ANNUAL REVIEWS

Annual Review of Cancer Biology

New Tools for Lineage Tracing in Cancer In Vivo

Matthew G. Jones,[1,2,3,4,5,*] Dian Yang,[1,2,*] and Jonathan S. Weissman[1,2]

[1]Whitehead Institute for Biomedical Research, Howard Hughes Medical Institute, David H. Koch Institute for Integrative Cancer Research, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA; email: weissman@wi.mit.edu, dyang@wi.mit.edu

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

[3]Biological and Medical Informatics Graduate Program and Integrative Program in Quantitative Biology, University of California, San Francisco, San Francisco, California, USA

[4]Center for Computational Biology, University of California, Berkeley, Berkeley, California, USA

[5]Center for Personal Dynamic Regulomes, Stanford University School of Medicine, Stanford, California, USA

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Annu. Rev. Cancer Biol. 2023. 7:111–29

First published as a Review in Advance on January 17, 2023

The *Annual Review of Cancer Biology* is online at cancerbio.annualreviews.org

https://doi.org/10.1146/annurev-cancerbio-061421-123301

Copyright © 2023 by the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See credit lines of images or other third-party material in this article for license information.

*These authors contributed equally to this article

Keywords

cancer biology, lineage tracing, single-cell analysis, tumor evolution, cancer metastasis, cell barcoding

Abstract

During tumor evolution, cancer cells can acquire the ability to proliferate, invade neighboring tissues, evade the immune system, and spread systemically. Tracking this process remains challenging, as many key events occur stochastically and over long times, which could be addressed by studying the phylogenetic relationships among cancer cells. Several lineage tracing approaches have been developed and employed in many tumor models and contexts, providing critical insights into tumor evolution. Recent advances in single-cell lineage tracing have greatly expanded the resolution, scale, and readout of lineage tracing toolkits. In this review, we provide an overview of static lineage tracing methods, and then focus on evolving lineage tracing technologies that enable reconstruction of tumor phylogenies at unprecedented resolution. We also discuss in vivo applications of these technologies to profile subclonal dynamics, quantify tumor plasticity, and track metastasis. Finally, we highlight outstanding questions and emerging technologies for building comprehensive cancer evolution roadmaps.

INTRODUCTION

Tumor growth and progression are shaped by evolutionary forces (Nowell 1976, Greaves & Maley 2012, Black & McGranahan 2021). From the transformation of a single cell, each successive cell division carries with it the possibility of introducing new genetic or epigenetic changes that can confer selective growth advantages in the daughter cells and are passed on through generations. A key consequence of this process is that the accumulation of these changes over time creates complex subpopulations in a tumor (i.e., intratumoral heterogeneity) where each subpopulation is the product of the unique set of changes in its ancestors (or lineage) (Shah et al. 2009, Marusyk & Polyak 2010, Navin & Hicks 2010, Navin et al. 2011, Gerlinger et al. 2012, McGranahan & Swanton 2017, Vendramin et al. 2021). Ultimately, this evolutionary process can lead to aggressive, resilient, and complex tumor masses with rare subpopulations capable of metastasizing throughout an individual or resisting targeted therapies (Hanahan & Weinberg 2011, Vogelstein et al. 2013, Gerlinger et al. 2014, Quintanal-Villalonga et al. 2020). Understanding the (epi)genetic changes in each lineage giving rise to aggressive subpopulations has been a central goal in cancer research, as it can reveal fundamental insights into tumor progression and nominate therapeutic targets (Greenman et al. 2007, Jamal-Hanjani et al. 2015, Amirouchene-Angelozzi et al. 2017).

Tumor lineages can be elucidated through lineage tracing approaches, which provide a suite of techniques enabling researchers to track the emergence of new cell subpopulations, as well as their proliferation, and migration in vivo (Woodworth et al. 2017). Although a variety of lineage tracing approaches have been employed in the context of cancer research, historically most efforts have taken one of two approaches: (*a*) the use of reporters or barcodes in model organisms, which enable one to follow all of the progeny of a clone marked at the beginning of the experiment, or (*b*) the use of naturally occurring genomic variation [often in the form of single-nucleotide variants or copy number variation (CNV)] as natural labels that can be used to reconstruct tumor lineages (Navin & Hicks 2010, Bailey et al. 2021, Tarabichi et al. 2021).

In this review, we focus on the development and application of new, high-resolution lineage tracing technologies in the context of cancer biology. We first summarize the development of different lineage tracing tools and then discuss the application of these technologies to trace tumor progression, with a focus on methods that enable a single-cell readout and high-resolution profiling of tumor evolution. Then, we provide a review of recent applications of these tools to profile key questions in tumor biology such as metastatic behavior and drug resistance. Finally, we highlight open questions in cancer biology and speculate on the potential technological innovations that will support these investigations and further increase the scope, dimension, and precision of cancer lineage tracing studies. For additional discussion of lineage tracing in human samples and additional technological overviews we refer the reader to several other reviews (Kretzschmar & Watt 2012, Woodworth et al. 2017, Wagner & Klein 2020, Bailey et al. 2021, Black & McGranahan 2021, Penter et al. 2021b, VanHorn & Morris 2021, Vendramin et al. 2021).

TECHNOLOGY

Overview of Lineage Tracing Technologies

Lineage tracing is a suite of methods for identifying all descendants of a single cell and is thus a powerful approach for understanding tissue development, homeostasis, and disease progression (Kretzschmar & Watt 2012). Generally, lineage tracing approaches are classified as either prospective, in which founder cells are experimentally labeled to track their descendants, or retrospective, in which shared spontaneous genetic variations are used to reconstruct lineage relationships (Woodworth et al. 2017). As mentioned above, studies of primary human tumors have generally used natural genetic variation to perform retrospective lineage tracing (Vogelstein et al. 2013,

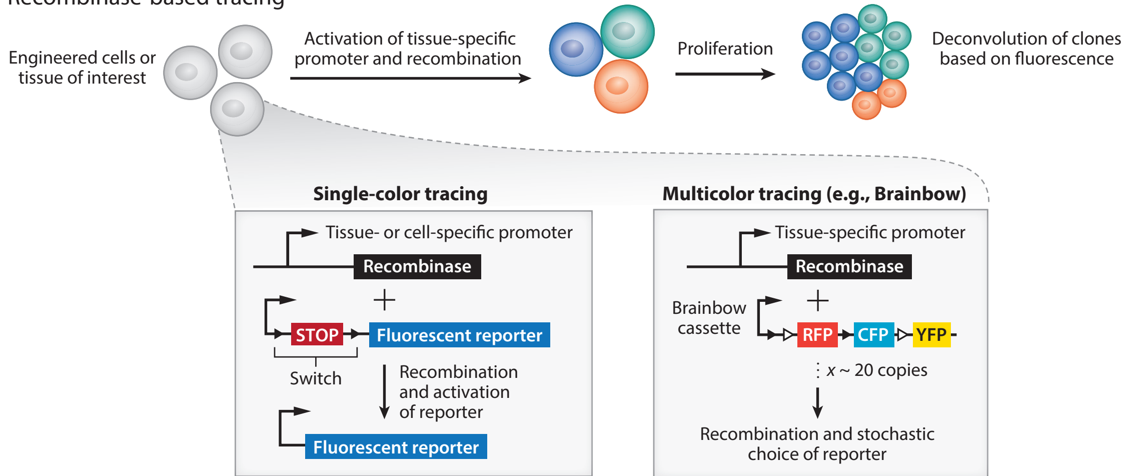
Turajlic et al. 2019, Abyzov & Vaccarino 2020, Gerstung et al. 2020, Bailey et al. 2021). To date, these human tumor lineage tracing efforts have led to the discovery of key principles underpinning tumor development, including the acquisition of subclonal genetic or epigenetic changes (Jones et al. 2008, Gerlinger et al. 2012, Sottoriva et al. 2015, Williams et al. 2018, Minussi et al. 2021), the timing and routes of metastatic spread (Yachida et al. 2010, Yates et al. 2017, Turajlic et al. 2018, Hu et al. 2020), and the development of therapeutic resistance (Abbosh et al. 2017, Kim et al. 2018, Powles et al. 2021, Salehi et al. 2021).

