M, Sun J, Calogero RA, Klein AM, and Camargo FD (2018). Clonal analysis of lineage fate in native haematopoiesis. Nature 553, 212–216. [PubMed: 29323290]

Rompolas P, Deschene ER, Zito G, Gonzalez DG, Saotome I, Haberman AM, and Greco V (2012). Live imaging of stem cell and progeny behaviour in physiological hair-follicle regeneration. Nature 487, 496–499. [PubMed: 22763436]

Rompolas P, Mesa KR, and Greco V (2013). Spatial organization within a niche as a determinant of stem-cell fate. Nature 502, 513–518. [PubMed: 24097351]

Rompolas P, Mesa KR, Kawaguchi K, Park S, Gonzalez D, Brown S, Boucher J, Klein AM, and Greco V (2016). Spatiotemporal coordination of stem cell commitment during epidermal homeostasis. Science 352, 1471–1474. [PubMed: 27229141]

Rosenberg AB, Roco CM, Muscat RA, Kuchina A, Sample P, Yao Z, Graybuck LT, Peeler DJ, 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. [PubMed: 29545511]

Rotem A, Ram O, Shoresh N, Sperling RA, Goren A, Weitz DA, and Bernstein BE (2015). Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. Nat Biotechnol 33, 1165–1172. [PubMed: 26458175]

Saka SK, Wang Y, Kishi JY, Zhu A, Zeng Y, Xie W, Kirli K, Yapp C, Cicconet M, Beliveau BJ, et al. (2019). Immuno-SABER enables highly multiplexed and amplified protein imaging in tissues. Nat Biotechnol 37, 1080–1090. [PubMed: 31427819]

Sancho M, Di Gregorio A, George N, Pozzi S, Sánchez JM, Pernaute B, and Rodríguez TA (2013). Competitive Interactions Eliminate Unfit Embryonic Stem Cells at the Onset of Differentiation. Developmental Cell 26, 19–30. [PubMed: 23867226]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Sandison JC (1924). A new method for the microscopic study of living growing tissues by the introduction of a transparent chamber in the rabbit's ear. The Anatomical Record 28, 281–287.

Scheele CL, Hannezo E, Muraro MJ, Zomer A, Langedijk NS, van Oudenaarden A, Simons BD, and van Rheenen J (2017). Identity and dynamics of mammary stem cells during branching morphogenesis. Nature 542, 313–317. [PubMed: 28135720]

Schepers AG, Snippert HJ, Stange DE, van den Born M, van Es JH, van de Wetering M, and Clevers H (2012). Lineage tracing reveals Lgr5+ stem cell activity in mouse intestinal adenomas. Science 337, 730–735. [PubMed: 22855427]

Schepers K, Swart E, van Heijst JW, Gerlach C, Castrucci M, Sie D, Heimerikx M, Velds A, Kerkhoven RM, Arens R, et al. (2008). Dissecting T cell lineage relationships by cellular barcoding. J Exp Med 205, 2309–2318. [PubMed: 18809713]

Schmidt M, Zickler P, Hoffmann G, Haas S, Wissler M, Muessig A, Tisdale JF, Kuramoto K, Andrews RG, Wu T, et al. (2002). Polyclonal long-term repopulating stem cell clones in a primate model. Blood 100, 2737–2743. [PubMed: 12351380]

Serra D, Mayr U, Boni A, Lukonin I, Rempfler M, Challet Meylan L, Stadler MB, Strnad P, Papasaikas P, Vischi D, et al. (2019). Self-organization and symmetry breaking in intestinal organoid development. Nature 569, 66–72. [PubMed: 31019299]

Shema E, Bernstein BE, and Buenrostro JD (2019). Single-cell and single-molecule epigenomics to uncover genome regulation at unprecedented resolution. Nat Genet 51, 19–25. [PubMed: 30559489]

Sheng XR, and Matunis E (2011). Live imaging of the Drosophilaspermatogonial stem cell niche reveals novel mechanisms regulating germline stem cell output. In Development, pp. 3367–3376. [PubMed: 21752931]

Sipkins DA, Wei X, Wu JW, Runnels JM, Cote D, Means TK, Luster AD, Scadden DT, and Lin CP (2005). In vivo imaging of specialized bone marrow endothelial microdomains for tumour engraftment. Nature 435, 969–973. [PubMed: 15959517]