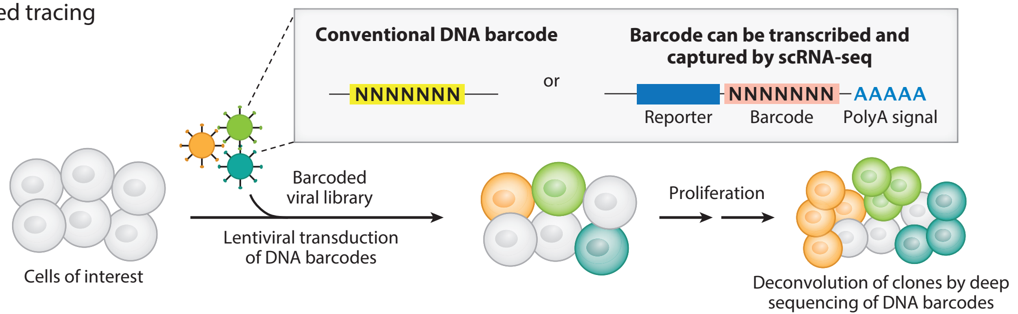
Though the importance of these studies cannot be overstated, there are several significant limitations to these approaches. First, several of the early studies have relied on averaging together signals across sampled regions (i.e., multiregion sampling), which can obfuscate critical intratumoral heterogeneity. Furthermore, although recent advances in single-cell sequencing technologies have enabled high-resolution tumor lineage profiling, whole-genome sequencing (WGS) of single cells is highly expensive and thus has limited scalability. To note, investigators have had success reconstructing tumor lineages without WGS by leveraging CNV (Patel et al. 2014, Gao et al. 2021), mitochondrial variation (Ludwig et al. 2019, Penter et al. 2021a), methylation states (Gabbutt et al. 2022), and targeted profiling of highly variable genomic features like short tandem repeats (Tao et al. 2021). Moreover, beyond technological considerations, human tumor studies are limited by confounding variables (e.g., environmental exposures and genetic background) and the inherent variability in the timing of naturally occurring mutations, and they are not amenable to in vivo perturbations or functional studies. As such, synthetic lineage tracing in model organisms serves as a critical complement to human tumor studies. Below, we focus on synthetic lineage tracing systems with consideration toward how these technologies can be applied to tumor tracing.

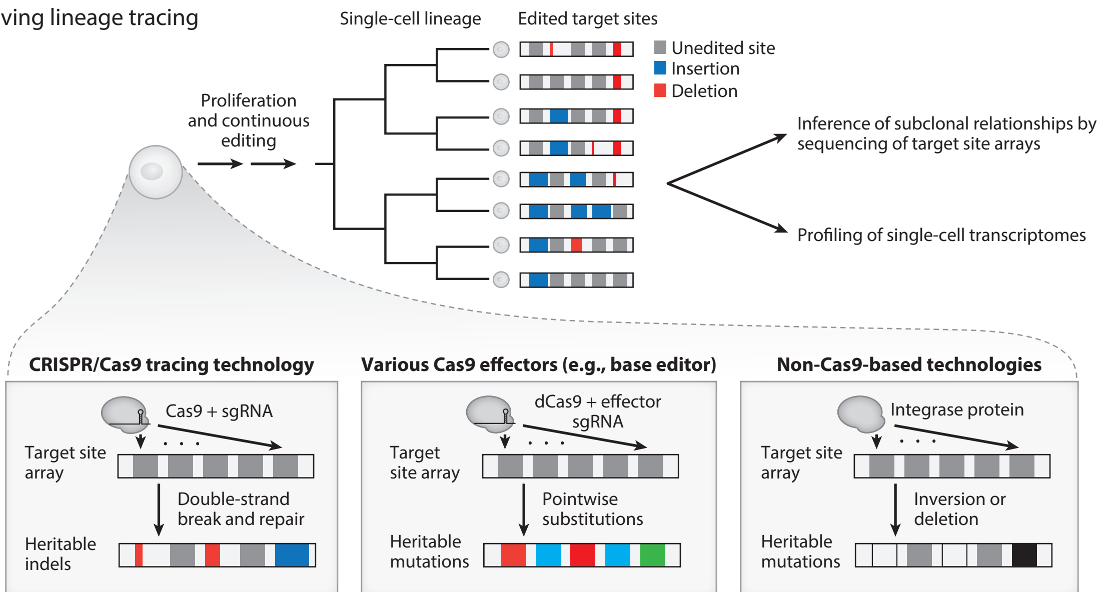
Engineered Static Lineage Tracing Systems

Conventional lineage tracing technologies rely on a reporter, such as a fluorescent protein or a DNA barcode. Specifically, cells are engineered with a lineage reporter (ideally introducing a unique reporter into each cell) that is passed on through cell division. Because these reporters are heritable (i.e., static), one can identify individual cell lineages based on progenies sharing the same reporters. Below, we provide a survey of static lineage tracing methods; we additionally refer the reader to other reviews for more in-depth discussion of these approaches (Kretzschmar & Watt 2012, Woodworth et al. 2017, VanHorn & Morris 2021).

Recombinase-based lineage tracing methods. One common approach for introducing static lineage reporters into cells has relied on recombinases that induce recombination between DNA target sites to create inversions or deletions in a predictable manner (Nagy 2000, Liu et al. 2020). Typically, a recombinase serves as a switch (e.g., Cre-loxP, FLP-FRT, or Dre-rox) that induces expression of a heritable reporter, such as a fluorescent protein (**Figure 1*a***). More recently, the ability to label specific cell types has been improved by using multiple recombinases (He et al. 2017, Liu et al. 2019) or by engineering split-recombinase systems where each fragment is driven by a different promoter and only cells expressing both promoters will successfully achieve recombination (Hirrlinger et al. 2009). Sparse labeling of rare cells, such as using Cre-ER (estrogen receptor) with low-dose tamoxifen, is one solution to increase the labeling resolution, with the caveat that the labeled cells do not represent the overall population. To overcome this and enable a more comprehensive analysis of individual clones, researchers use multicolor reporters, such as Confetti and Brainbow, that leverage stochastic Cre-mediated recombination to induce single or combinatorial expression of multiple fluorescent reporters (Livet et al. 2007, Weissman & Pan 2015, Snippert et al. 2010).







(Caption appears on following page)

Figure 1 (*Figure appears on preceding page*)

Summary of major engineered lineage tracing technologies. (*a*,*b*) Static lineage tracers report on the clonal relationship of cells. (*a*) Recombinase-based lineage tracing approaches engineer cells or tissues to activate a heritable, fluorescent reporter or a combination of multiple fluorescent proteins upon expression of a recombinase. (*b*) Static DNA barcoding technologies integrate a random DNA barcode into cells (for example, through viral transduction or transposon integration) that can be read out using high-throughput sequencing approaches. Linking barcodes to expressed reporter genes enables the simultaneous capture of transcriptomic profile and barcode identity. (*c*) Evolving lineage tracing technologies leverage gene-editing machinery (such as CRISPR/Cas9, one of its various effector variants, site-specific integrases, or recombinases) to continuously edit designated target site regions to convey information about the evolutionary history of a cell population and reveal intraclonal relationships of labeled cells. Abbreviations: dCas9, dead (endonuclease-deficient) Cas9; indels, insertion and deletion mutations; RFP (CFP, YFP), red (cyan, yellow) fluorescent protein; scRNA-seq, single-cell RNA sequencing; sgRNA, single guide RNA.

Synthetic barcode–based lineage tracing methods. Important challenges of diversity intrinsic to fluorescent reporters have been solved by the rapid development of next-generation sequencing technologies and high-diversity DNA barcoding technologies. Early versions of this technology utilized vectors carrying defined DNA sequences that could be stably introduced into cells for distinguishing clones (Walsh & Cepko 1992, Schepers et al. 2008); further improvements replaced defined sequences with high-diversity random barcode sequences (Gerrits et al. 2010). The barcodes are heritable and all descendants derived from each labeled cell clone will inherit the same barcode (**Figure 1*b***). This technology has been instrumental in enabling the high-resolution and quantitative analysis in a broad range of cancer-related studies, including analyzing tumor progression and clonal dynamics, mapping of primary tumor–metastasis relationships (Chuang et al. 2017, Merino et al. 2019), and identifying the origins of drug resistance (Bhang et al. 2015, Hata et al. 2016).

The recent emergence of single-cell technologies has led to a new wave of DNA barcode-based lineage tracing technologies. A fundamental advance has been to pair DNA barcodes with RNA reporters (**Figure 1*b***) (e.g., by inserting a random barcode in the 3′ UTR of a transgene). This results in cells containing multiple mRNA copies of the DNA barcode, thus enabling observation with single-cell RNA-seq (scRNA-seq) technologies (Adamson et al. 2016, Dixit et al. 2016, Yao et al. 2017, Biddy et al. 2018, Weinreb et al. 2020). The pairing of scRNA-seq and DNA barcoding has yielded rich insights into clone-specific differentiation hierarchies (Wagner et al. 2018), kinetics (Biddy et al. 2018), and biases (Weinreb et al. 2020). To expand the compatibility with different data modalities, researchers have developed a variety of barcodes to enable epigenomic (Pierce et al. 2021) or proteomic analysis (Wroblewska et al. 2018, Rovira-Clavé et al. 2021).

Coupling lineage tracing with other genomic technologies can be powerful. For example, barcoding combined with pooled CRISPR-based perturbations has enabled the high-throughput dissection of gene function both in cell culture and in vivo (Michlits et al. 2017, Rogers et al. 2017, Schmierer et al. 2017, Winters et al. 2017). Moreover, retroactive characterization or isolation of rare clones of interest from an initial pool of cells has been enabled by the integration of single-cell barcoding technologies with RNA FISH (fluorescence in situ hybridization) (Emert et al. 2021), CRISPR/Cas9-induced frameshift mutations (Feldman et al. 2020) or CRISPR activation (CRISPRa) (Al'Khafaji et al. 2018, Gutierrez et al. 2021, Umkehrer et al. 2021). Typically, these systems sequence cells after some experimental selection (e.g., drug treatment) to identify enriched or depleted barcodes and then retroactively isolate clones from the initial cell line (e.g., using CRISPRa and single-guide RNAs (sgRNAs) complementary to the barcodes of interest to specifically activate a fluorescent reporter in the cells of interest). Together, this rapidly expanding toolkit offers a variety of approaches for studying the genetic control over complex behavior and the emergence of rare subpopulations in tumors.

Engineered Evolving Lineage Tracing Systems

The approaches discussed above rely on introducing heritable marks at the clonal level and prospectively labeling all progeny. While some aforementioned approaches are amenable to sequential rounds of barcoding [e.g., with CellTagging (Biddy et al. 2018)] to gain subclonal resolution, they are challenging to implement for in vivo study. To address this, researchers have begun to develop evolving lineage tracing technologies that combine concepts from both prospective tracing approaches, whereby heritable marks are used to track clonal populations, and retrospective tracing approaches, whereby random, heritable genetic diversity is introduced over time that can be used to infer lineage relationships.

CRISPR/Cas9-based evolving barcoding technologies. At their core, these evolving barcoding tools require randomness to be continually introduced into a defined genetic locus over the course of the experiment in a manner that is inherited throughout cell divisions (as reviewed in McKenna & Gangon 2019). The most common version of such approaches has leveraged CRISPR/Cas9 gene-editing technologies to introduce random mutations at specific, synthetic DNA scratchpads or target sites (McKenna et al. 2016, Frieda et al. 2017, Kalhor et al. 2017, Garcia-Marques et al. 2020). While several elegant strategies have been developed, each relies on three basic components: Cas9 nuclease, DNA target sites that are integrated into the genomes, and single-guide RNAs (sgRNAs) that specifically target Cas9 to the target sites. Together, these three components enable Cas9 to introduce targeted double-strand breaks at the DNA target site locus that are subsequently repaired in an error-prone fashion to generate heritable insertions or deletions (indels) (**Figure 1*c***). Importantly, these DNA target sites can be transcribed into poly-adenylated mRNAs, allowing them to be captured and profiled along with all other cellular mRNAs using massively parallel scRNA-seq techniques (e.g., droplet-based or split-pool strategies). In doing so, this approach makes it possible to directly link the current cell state (as measured by scRNA-seq) with its past history (as captured by the lineage recorder), and to do so on a massive scale (Alemany et al. 2018, Raj et al. 2018, Spanjaard et al. 2018, Chan et al. 2019, Bowling et al. 2020).

Concomitant with the development of these technologies has been the emergence of new analytical methods for handling the scale and complexity of the data. While initial studies had success with reconstructing lineages using existing approaches like Camin-Sokal maximum parsimony (Camin & Sokal 1965) and neighbor-joining (Saitou & Nei 1987), there has been significant interest in developing algorithms specifically tailored to engineered systems (Gong et al. 2021). Several new approaches have emerged for reconstructing lineages from engineered systems, such as Cassiopeia (Jones et al. 2020), LinTIMat (lineage tracing by integrating mutation and transcriptomic data; Zafar et al. 2020), GAPML (GESTALT analysis using penalized maximum likelihood; Feng et al. 2021), PhyloTime (Fang et al. 2022), and DCLEAR (distance-based cell lineage reconstruction; Gong et al. 2022). Moreover, investigations have provided theoretical results for improving reconstruction with better lineage tracing designs (Salvador-Martínez et al. 2019, Wang et al. 2021). Beyond reconstructing lineages, significant effort has been invested in creating computational approaches for inferring so-called fate maps of how specific progenitor populations give rise to various cell types (Fang et al. 2022, Wang et al. 2022, Yang et al. 2022), inferring ancestral gene expression states (Forrow & Schiebinger 2021, Ouardini et al. 2021), and providing tools for interactively exploring these lineages (Salvador-Martínez et al. 2021, Jones et al. 2022).

Other evolving lineage tracing tools. While Cas9 is an efficient, flexible, and commonly used tool for lineage tracing, there are three limitations. First, DNA double-strand breaks generated by Cas9 editing can cause cell stress and potential cytotoxicity (Ihry et al. 2018, Geisinger & Stearns 2020). Second, there is little control over the types and sizes of indels, and large edits are

challenging to measure (McKenna et al. 2016, Hussmann et al. 2021). Third, the use of Cas9 to edit arrays of adjacent DNA target sites presents the risk of collapsing several sites together and thus losing intermediate information (McKenna et al. 2016).

A variety of tools have been developed to overcome some of these issues. To minimize unintended cytotoxicity, technologies that avoid double-stranded breakage and endogenous repair of DNA can be favorable, such as those using site-specific mutations with deaminases (Hwang et al. 2019, Askary et al. 2020, Cravens et al. 2021, Liu et al. 2021), site-specific serine integrases (Chow et al. 2021), Polylox recombination cassettes (Pei et al. 2017), or transposon-based barcoding (Wagner et al. 2018) (**Figure 1*c***). To better control the type, size, and temporal order of edits, clever approaches have recently been developed by pairing Cas9 with template-independent DNA polymerases (Loveless et al. 2021b) or leveraging prime editing for template-based DNA insertions (Loveless et al. 2021a, Choi et al. 2022).

Signal recording tools. Similar to lineage information, because of their transient nature, many important cell signaling events cannot be detected at the end point of experiments. To capture and reconstruct intermediate cell states during the evolutionary process, researchers have expanded molecular recording toolkits with a variety of signal recording technologies. For example, elegant approaches leveraging the CRISPR spacer acquisition process have enabled transcriptional recording in bacteria (Shipman et al. 2016, Schmidt et al. 2018). In this section, we focus on signal recording technologies in mammalian cells.

One effective approach for recording molecular signals is to engineer barcodes under a tissue- or gene-specific promoter. Conventional recombinase-based lineage tracing methods have leveraged this approach by pairing the expression of a recombinase with the promoter of a desired lineage marker (Kretzschmar & Watt 2012). Recently, a dual-recombinase-based fate mapping system was developed to record transient epithelial-to-mesenchymal (EMT) fluctuations during metastasis and revealed that N-cadherin, but not vimentin, labeled metastatic-initiating cells in the MMTV-PyMT (mouse mammary tumor virus–polyoma middle tumor-antigen) mouse model (Li et al. 2020). In another example, investigators introduced a barcoding method dubbed "Watermelon," which leverages a dual-fluorescent reporter to not only trace drug-resistant clones but also report on the proliferative state of the clone by monitoring the dilution by cell division of an inducible H2B-mCherry transgene (Oren et al. 2021).

Due to their high modularity, Cas9-based evolving lineage tracers can be readily adapted to be signal recorders by coupling sgRNA or Cas9 expression with signal response elements (Perli et al. 2016, Frieda et al. 2017). Variants of Cas9-based editors enable multiplexed signal recording: for example CAMERA, a base-editor system that introduces C • G to T • A edits to different sites of a safe harbor locus (Tang & Liu 2018), and ENGRAM, which utilizes a prime editor to record activity and dynamics of multiple transcriptional reporters in DNA (Chen et al. 2021). In addition to these CRISPR-related recording tools, Lin et al. (2021) engineered intracellular protein fibers that can grow slowly in cells by incorporating diverse fluorescent marks, allowing for accurate time-stamping. Overall, these tools for capturing critical intermediate signals offer a powerful opportunity to expand our understanding of unobserved cellular history.

APPLICATIONS

Efforts over the past decade have underscored the value of lineage tracing tools to answer fundamental questions around cancer development: from elucidating the requisite genetic or epigenetic changes for tumor initiation and progression, to generating insights into how cellular plasticity arises and influences tumor progression and therapeutic resistance, to characterizing the patterns

of metastatic dissemination. In this section, we discuss major discoveries with a focus on recent in vivo tumor studies empowered by evolving lineage tracing technologies (**Figure 2**).

Monitoring Tumor Progression

As described above, tumorigenesis typically unfolds over long periods of time. Over the past decade, static lineage tracing has been widely used to determine the cell type origins of cancers (Driessens et al. 2012, Schepers et al. 2012), the relative growth rate of different tumor clones (Lamprecht et al. 2017, Rogers et al. 2017, Reeves et al. 2018), the influence of oncogenes and tumor suppressors on tumor progression (Nguyen et al. 2015, Winters et al. 2017, Cai et al. 2021), and the molecular nature of tumor-initiating cellular states (Fennell et al. 2022) (**Figure 2**).