Snippert HJ, van der Flier LG, Sato T, van Es JH, van den Born M, Kroon-Veenboer C, Barker N, Klein AM, van Rheenen J, Simons BD, et al. (2010). Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. Cell 143, 134–144. [PubMed: 20887898]

Sobolik T, Su YJ, Ashby W, Schaffer DK, Wells S, Wikswo JP, Zijlstra A, and Richmond A (2016). Development of novel murine mammary imaging windows to examine wound healing effects on leukocyte trafficking in mammary tumors with intravital imaging. Intravital 5, e1125562. [PubMed: 28243517]

Spanjaard B, Hu B, Mitic N, Olivares-Chauvet P, Janjuha S, Ninov N, and Junker JP (2018). Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. Nat Biotechnol 36, 469–473. [PubMed: 29644996]

Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, Satija R, and Smibert P (2017). Simultaneous epitope and transcriptome measurement in single cells. Nat Methods 14, 865–868. [PubMed: 28759029]

Stoeckius M, Zheng S, Houck-Loomis B, Hao S, Yeung BZ, Mauck WM 3rd, Smibert P, and Satija R (2018). Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. Genome Biol 19, 224. [PubMed: 30567574]

Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, Hao Y, Stoeckius M, Smibert P, and Satija R (2019). Comprehensive Integration of Single-Cell Data. Cell 177, 1888–1902 e1821. [PubMed: 31178118]

Sulston JE, and White JG (1980). Regulation and cell autonomy during postembryonic development of Caenorhabditis elegans. Dev Biol 78, 577–597. [PubMed: 7190941]

Sun J, Ramos A, Chapman B, Johnnidis JB, Le L, Ho YJ, Klein A, Hofmann O, and Camargo FD (2014). Clonal dynamics of native haematopoiesis. Nature 514, 322–327. [PubMed: 25296256]

Swirski FK, Nahrendorf M, Etzrodt M, Wildgruber M, Cortez-Retamozo V, Panizzi P, Figueiredo JL, Kohler RH, Chudnovskiy A, Waterman P, et al. (2009). Identification of splenic reservoir monocytes and their deployment to inflammatory sites. Science 325, 612–616. [PubMed: 19644120]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Ulman V, Maska M, Magnusson KEG, Ronneberger O, Haubold C, Harder N, Matula P, Matula P, Svoboda D, Radojevic M, et al. (2017). An objective comparison of cell-tracking algorithms. Nat Methods 14, 1141–1152. [PubMed: 29083403]

Valm AM, Cohen S, Legant WR, Melunis J, Hershberg U, Wait E, Cohen AR, Davidson MW, Betzig E, and Lippincott-Schwartz J (2017). Applying systems-level spectral imaging and analysis to reveal the organelle interactome. Nature 546, 162–167. [PubMed: 28538724]

van Es JH, Sato T, van de Wetering M, Lyubimova A, Yee Nee AN, Gregorieff A, Sasaki N, Zeinstra L, van den Born M, Korving J, et al. (2012). Dll1+ secretory progenitor cells revert to stem cells upon crypt damage. Nat Cell Biol 14, 1099–1104. [PubMed: 23000963]

van Es JH, Wiebrands K, Lopez-Iglesias C, van de Wetering M, Zeinstra L, van den Born M, Korving J, Sasaki N, Peters PJ, van Oudenaarden A, et al. (2019). Enteroendocrine and tuft cells support Lgr5 stem cells on Paneth cell depletion. Proc Natl Acad Sci U S A.

Wagner DE, Weinreb C, Collins ZM, Briggs JA, Megason SG, and Klein AM (2018). Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. Science 360, 981–987. [PubMed: 29700229]

Wan Y, Wei Z, Looger LL, Koyama M, Druckmann S, and Keller PJ (2019). Single-Cell Reconstruction of Emerging Population Activity in an Entire Developing Circuit. Cell 179, 355–372 e323. [PubMed: 31564455]

Weinreb C, Rodriguez-Fraticelli A, Camargo FD, and Klein AM (2020). Lineage tracing on transcriptional landscapes links state to fate during differentiation. Science, eaaw3381. [PubMed: 31974159]