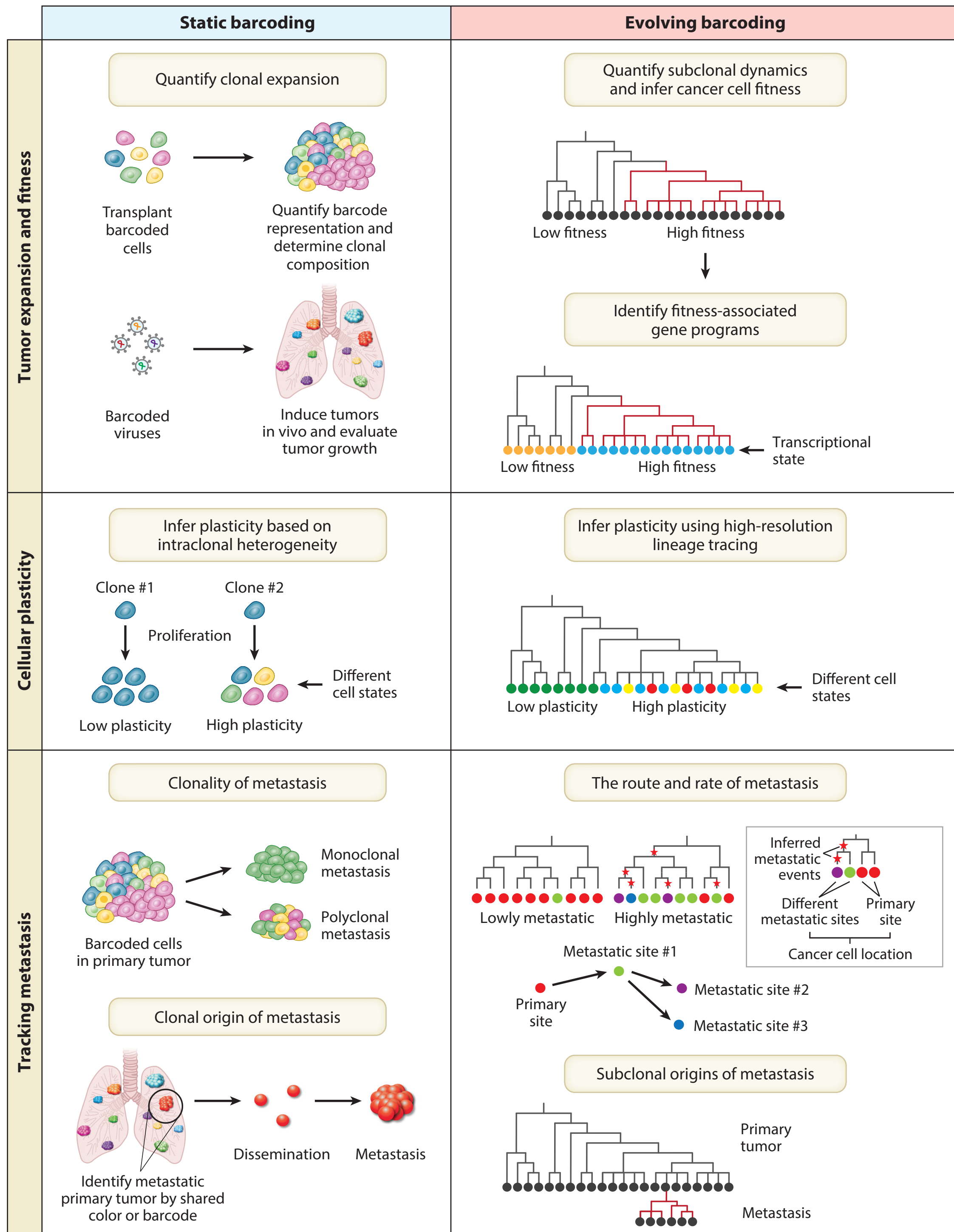
The use of evolving lineage tracing systems has provided a unique opportunity to dissect how distinct subclones contribute to tumor development at a much higher resolution. Yang et al. (2022) engineered the $Kras^{LSL\text{-}G12D/+}$; $Trp53^{fl/fl}$ (KP) lung cancer mouse model, developed by the Jacks lab, with an evolving lineage tracer such that the addition of Cre recombinase would simultaneously activate evolving lineage tracing and the oncogenic *Kras* and *Trp53* mutations in individual lung epithelial cells. This system enabled the continuous tracking of tumor evolution from single transformed cells to metastatic tumors at unprecedented resolution. From the deeply resolved lineages, the authors found that tumor progression is characterized by rare subclonal expansions driven by the adoption of distinct transcriptional programs (**Figure 2**). Building on previous efforts by multi–time point sampling of KP tumors (Marjanovic et al. 2020), this high-resolution phylogenetic analysis offered a phenotypic readout of tumor fitness by single–time point sampling and generated a comprehensive fitness landscape of tumor evolution.

Quantifying Tumor Plasticity

Tumor plasticity, or the rate at which cancer cells change their molecular and phenotypic features, has been implicated as a hallmark of tumor progression and therapeutic resistance (Flavahan et al. 2017, Le Magnen et al. 2018, Quintanal-Villalonga et al. 2020, Gutierrez et al. 2022). Several recent efforts have highlighted how lineage tracing approaches offer a powerful tool to characterize the nature and consequences of plasticity in tumor progression.

Evolving lineage tracers and the resulting high-resolution phylogenies have been highly for quantifying tumor plasticity, as they enable one to assess the transcriptomic differences between closely related cells (Rios et al. 2019, Chaligne et al. 2021, Yang et al. 2022). The first study to leverage this property used an inducible evolving tracer in a metastatic pancreatic cancer cell line to demonstrate that aggressive tumors occupy a continuous spectrum of EMT cell states with the late-hybrid EMT state being highly proliferative and metastatic (Simeonov et al. 2021), corroborating previous studies (Yu et al. 2013, Lüönd et al. 2021, Yang et al. 2020). More recently, Yang et al. (2022) developed the autochthonous lung cancer model with an evolving lineage tracing tool, discussed above, and revealed that the loss of the initial alveolar type 2 cell state was accompanied by a transient increase of transcriptome plasticity, leading to increased intratumoral heterogeneity and providing a substrate for the selection of aggressive subclones. Although early, these studies highlight the power of evolving lineage tracing technologies in improving our understanding of how tumor plasticity arises and is maintained and regulated.

Lineage tracing has also been critical for illuminating how plasticity contributes to therapeutic resistance (Boumahdi & De Sauvage 2020, Tulpule & Bivona 2020, Torborg et al. 2022). Static tracing with fluorescent reporters or DNA barcoding has been widely used for identifying the tumor-initiating cells of tumor relapse after drug treatment (Chen et al. 2012), determining clonal dynamics and fitness (Bhang et al. 2015, Grüner et al. 2016, Roh et al. 2018, Walens et al. 2020,



(Caption appears on following page)

Figure 2 (*Figure appears on preceding page*)

Illustration and comparison of how static and evolving lineage tracing technologies contribute to our understanding of tumor progression, plasticity, and metastasis. (*Left*) Static barcoding enables the study of evolutionary dynamics at a clonal level. For example, static barcodes enable quantification of the composition and relative growth rate of different tumor clones. It also enables comparison of the phenotypic heterogeneity between clonal populations and can be further applied to determine the clonality of metastasis and the relationships between metastatic lesions and their original primary tumor clones. (*Right*) By contrast, evolving barcoding approaches allow for a more quantitative, high-resolution assessment of the tumor's evolutionary dynamics at a subclonal level. For example, evolving lineage tracing technologies enable one to determine the fine subclonal lineage structure of cells derived from individual tumor clones and to identify the relative fitness and expansion of different subclones. Evolving lineage tracers also allow for a quantitative measurement of cellular plasticity by directly assessing the frequencies of cell-state transitions in related lineages. Finally, evolving lineage tracers can reveal the rates and routes of metastasis and track the origins of metastases to specific subclonal lineages within the primary tumors.

Fennell et al. 2022), and quantifying pharmacogenomic interactions (Foggetti et al. 2021, Li et al. 2021). More recently, pairing such approaches with single-cell transcriptomics has enabled the detailed dissection of the response to therapies and the development of drug resistance (Eyler et al. 2020, Oren et al. 2021, Chang et al. 2022). Moreover, retrospective lineage tracing approaches enable one to study the rare cells that give rise to drug-resistant tumors. For example, recent studies using the CaTCH (CRISPRa tracing of clones in heterogeneous cell populations) and Rewind retrospective lineage tracing technologies isolated and studied rare cell clones from treatment-naive populations to gain mechanistic insights into the genomic and transcriptomic alterations that drive melanoma resistance to targeted therapies (Umkehrer et al. 2021, Emert et al. 2021). Collectively, the success of applying lineage tracing to study these rare and transient events will provide a firm foundation for future endeavors integrating these technologies.

Tracking Tumor Metastasis

Metastasis accounts for approximately 90% of cancer-related deaths (Ganesh & Massagué 2021). Despite its immense clinical importance, the metastatic process remains difficult to study due to its temporally and spatially sporadic nature. Consequently, major open questions remain, including how metastases arise, what the rates and routes of metastatic spread are, and what the molecular features of metastasis-initiating cells are.

Several studies have demonstrated the promise of lineage tracing to address these questions. Early efforts employing static lineage tracing have illuminated the clonality of metastases. For example, studies utilizing multicolor tracing revealed that the metastases in some cancer models could arise from polyclonal seeding events (Maddipati & Stranger 2015, Reeves et al. 2018), whereas in other tumor types metastasis lesions were largely monoclonal (Caswell et al. 2014, Chiou et al. 2015). Other reports have leveraged DNA barcoding to improve the scale and readout in their studies. For example, one study used DNA barcoding and paired RNA sequencing to associate metastatic spread with specific transcriptomic programs (Chuang et al. 2017). More recently, a large-scale effort barcoded 500 cell lines spanning 21 types of solid tumors to create a comprehensive map of the metastatic routes and organ tropisms of different cancer types at unprecedented scale (Jin et al. 2020).

High-resolution, continuous lineage tracing systems have further enabled the quantitative measurement of metastatic rates and routes. Using a Cas9-based evolving recorder, Quinn et al. (2021) tracked metastatic behavior in a xenograft model of lung adenocarcinoma. The high-resolution lineages uncovered a remarkable heterogeneity in metastatic ability, which was characterized by preexisting and heritable gene expression differences of individual clones, but not their proliferative potential. Quinn et al. also systematically mapped the detailed steps and routes of metastatic spread for each clone and revealed that the mediastinal lymph node consistently

served as a major hub for lung cancer metastasis. Similarly, Zhang et al. (2021) tracked the relative timing of metastasis using an evolving lineage tracer and found in a breast cancer model that most metastases were polyclonal and primarily metastasized to the bone before forming secondary metastases. They have further uncovered that the bone microenvironment promotes tumor cell stemness and plasticity, thereby invigorating secondary metastases.

Evolving lineage tracing has also shed light on the origins of metastasis. A recent study followed pancreatic cancer progression and metastasis by engineering an inducible CRISPR/Cas9 lineage tracer (macsGESTALT) into a mouse pancreatic cancer cell line. This work revealed that cancer cells went through multiple rounds of bottlenecking from engraftment to metastasis and only rare clones with late hybrid EMT transcriptional states are the major source of metastasis (Simeonov et al. 2021). Using similar approaches in the KP lung cancer model revealed that metastases almost always arose from the expanding subclones in primary tumors (Yang et al. 2022). Both of these studies suggest that progression and selection at the primary site are important for metastasis. Collectively, these recent efforts serve to underscore the power of evolving lineage tracing approaches for elucidating fine-scale metastatic behaviors.

OUTLOOK AND FUTURE DIRECTIONS

The rapid development of lineage tracing technologies has enabled the reconstruction of tumor phylogenies at unprecedented resolution and scale, providing critical insights into the major steps of tumor evolution. Below, we outline several key avenues for future investigation that will further expand lineage tracing toolkits and facilitate comprehensive mapping of tumor evolution across different cancers.

Interrogating Tumor-Stromal Interactions

Cancer cells evolve within a dynamic and diverse ecological niche, commonly referred to as the tumor microenvironment (TME), that constantly exerts several selective pressures, including hypoxia, competition for space and nutrients, and inflammation, that collectively shape tumor evolution (Binnewies et al. 2018). However, cancer lineage tracing studies have largely focused on cancer cells without taking into consideration the spatial context and other cells and signals coexisting in the TME. A promising direction to preserve the TME information in cancer lineage tracing analysis is to capture tumor evolution in its native spatial context by integrating lineage tracing approaches with high-resolution spatial transcriptomics technologies enabled by sequencing or imaging (Chen et al. 2015, Moffitt et al. 2016, Ståhl et al. 2016, Eng et al. 2019, Stickels et al. 2021, Chen et al. 2022). Already, evolving lineage tracing approaches have successfully been paired with spatial transcriptome sequencing (He et al. 2022) or FISH-based methods (Frieda et al. 2017, Askary et al. 2020, Chow et al. 2021) for simultaneously profiling lineage, transcriptome, and spatial context. We anticipate that integrative spatial and lineage analysis of tumor evolution will provide insights into key questions around how the spatial arrangement of cancer cells and various cancer-stromal interactions influence tumor progression, plasticity, and metastasis.

Capturing Intermediate, Transient Cell States During Tumor Evolution

Employing concepts from classic lineage tracing efforts, an overarching goal is to identify stereotypical paths that a transformed cell takes to develop into an aggressive tumor. Various approaches have been employed to reconstruct these paths from a single sampling of cells with paired single-cell transcriptomic and lineage tracing readouts (Weinreb et al. 2020, Yang et al. 2022), although without prior knowledge of the developmental process these efforts are inherently limited by the possibility of unobserved intermediate states. While dense temporal sampling could partially

address this, incorporating multichannel molecular recorders capable of recording transient expression of genes that mark important intermediate cell states could significantly improve fate mapping efforts. Additionally, recent computational advances that infer fate maps (Fang et al. 2022) or ancestral gene expression states (Ouardini et al. 2021) will be useful for this task. Together, we anticipate that these tools will not only reveal what subclonal structures and evolutionary paths gave rise to aggressive tumors but also provide a mechanistic understanding into how and why each tumor made specific fate decisions by sensing and responding to different evolutionary forces.

Developing a Comprehensive, Multiomic, and Predictive Roadmap of Tumor Development

The efforts described above have focused on the relationship between transcriptome and lineage; however, a cell's identity is a product of the interplay between several molecular species that give rise to complex relationships between various gene regulatory circuits (Tanay & Regev 2017). In addition to transcriptome analysis, we must integrate other levels of the cell's genetic, epigenetic, and functional state to build comprehensive maps of tumor evolution (Granja et al. 2019, Neftel et al. 2019, LaFave et al. 2020). To this end, pairing lineage tracing with multiomic assays is likely to shed light on the regulatory features underlying consequential changes in the tumor (Ogbeide et al. 2022).

Beyond a descriptive understanding, we believe that a predictive model of tumor progression will require a fundamental understanding of how gene function affects evolution. Recent work pairing genetic perturbations with an evolving lineage tracer demonstrated how tumors' evolutionary trajectories can be altered by genetic perturbations (Yang et al. 2022). Future efforts combining functional genomics and lineage tracing will shed light on complex gene functions in the context of tumor evolution and nominate gene candidates that may regulate tumor expansion, plasticity, and metastasis. More broadly, integrating these technologies will aid in the construction of a comprehensive, multiomic roadmap of tumor development, from which we can build predictive models of tumor evolution, reveal key drivers of tumor progression, and develop new therapeutic strategies.

DISCLOSURE STATEMENT

J.S.W. declares outside interest in 5 AM Ventures, Amgen, Chroma Medicine, DEM Biosciences, KSQ Therapeutics, Maze Therapeutics, Tenaya Therapeutics, Tessera Therapeutics, and Velia Therapeutics. D.Y. declares outside interest in DEM Biosciences. M.G.J. declares outside interest in Vevo Therapeutics.

ACKNOWLEDGMENTS

The Weissman lab was supported in part by the NCI (National Cancer Institute) Cancer Target Discovery and Development (CTD^2) and the NIH (National Institutes of Health) Centers of Excellence in Genomic Science (CEGS), the NCI Cancer Center Support (core) grant P30-CA14051, the Howard Hughes Medical Institute, and the Ludwig Center at MIT. D.Y. is supported by a Damon Runyon Cancer Research Foundation Postdoctoral Fellowship (DRG-2238–18). M.G.J. is supported by a UCSF Discovery Fellowship.

LITERATURE CITED

Abbosh C, Bikbak NJ, Wilson GA, Jamal-Hanjani M, Constantin T, et al. 2017. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. *Nature* 545(7655):446–51

Abyzov A, Vaccarino FM. 2020. Cell lineage tracing and cellular diversity in humans. *Annu. Rev. Genom. Hum. Genet.* 21:101–16

Adamson B, Norman TM, Jost M, Cho MY, Nuñez JK, et al. 2016. A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell* 167(7):1867–82.e21

Al'Khafaji AM, Deatherage D, Brock A. 2018. Control of lineage-specific gene expression by functionalized gRNA barcodes. *ACS Synthet. Biol.* 7(10):2468–74

Alemany A, Florescu M, Baron CS, Peterson-Maduro J, van Oudenaarden A. 2018. Whole-organism clone tracing using single-cell sequencing. *Nature* 556(7699):108–12

Amirouchene-Angelozzi N, Swanton C, Bardelli A. 2017. Tumor evolution as a therapeutic target. *Cancer Discov*. 7(8):805–17

Askary A, Sanchez-Guardado L, Linton JM, Chadly DM, Budde MW, et al. 2020. In situ readout of DNA barcodes and single base edits facilitated by in vitro transcription. *Nat. Biotechnol.* 38(1):66–75

Bailey C, Black JRM, Reading JL, Litchfield K, Turajlic S, et al. 2021. Tracking cancer evolution through the disease course. *Cancer Discov*. 11(4):916–32

Bhang HC, Ruddy D, Radhakrishna VK, Caushi JX, Zhao R, et al. 2015. Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. *Nat. Med.* 21(5):440–48

Biddy BA, Kong W, Kamimoto K, Guo C, Waye SE, et al. 2018. Single-cell mapping of lineage and identity in direct reprogramming. *Nature* 564(7735):219–24