Welch JD, Kozareva V, Ferreira A, Vanderburg C, Martin C, and Macosko EZ (2019). Single-Cell Multi-omic Integration Compares and Contrasts Features of Brain Cell Identity. Cell 177, 1873–1887 e1817. [PubMed: 31178122]

Winter MR, Liu M, Monteleone D, Melunis J, Hershberg U, Goderie SK, Temple S, and Cohen AR (2015). Computational Image Analysis Reveals Intrinsic Multigenerational Differences between Anterior and Posterior Cerebral Cortex Neural Progenitor Cells. Stem Cell Reports 5, 609–620. [PubMed: 26344906]

Wolff C, Tinevez JY, Pietzsch T, Stamataki E, Harich B, Guignard L, Preibisch S, Shorte S, Keller PJ, Tomancak P, et al. (2018). Multi-view light-sheet imaging and tracking with the MaMuT software reveals the cell lineage of a direct developing arthropod limb. Elife 7.

Wu AM, Till JE, Siminovitch L, and McCulloch EA (1968). Cytological evidence for a relationship between normal hemotopoietic colony-forming cells and cells of the lymphoid system. J Exp Med 127, 455–464. [PubMed: 5636553]

Wu YE, Pan L, Zuo Y, Li X, and Hong W (2017). Detecting Activated Cell Populations Using Single-Cell RNA-Seq. Neuron 96, 313–329 e316. [PubMed: 29024657]

Xia C, Fan J, Emanuel G, Hao J, and Zhuang X (2019). Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression. Proc Natl Acad Sci U S A.

Yao Z, Mich JK, Ku S, Menon V, Krostag AR, Martinez RA, Furchtgott L, Mulholland H, Bort S, Fuqua MA, et al. (2017). A Single-Cell Roadmap of Lineage Bifurcation in Human ESC Models of Embryonic Brain Development. Cell Stem Cell 20, 120–134. [PubMed: 28094016]

Yu KR, Espinoza DA, Wu C, Truitt L, Shin TH, Chen S, Fan X, Yabe IM, Panch S, Hong SG, et al. (2018). The impact of aging on primate hematopoiesis as interrogated by clonal tracking. Blood 131, 1195–1205. [PubMed: 29295845]

Zechel S, Zajac P, Lonnerberg P, Ibanez CF, and Linnarsson S (2014). Topographical transcriptome mapping of the mouse medial ganglionic eminence by spatially resolved RNA-seq. Genome Biol 15, 486. [PubMed: 25344199]

Zhu C, Preissl S, and Ren B (2020). Single-cell multimodal omics: the power of many. Nature Methods 17, 11–14. [PubMed: 31907462]

Zhu Y, Huang YF, Kek C, and Bulavin DV (2013). Apoptosis differently affects lineage tracing of Lgr5 and Bmi1 intestinal stem cell populations. Cell Stem Cell 12, 298–303. [PubMed: 23415913]

Ziffra RS, Kim CN, Wilfert A, Haeussler M, Casella AM, Przytycki PF, Kreimer A, Pollard KS, Ament SA, Eichler EE, et al. (2019). Single cell epigenomic atlas of the developing human brain and organoids. bioRxiv, 2019.2012.2030.891549.

Zimmermann T (2005). Spectral imaging and linear unmixing in light microscopy. Adv Biochem Eng Biotechnol 95, 245–265. [PubMed: 16080271]

Zong H, Espinosa JS, Su HH, Muzumdar MD, and Luo L (2005). Mosaic analysis with double markers in mice. Cell 121, 479–492. [PubMed: 15882628]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