Binnewies M, Robetts EW, Kersten K, Chan V, Fearon D, et al. 2018. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat. Med.* 24(5):541–50

Black JRM, McGranahan N. 2021. Genetic and non-genetic clonal diversity in cancer evolution. *Nat. Rev. Cancer* 21(6):379–92

Boumahdi S, de Sauvage FJ. 2020. The great escape: tumour cell plasticity in resistance to targeted therapy. *Nat. Rev. Drug Discov.* 19(1):39–56

Bowling S, Sritharan D, Osorio FG, Nguyen M, Cheung P, Rodriguez-Fraticelli A, et al. 2020. An engineered CRISPR-Cas9 mouse line for simultaneous readout of lineage histories and gene expression profiles in single cells. *Cell* 181(6):1410–22.e27

Cai H, Chew SK, Li C, Tsai MK, Andrejka L, et al. 2021. A functional taxonomy of tumor suppression in oncogenic KRAS–driven lung cancer. *Cancer Discov*. 11(7):1754–73

Camin JH, Sokal RR. 1965. A method for deducing branching sequences in phylogeny. *Evol. Int. J. Org. Evol.* 19(3):311–26

Caswell DR, Chuang C, Yang D, Chiou S, Cheemalavagu S, et al. 2014. Obligate progression precedes lung adenocarcinoma dissemination. *Cancer Discov*. 4(7):781–89

Chaligne R, Gaiti F, Silverbush D, Schiffman JS, Weisman H, et al. 2021. Epigenetic encoding, heritability and plasticity of glioma transcriptional cell states. *Nat. Genet.* 53(10):1469–79

Chan MM, Smith ZD, Grosswendt S, Kretzmer H, Norman TM, et al. 2019. Molecular recording of mammalian embryogenesis. *Nature* 570(7759):77–82

Chang MT, Shanahan F, Nguyen TTT, Staben ST, Gazzard L, et al. 2022. Identifying transcriptional programs underlying cancer drug response with TraCe-Seq. *Nat. Biotechnol.* 40(1):86–93

Chen A, Liao S, Cheng M, Ma K, Wu L, et al. 2022. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* 185(10):1777–92.e21

Chen J, Li Y, Yu T, McKay RM, Burns DK, et al. 2012. A restricted cell population propagates glioblastoma growth after chemotherapy. *Nature* 488(7412):522–26

Chen KH, Boettiger AN, Moffitt JR, Wang S, Zhuang X. 2015. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* 348(6233):aaa6090

Chen W, Choi J, Nathans JF, Agarwal V, Martin B, et al. 2021. Multiplex genomic recording of enhancer and signal transduction activity in mammalian cells. bioRxiv 2021.11.05.467434. **https://doi.org/10.1101/2021.11.05.467434**

Chiou S, Winters IP, Wang J, Naranjo S, Dudgeon C, et al. 2015. Pancreatic cancer modeling using retrograde viral vector delivery and in vivo CRISPR/Cas9-mediated somatic genome editing. *Genes Dev.* 29(14):1576–85

Choi J, Chen W, Minkina A, Chardon FM, Suiter CC, et al. 2022. A time-resolved, multi-symbol molecular recorder via sequential genome editing. *Nature* 608(7921):98–107

Chow KK, Budde MW, Granados AA, Cabrera M, Yoon S, et al. 2021. Imaging cell lineage with a synthetic digital recording system. *Science* 372(6538):abb3099

Chuang C, Greenside PG, Rodgers ZN, Brady JJ, Yang D, et al. 2017. Molecular definition of a metastatic lung cancer state reveals a targetable CD109–Janus kinase–Stat axis. *Nat. Med.* 23(3):291–300

Cravens A, Jamil O, Kong D, Sockolosky JT, Smolke CD. 2021. Polymerase-guided base editing enables in vivo mutagenesis and rapid protein engineering. *Nat. Commun.* 12:1579

Dixit A, Parnas O, Li B, Chen J, Fulco CP, et al. 2016. Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* 167(7):1853–66.e17

Driessens G, Beck B, Caauwe A, Simons BD, Blanpain C. 2012. Defining the mode of tumour growth by clonal analysis. *Nature* 488(7412):527–30

Emert BL, Cote CJ, Torre DA, Dardani IP, Jiang CL, et al. 2021. Variability within rare cell states enables multiple paths toward drug resistance. *Nat. Biotechnol.* 39(7):865–76

Eng CL, Lawson M, Zhu Q, Dries R, Koulena N, et al. 2019. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. *Nature* 568(7751):235–39

Eyler CE, Matsunaga H, Hovestadt V, Vantine SJ, van Galen P, Bernstein BE. 2020. Single-cell lineage analysis reveals genetic and epigenetic interplay in glioblastoma drug resistance. *Genome Biol.* 21:174

Fang W, Bell C, Sapirstein A, Asami S, Leeper K, et al. 2022. Quantitative fate mapping: reconstructing progenitor field dynamics via retrospective lineage barcoding. bioRxiv 10.1101/2022.02.13.480215. **https://doi.org/10.1101/2022.02.13.480215**

Feldman D, Tsai F, Garrity AJ, O'Rourke R, Brenan L, et al. 2020. CloneSifter: enrichment of rare clones from heterogeneous cell populations. *BMC Biol.* 18:177

Feng J, DeWitt WS 3rd, McKenna A, Simon N, Willis AD, Matsen FA IV. 2021. Estimation of cell lineage trees by maximum-likelihood phylogenetics. *Ann. Appl. Stat.* 15:343–62

Fennell KA, Vassiliadis D, Lam EYN, Martelotto LG, Balic JJ, et al. 2022. Non-genetic determinants of malignant clonal fitness at single-cell resolution. *Nature* 601(7891):125–31

Flavahan WA, Gaskell E, Bernstein BE. 2017. Epigenetic plasticity and the hallmarks of cancer. *Science* 357(6348):aal2380

Foggetti G, Li C, Cai H, Hellyer JA, Lin W, et al. 2021. Genetic determinants of EGFR-driven lung cancer growth and therapeutic response in vivo. *Cancer Discov.* 11(7):1736–53

Forrow A, Schiebinger G. 2021. LineageOT is a unified framework for lineage tracing and trajectory inference. *Nat. Commun.* 12:4940

Frieda KL, Linton JM, Hormoz S, Choi J, Chow KK, et al. 2017. Synthetic recording and in situ readout of lineage information in single cells. *Nature* 541(7635):107–11

Gabbutt C, Schenck RO, Weisenberger DJ, Kimberley C, Berner A, et al. 2022. Fluctuating methylation clocks for cell lineage tracing at high temporal resolution in human tissues. *Nat. Biotechnol.* 40(5):720–30

Ganesh K, Massagué J. 2021. Targeting metastatic cancer. *Nat. Med.* 27(1):34–44

Gao R, Bai S, Henderson YC, Lin Y, Schalck A, et al. 2021. Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. *Nat. Biotechnol.* 39(5):599–608

Garcia-Marques J, Epinosa-Medina I, Ku K, Yang C, Koyama M, et al. 2020. A programmable sequence of reporters for lineage analysis. *Nat. Neurosci.* 23(12):1618–28

Geisinger JM, Stearns T. 2020. CRISPR/Cas9 treatment causes extended TP53-dependent cell cycle arrest in human cells. *Nucleic Acids Res.* 48(16):9067–81

Gerlinger M, McGranahan N, Dewhurst SM, Burrel RA, Tomlinson I, Swanton C. 2014. Cancer: evolution within a lifetime. *Annu. Rev. Genet.* 48:215–36

Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, et al. 2012. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* 366(10):883–92

Gerrits A, Dykstra B, Kalmykowa OJ, Klauke K, Verovskaya E, et al. 2010. Cellular barcoding tool for clonal analysis in the hematopoietic system. *Blood* 115(13):2610–18

Gerstung M, Jolly C, Leshchiner I, Dentro SC, Gonzalez S, et al. 2020. The evolutionary history of 2,658 cancers. *Nature* 578(7793):122–28

Gong W, Granados AA, Hu J, Jones MG, Raz O, et al. 2021. Benchmarked approaches for reconstruction of in vitro cell lineages and in silico models of *C. elegans* and *M. musculus* developmental trees. *Cell Syst.* 12(8):810–26.e4

Gong W, Kim HJ, Garry DJ, Kwak I. 2022. Single cell lineage reconstruction using distance-based algorithms and the R package, DCLEAR. *BMC Bioinformat*. 23:103

Granja JM, Klemm S, McGinnis LM, Kathiria AS, Mezger A, et al. 2019. Single-cell multiomic analysis identifies regulatory programs in mixed-phenotype acute leukemia. *Nat. Biotechnol.* 37(12):1458–65

Greaves M, Maley CC. 2012. Clonal evolution in cancer. *Nature* 481(7381):306–13

Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, et al. 2007. Patterns of somatic mutation in human cancer genomes. *Nature* 446(7132):153–58