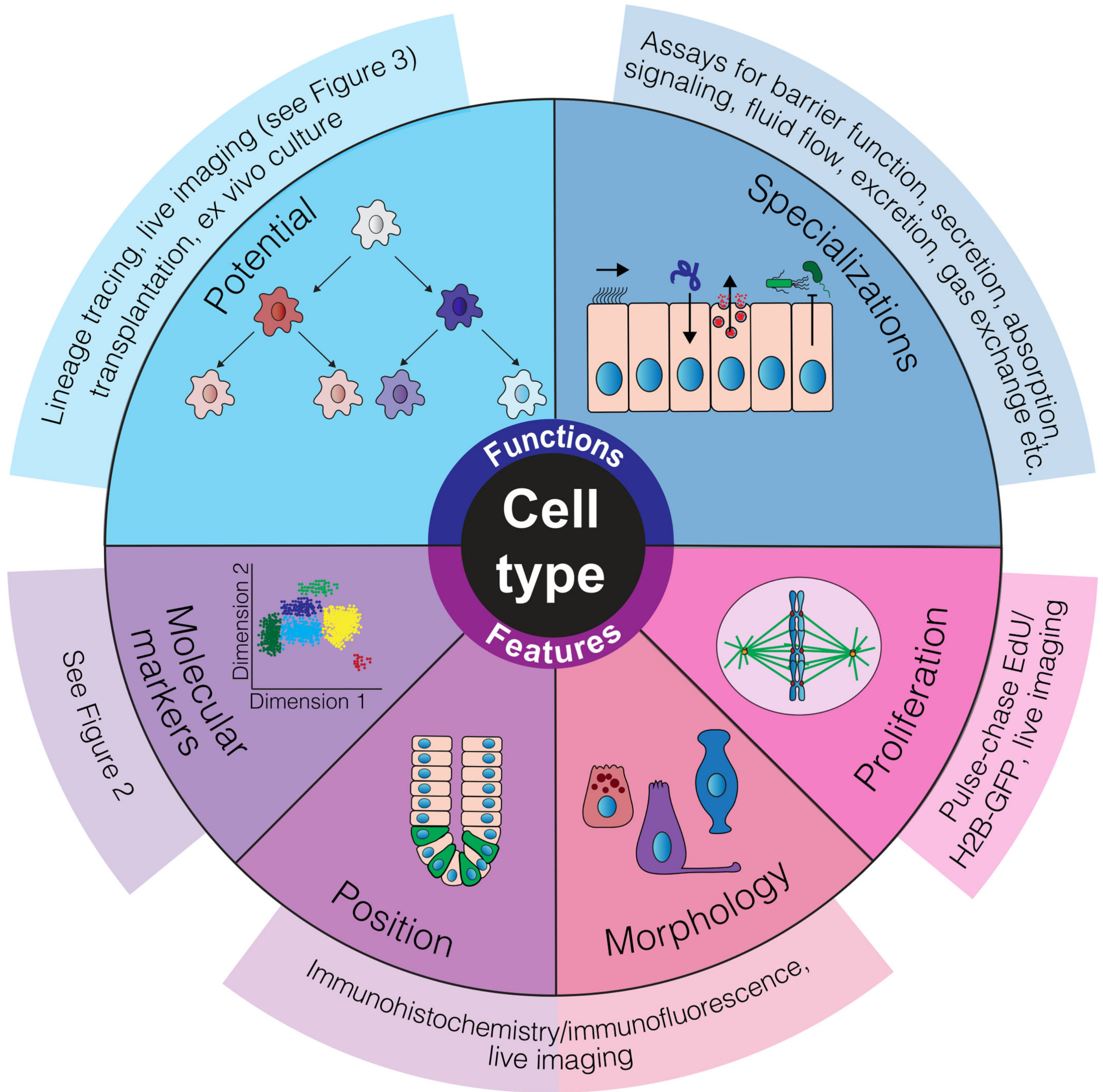

Figure 1. Defining cell types.

A) Attributes used to categorize cells into types (inner wedges) and strategies to assay them (outer wedges). Cells can be classified according to their functions (top), and their physical features (bottom).