Grüner BM, Schulze CJ, Yang D, Ogasawara D, Dix MM, et al. 2016. An in vivo multiplexed small-molecule screening platform. *Nat. Methods* 13(10):883–89

Gutierrez C, Al'Khafaji AM, Brenner E, Johson KE, Gohil SH, et al. 2021. Multifunctional barcoding with ClonMapper enables high-resolution study of clonal dynamics during tumor evolution and treatment. *Nat. Cancer* 2(7):758–72

Gutierrez C, Vilas CK, Wu CJ, Al'Khafaji AM. 2022. Functionalized lineage tracing can enable the development of homogenization-based therapeutic strategies in cancer. *Front. Immunol.* 13:859032

Hanahan D, Weinberg RA. 2011. Hallmarks of cancer: the next generation. *Cell* 144(5):646–74

Hata AN, Niederst MJ, Archibald HL, Gomez-Caraballo M, Siddiqui FM, et al. 2016. Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. *Nat. Med.* 22(3):262–69

He L, Li Y, Li Y, Pu W, Huang X, et al. 2017. Enhancing the precision of genetic lineage tracing using dual recombinases. *Nat. Med.* 23(12):1488–98

He Z, Maynard A, Jain A, Gerber T, Petri R, et al. 2022. Lineage recording in human cerebral organoids. *Nat. Methods* 19:90–99

Hirrlinger J, Scheller A, Hirrlinger PG, Kellert B, Tang W, et al. 2009. Split-Cre complementation indicates coincident activity of different genes in vivo. *PLOS ONE* 4:e4286

Hu Z, Li Z, Zhicheng M, Curtis C. 2020. Multi-cancer analysis of clonality and the timing of systemic spread in paired primary tumors and metastases. *Nat. Genet.* 52(7):701–8

Hussmann JA, Ling J, Ravisankar P, Yan J, Cirincione A, et al. 2021. Mapping the genetic landscape of DNA double-strand break repair. *Cell* 184(22):5653–69.e25

Hwang B, Lee W, Lee W, Yum S, Jeon Y, et al. 2019. Lineage tracing using a Cas9-deaminase barcoding system targeting endogenous L1 elements. *Nat. Commun.* 10:1234

Ihry RJ, Worringer KA, Salick MR, Frias E, Ho D, et al. 2018. p53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells. *Nat. Med.* 24(7):939–46

Jamal-Hanjani M, Quezada SA, Larkin J, Swanton C. 2015. Translational implications of tumor heterogeneity. *Clin. Cancer Res.* 21(6):1258–66

Jin X, Demere Z, Nair K, Ali A, Ferraro G, et al. 2020. A metastasis map of human cancer cell lines. *Nature* 588(7837):331–36

Jones MG, Khodaverdian A, Quinn JJ, Chan MM, Hussman JA, et al. 2020. Inference of single-cell phylogenies from lineage tracing data using Cassiopeia. *Genome Biol.* 21:92

Jones MG, Rosen Y, Yosef N. 2022. Interactive, integrated analysis of single-cell transcriptomic and phylogenetic data with PhyloVision. *Cell Rep. Methods* 2(4):100200

Jones S, Chen W, Parmigiani G, Diehl F, Beerenwinkel N, et al. 2008. Comparative lesion sequencing provides insights into tumor evolution. *PNAS* 105(11):4283–88

Kalhor R, Mali P, Church GM. 2017. Rapidly evolving homing CRISPR barcodes. *Nat. Methods* 14(2):195–200

Kim C, Gao R, Sei E, Brandt R, Hartman J, et al. 2018. Chemoresistance evolution in triple-negative breast cancer delineated by single-cell sequencing. *Cell* 173(4):879–93.e13

Kretzschmar K, Watt FM. 2012. Lineage TRACING. *Cell* 148(1–2):33–45

LaFave LM, Kartha VK, Ma S, Meli K, Del Priore I, et al. 2020. Epigenomic state transitions characterize tumor progression in mouse lung adenocarcinoma. *Cancer Cell* 38(2):212–28.e13

Lamprecht S, Schmidt EM, Blaj C, Hermeking H, Jung A, et al. 2017. Multicolor lineage tracing reveals clonal architecture and dynamics in colon cancer. *Nat. Commun.* 8:1406

Le Magnen C, Shen MM, Abate-Shen C. 2018. Lineage plasticity in cancer progression and treatment. *Annu. Rev. Cancer Biol.* 2:271–89

Li C, Lin W, Rizvi H, Cai H, McFarland CD, et al. 2021. Quantitative in vivo analyses reveal a complex pharmacogenomic landscape in lung adenocarcinoma. *Cancer Res.* 81(17):4570–80

Li Y, Lv Z, Zhang S, Wang Z, He L, et al. 2020. Genetic fate mapping of transient cell fate reveals N-cadherin activity and function in tumor metastasis. *Dev. Cell* 54(5):593–607.e5

Lin D, Li X, Park P, Tang B, Shen H, et al. 2021. Time-tagged ticker tapes for intracellular recordings. bioRxiv 10.1101/2021.10.13.463862. **https://doi.org/10.1101/2021.10.13.463862**

Liu K, Deng S, Ye C, Yao Z, Wang J, et al. 2021. Mapping single-cell-resolution cell phylogeny reveals cell population dynamics during organ development. *Nat. Methods* 18(12):1506–14

Liu K, Jin H, Zhou B. 2020. Genetic lineage tracing with multiple DNA recombinases: a user's guide for conducting more precise cell fate mapping studies. *J. Biol. Chem.* 295(19):6413–24

Liu Q, Liu K, Cui G, Huang X, Yao S, et al. 2019. Lung regeneration by multipotent stem cells residing at the bronchioalveolar-duct junction. *Nat. Genet.* 51(4):728–38

Livet J, Weissman TA, Kang H, Draft RW, Lu J, et al. 2007. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* 450(7166):56–62

Loveless TB, Carlson CK, Hu VJ, Helmy CAD, Liang G, et al. 2021a. Molecular recording of sequential cellular events into DNA. bioRxiv 10.1101/2021.11.05.467507. **https://doi.org/10.1101/2021.11.05.467507**

Loveless TB, Grotts JH, Schecter MW, Forouzmand E, Carlson CK, et al. 2021b. Lineage tracing and analog recording in mammalian cells by single-site DNA writing. *Nat. Chem. Biol.* 17(6):739–47

Ludwig LS, Lareau CA, Ulirsch JC, Christian E, Muus C, et al. 2019. Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. *Cell* 176(6):1325–39.e22

Lüönd F, Sugiyama N, Bill R, Bornes L, Hager C, et al. 2021. Distinct contributions of partial and full EMT to breast cancer malignancy. *Dev. Cell* 56(23):3203–21.e11

Maddipati R, Stranger BZ. 2015. Pancreatic cancer metastases harbor evidence of polyclonality. *Cancer Discov.* 5(10):1086–97

Marjanovic ND, Hofree M, Chan JE, Canner D, Wu K, et al. 2020. Emergence of a high-plasticity cell state during lung cancer evolution. *Cancer Cell* 38(2):229–46.e13

Marusyk A, Polyak K. 2010. Tumor heterogeneity: causes and consequences. *Biochim. Biophys. Acta Rev. Cancer* 1805:105–17

McGranahan N, Swanton C. 2017. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell* 168(4):613–28

McKenna A, Findlay GM, Gagnon JA, Horwitz MS, Schier AF, Shendure J. 2016. Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science* 353(6298):aaf7907

McKenna A, Gangon JA. 2019. Recording development with single cell dynamic lineage tracing. *Development* 146(12):dev169730

Merino D, Weber TS, Serrano A, Vaillant F, Liu K, et al. 2019. Barcoding reveals complex clonal behavior in patient-derived xenografts of metastatic triple negative breast cancer. *Nat. Commun.* 10:766

Michlits G, Hubmann M, Wu S, Vainorius G, Boudusan E, et al. 2017. CRISPR-UMI: single-cell lineage tracing of pooled CRISPR–Cas9 screens. *Nat. Methods* 14(12):1191–97

Minussi DC, Nicholson MD, Ye H, Davis A, Wang K, et al. 2021. Breast tumours maintain a reservoir of subclonal diversity during expansion. *Nature* 592(7853):302–8

Moffitt JR, Hao J, Wang G, Chen KH, Babcok HP, Zhuang X. 2016. High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. *PNAS* 113(39):11046–51

Nagy A. 2000. Cre recombinase: the universal reagent for genome tailoring. *Genesis* 26(2):99–109

Navin N, Kendall J, Troge J, Andrews P, Rodgers L, et al. 2011. Tumour evolution inferred by single-cell sequencing. *Nature* 472(7341):90–94

Navin NE, Hicks J. 2010. Tracing the tumor lineage. *Mol. Oncol.* 4(3):267–83

Neftel C, Laffy J, Filbin MG, Hara T, Shore ME, et al. 2019. An integrative model of cellular states, plasticity, and genetics for glioblastoma. *Cell* 178(4):835–49.e21