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

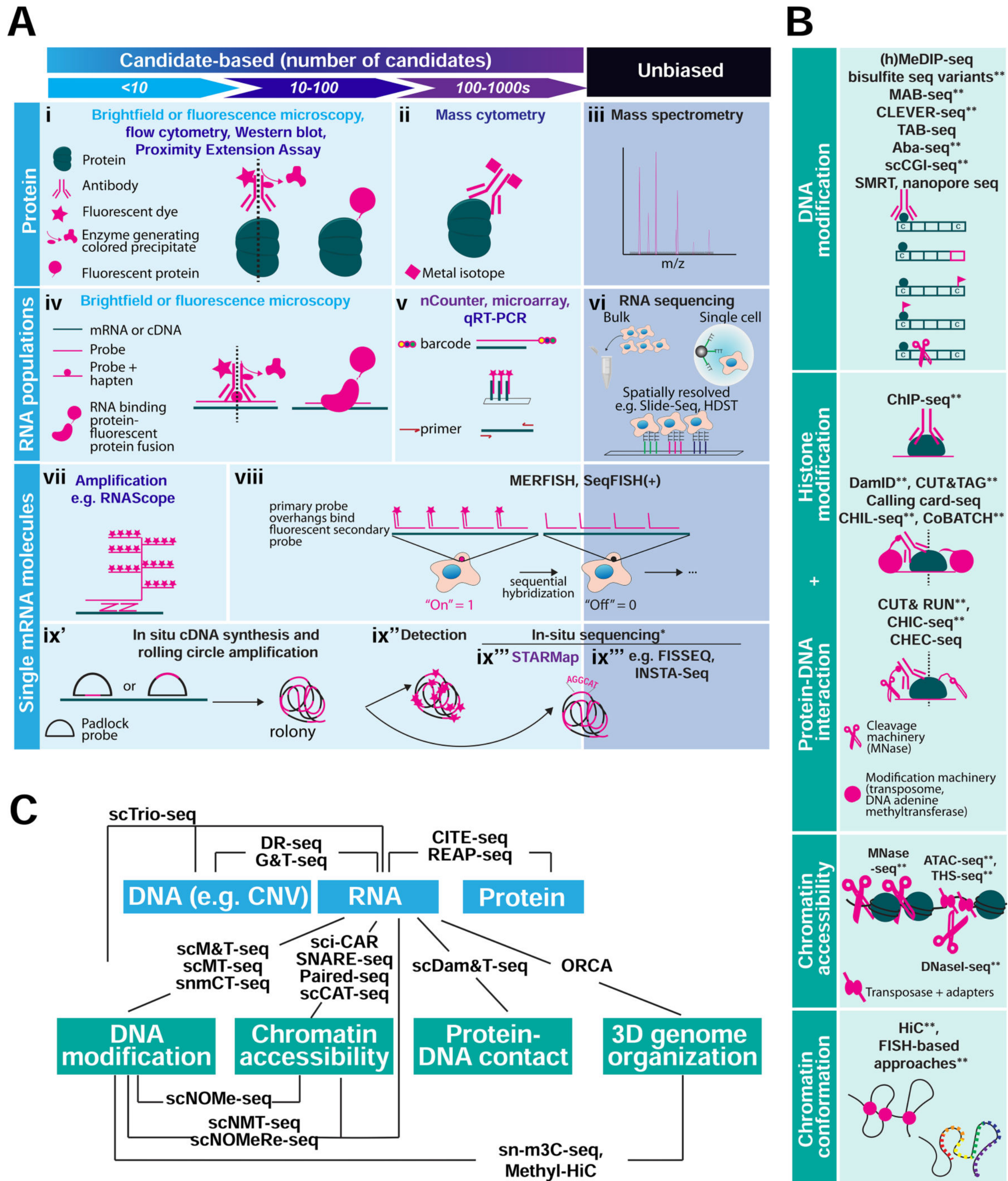

Figure 2. Strategies to detect molecular features associated with a cell type.

A) Common approaches to detect proteins (i-iii) or RNAs (iv-ix) associated with cells. Strategies are grouped according to whether they require up-front selection of candidates (left columns) and/or allow for unbiased profiling of the proteome or transcriptome (right column). The approximate range of features that can be detected by each strategy is indicated by color, corresponding to the blue-to-purple scale. Targets are indicated in teal; detection reagents are indicated in pink. i-iii) Approaches for protein detection. iv) Microscopic approaches for RNA detection. v) Approaches to quantify a panel of mRNAs

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

from populations of cells. vi) Unbiased approaches to detect RNAs. vii-viii) Approaches to detect single mRNA molecules by *in situ* hybridization. Either the primary probe that directly binds the sequence can be fluorescently labeled, or the sequence can be first bound with an unlabeled primary probe containing overhangs, which function as landing pads for hybridization of fluorescent secondary "readout" probes. ix) Approaches to detect single mRNA molecules *in situ* by first generating rolling circle-amplified cDNA. * *In situ* sequencing can be used both for candidate-based approaches, to read out barcode sequences, or for unbiased approaches, to read out short sequences of the transcripts. B) Common approaches to profile epigenomic features associated with a cell type. **: Strategies currently available to profile at the single cell level. C) Approaches to simultaneously profile more than one modality in a sample.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