Nguyen LV, Pellacani D, Lefort S, Kannan N, Osako T, et al. 2015. Barcoding reveals complex clonal dynamics of de novo transformed human mammary cells. *Nature* 528(7581):267–71

Nowell PC. 1976. The clonal evolution of tumor cell populations. *Science* 194(4260):23–28
Ogbeide S, Giannese F, Mincarelli L, Macaulay IC. 2022. Into the multiverse: advances in single-cell multiomic profiling. *Trends Genet*. 38(8):831–43
Oren Y, Tsabar M, Cuoco MS, Amir-Zilberstein L, Cabanos HF, et al. 2021. Cycling cancer persister cells arise from lineages with distinct programs. *Nature* 596(7873):576–82
Ouardini K, Lopez R, Jones MG, Prillo S, Zhang R, et al. 2021. Reconstructing unobserved cellular states from paired single-cell lineage tracing and transcriptomics data. bioRxiv 10.1101/2021.05.28.446021. **https://doi.org/10.1101/2021.05.28.446021**
Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, et al. 2014. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 344(6190):1396–401
Pei W, Feyerabend TB, Rössler J, Wang X, Postrach D, et al. 2017. Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. *Nature* 548(7668):456–60
Penter L, Gohil SH, Lareau C, Ludwig LS, Parry EM, et al. 2021a. Longitudinal single-cell dynamics of chromatin accessibility and mitochondrial mutations in chronic lymphocytic leukemia mirror disease history. *Cancer Discov*. 11(12):3048–63
Penter L, Gohil SH, Wu CJ. 2021b. Natural barcodes for longitudinal single cell tracking of leukemic and immune cell dynamics. *Front. Immunol.* 12:788891
Perli SD, Cui CH, Lu TK. 2016. Continuous genetic recording with self-targeting CRISPR-Cas in human cells. *Science* 353(6304):aag0511
Pierce SE, Granja JM, Greenleaf WJ. 2021. High-throughput single-cell chromatin accessibility CRISPR screens enable unbiased identification of regulatory networks in cancer. *Nat. Commun.* 12:2969
Powles T, Assaf ZJ, Davarpanah N, Bancheraeu R, Szabados B, et al. 2021. ctDNA guiding adjuvant immunotherapy in urothelial carcinoma. *Nature* 595:432–37
Quinn JJ, Jones MG, Okomito RA, Najo S, Chan MM, et al. 2021. Single-cell lineages reveal the rates, routes, and drivers of metastasis in cancer xenografts. *Science* 371(6532):abc1944
Quintanal-Villalonga A, Chan JM, Yu HA, Pe'er D, Sawyers CL, et al. 2020. Lineage plasticity in cancer: a shared pathway of therapeutic resistance. *Nat. Rev. Clin. Oncol.* 17(6):360–71
Raj B, Wagner DE, McKenna A, Pandey S, Klein A, et al. 2018. Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. *Nat. Biotechnol.* 36(5):442–50
Reeves MQ, Kandyba E, Harris S, Del Rosario R, Balmain A. 2018. Multicolour lineage tracing reveals clonal dynamics of squamous carcinoma evolution from initiation to metastasis. *Nat. Cell Biol.* 20(6):699–709
Rios AC, Capaldo BD, Vaillant F, Pal B, van Ineveld R, et al. 2019. Intraclonal plasticity in mammary tumors revealed through large-scale single-cell resolution 3D imaging. *Cancer Cell* 35(6):953
Rogers ZN, McFarland CD, Winters IP, Naranjo S, Chuang C, et al. 2017. A quantitative and multiplexed approach to uncover the fitness landscape of tumor suppression in vivo. *Nat. Methods* 14(7):737–42
Roh V, Abramowski P, Hiou-Feige A, Cornils K, Rivals J, et al. 2018. Cellular barcoding identifies clonal substitution as a hallmark of local recurrence in a surgical model of head and neck squamous cell carcinoma. *Cell Rep*. 25(8):2208–22.e7
Rovira-Clavé X, Drainas AP, Jiang S, Bai Y, Baron M, et al. 2021. Spatial epitope barcoding reveals subclonal tumor patch behaviors. bioRxiv 10.1101/2021.06.29.449991. **https://doi.org/10.1101/2021.06.29.449991**
Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4(4):406–25
Salehi S, Kabeer F, Ceglia N, Andronesu M, Williams M, et al. 2021. Clonal fitness inferred from time-series modelling of single-cell cancer genomes. *Nature* 595(7868):585–90
Salvador-Martínez I, Grillo M, Averof M, Telford MJ. 2019. Is it possible to reconstruct an accurate cell lineage using CRISPR recorders? *eLife* 8:e40292
Salvador-Martínez I, Grillo M, Averof M, Telford MJ. 2021. CeLaVi: an interactive cell lineage visualization tool. *Nucleic Acids Res*. 49(W1):W80–85
Schepers AG, Snipper HJ, Stange DE, van den Born M, van Es JH, et al. 2012. Lineage tracing reveals $Lgr5^+$ stem cell activity in mouse intestinal adenomas. *Science* 337(6095):730–35
Schepers K, Swart E, van Heijst JWJ, Gerlach C, Castrucci M, et al. 2008. Dissecting T cell lineage relationships by cellular barcoding. *J. Exp. Med.* 205(10):2309–18

Schmidt F, Cherepkova MY, Platt RJ. 2018. Transcriptional recording by CRISPR spacer acquisition from RNA. *Nature* 562(7727):380–85

Schmierer B, Botla SK, Zhang J, Turnen M, Kivioja T, Taipale J. 2017. CRISPR/Cas9 screening using unique molecular identifiers. *Mol. Syst. Biol.* 13(10):945

Shah SP, Morin RD, Khattra J, Prentice L, Puge T, et al. 2009. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* 461(7265):809–13

Shipman SL, Nivala J, Macklis JD, Church GM. 2016. Molecular recordings by directed CRISPR spacer acquisition. *Science* 353(6298):aaf1175

Simeonov KP, Byrns CN, Clark ML, Norgard RJ, Martin B, et al. 2021. Single-cell lineage tracing of metastatic cancer reveals selection of hybrid EMT states. *Cancer Cell* 39(8):1150–62.e9

Snippert HJ, van der Flier LG, Sato T, van Es JH, van den Born M, et al. 2010. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* 143(1):134–44

Sottoriva A, Kang H, Ma Z, Graham T, Salomon MP, et al. 2015. A big bang model of human colorectal tumor growth. *Nat. Genet.* 47(3):209–16

Spanjaard B, Hu B, Mitic N, Olivares-Chauvet P, Janjuha S, et al. 2018. Simultaneous lineage tracing and cell-type identification using CRISPR–Cas9-induced genetic scars. *Nat. Biotechnol.* 36(5):469–73

Ståhl PL, Salmén F, Vickovic S, Lundmark A, Navaro JF, et al. 2016. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* 353(6294):78–82

Stickels RR, Murray E, Kuman P, Li J, Marshall JL, et al. 2021. Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nat. Biotechnol.* 39(3):313–19

Tanay A, Regev A. 2017. Scaling single-cell genomics from phenomenology to mechanism. *Nature* 541(7637):331–38

Tang W, Liu DR. 2018. Rewritable multi-event analog recording in bacterial and mammalian cells. *Science* 360(6385):6385

Tao L, Raz O, Marx Z, Ghosh MS, Huber S, Greindl-Junghans J, et al. 2021. Retrospective cell lineage reconstruction in humans by using short tandem repeats. *Cell Rep. Methods* 1(3):100054

Tarabichi M, Salcedo A, Deshwar AG, Leathlobhair M, Wintersinger J, et al. 2021. A practical guide to cancer subclonal reconstruction from DNA sequencing. *Nat. Methods* 18(2):144–55

Torborg SR, Li Z, Chan JE, Tammela T. 2022. Cellular and molecular mechanisms of plasticity in cancer. *Trends Cancer Res.* 8(9):735–46

Tulpule A, Bivona TG. 2020. Acquired resistance in lung cancer. *Annu. Rev. Cancer Biol.* 4:279–97

Turajlic S, Sottoriva A, Graham T, Swanton C. 2019. Resolving genetic heterogeneity in cancer. *Nat. Rev. Genet.* 20(7):404–16

Turajlic S, Xu H, Litchfield K, Rowan A, Chambers T, et al. 2018. Tracking cancer evolution reveals constrained routes to metastases: TRACERx renal. *Cell* 173(3):581–94.e12

Umkehrer C, Holstein F, Formenti L, Jude J, Froussious K, et al. 2021. Isolating live cell clones from barcoded populations using CRISPRa-inducible reporters. *Nat. Biotechnol.* 39(2):174–78

VanHorn S, Morris SA. 2021. Next-generation lineage tracing and fate mapping to interrogate development. *Dev. Cell* 56:7–21

Vendramin R, Litchfield K, Swanton C. 2021. Cancer evolution: Darwin and beyond. *EMBO J.* 40(18):e108389

Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Dias LA, Kinzler KW. 2013. Cancer genome landscapes. *Science* 339(6127):1546–58

Wagner DE, Klein AM. 2020. Lineage tracing meets single-cell