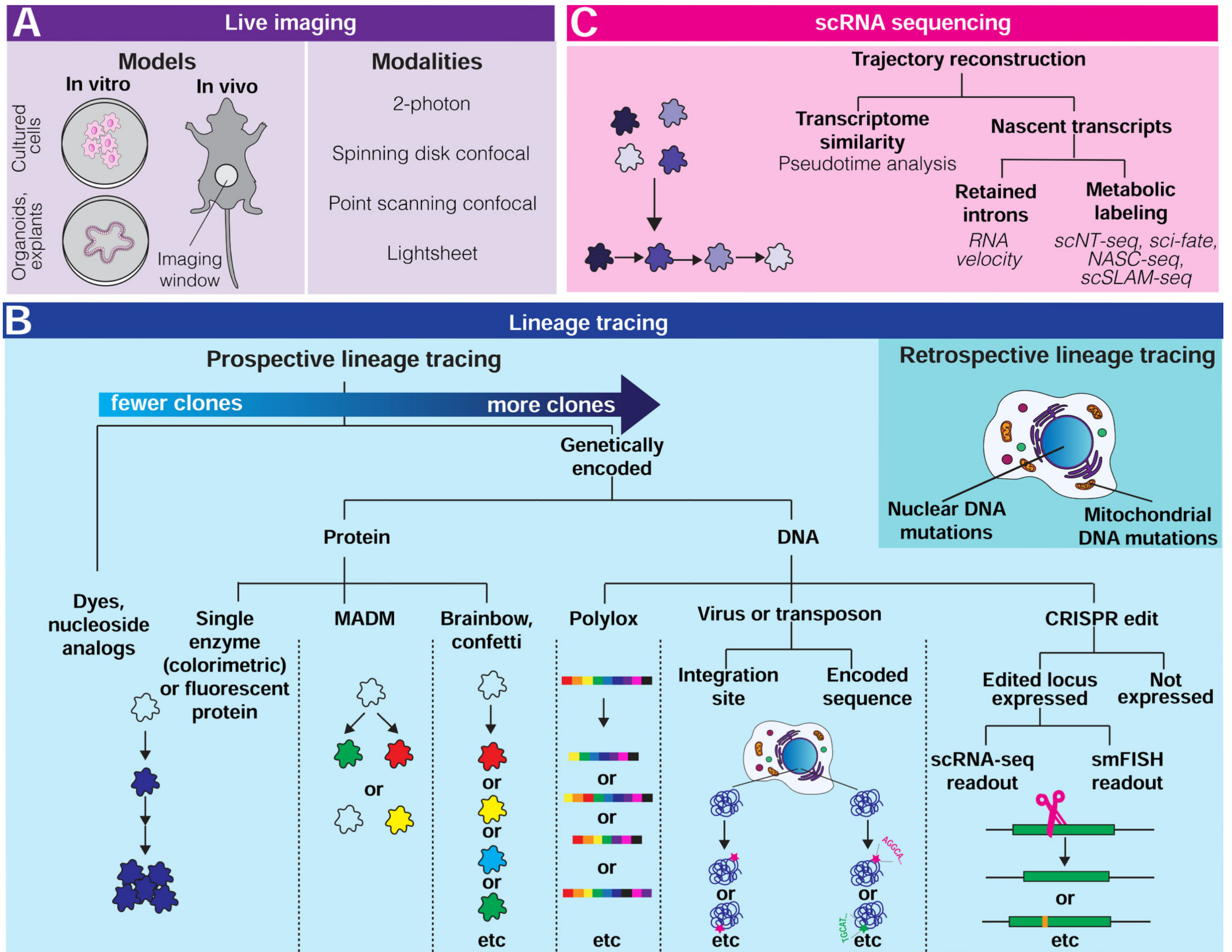

Figure 3. Strategies to define cellular relationships.

A) Models and modalities to facilitate live imaging of cellular relationships. B) Lineage tracing approaches, grouped according to the kind of reporter used by the system. C) Approaches for defining cellular relationships with scRNA sequencing.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

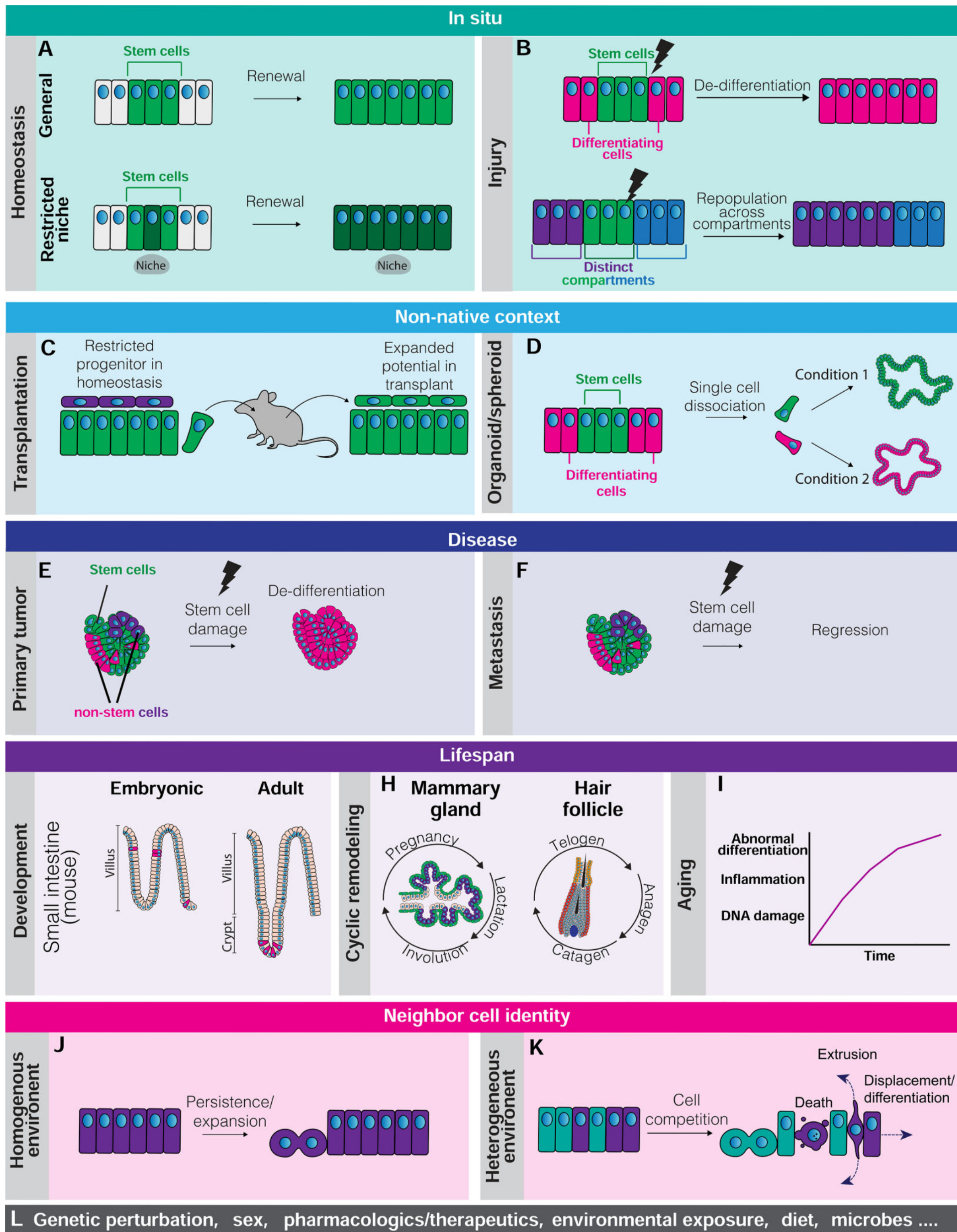

Figure 4. Cellular functions vary with context.

A) In homeostasis, canonical stem cells (green) give rise to all cells of the epithelium, including new stem cells. In some tissues, stem cells in close proximity to niche factors (dark green) expand preferentially over more peripheral stem cells (light green). B-D) Cells with diverse lineage histories can exhibit stem-like functions in different physiological and experimental contexts. Experimental and pathological alterations can expand the repopulation capacity of a cell type, either endowing cells other than the canonical stem cell with the capacity to repopulate tissue (pink cells), or allowing cells to repopulate across

previously non-permissive lineage boundaries. E-F) Cell function and contribution can vary with disease stage. G) Cellular relationships can be altered between developing tissues and adult tissues. For example, in the mouse small intestine, during development all cells of the epithelium can become stem cells, whereas in the adult, stem cells are restricted to the base of crypts. H) A subset of tissues exhibit cyclic differences in cellular composition and function. I) Cells in many tissues such as the skin, intestine, oral mucosa and hematopoietic system exhibit a variety of alterations with age. J-K) Cell competition can cause cells that can normally sustain and repopulate a tissue to be selectively eliminated when combined with fitter neighbors. L) In addition to the contexts presented in A-K, diverse additional sources of variability can alter cellular functions.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

	Genotype	Win or Lose (compared to wt cells)	References
Hyperplasia or oncogenesis	*yorkie*(high dose, in testis cyst stem cells)	W	(Amoyel et al., 2014) (Huang et al., 2005)
	ras^{v12}(with *scrib*−/− or *lgl*−/− or *dlg*−/−)	W	(Pagliarini and Xu, 2003)
	ptch$^{-/-}$ (in testis cyst stem cells)	W	(Amoyel et al., 2014)
	ex$^{+/+}$, *ft*$^{+/+}$, *sav*$^{+/+}$, *hpo*$^{+/+}$, or *wts*$^{+/+}$	W	(Tyler et al., 2007)
	apc$^{-/-}$ or *axin*$^{-/-}$	W	(Vincent et al., 2011) (Suijkerbuijk et al., 2016)
	egfr(high dose)	W	(Eichenlaub et al., 2016)
	dco^{3}	W	(Jursnich et al., 1990)
	ago$^{-/-}$	W	(Moberg et al., 2001)
	avl(low dose) and *crb*(high dose)	W	(Hafezi et al., 2012) (Lu and Bilder, 2005)
	bam$^{\Delta 86}$, *bgcn*^{20093} or *bgcn*^{20915} (in female germline stem cells)	W	(Jin et al., 2008)
Organ size maintained	*minute*$^{+/-}$	L	(Kucinski et al., 2017; Morata and Ripoll, 1975)
	yorkie$^{-/-}$	L	(Huang et al., 2005)
	scrib$^{-/-}$	L	(Brumby and Richardson, 2003)
	mahj$^{-/-}$	L	(Kucinski et al., 2017) (Tamori et al., 2010)
	lgl$^{-/-}$	L	(Tamori et al., 2010) (Menéndez et al., 2010)
	ept^{2}	L	(Moberg et al., 2005)
	tsc1(high dose) and *tsc2*(high dose)	L	(Potter et al., 2001)
	nrf2(high dose)	L	(Kucinski et al., 2017)
	brk(high dose)	L	(Moreno et al., 2002)
	dlg$^{-/-}$	L	(Igaki et al., 2009)
	azot$^{+/+}$	L	(Merino et al., 2015)
	vps25^{PB2931}	L	(Thompson et al., 2005)
	rab5(low dose)	L	(Ballesteros-Arias et al., 2014)
	fwe(low dose)	L	(Rhiner et al., 2010)
	csk$^{-/-}$	L	(Vidal et al., 2006)
	fz$^{-/-}$ and *fz2*$^{-/-}$	L	(Vincent et al., 2011)
	myc(low dose)/ *myc*(high dose)	L/W	(de la Cova et al., 2004) (Moreno and Basler, 2004)
	stat(low dose)/ *stat*(high dose)	L/W	(Rodrigues et al., 2012)
	lgl$^{-/-}$ or *dlg*$^{-/-}$ (in ovary follicle stem cells)	W	(Kronen et al., 2014)
	socs36E(high dose, in male germline stem cells)	W	(Issigonis et al., 2009)

Figure 5. OverView of cell competition mutants in Drosophila.

Studies of cell competition in mammals have taken advantage of the extensive groundwork laid by research in flies. The table summarizes mutations that have been identified in fly to confer “winner” (W) or “loser” (L) status when combined with wild type cells. Mutations in which expansion of the winner cells drives organ enlargement are indicated in purple; mutations that result in competition but maintain organ size are indicated in blue.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript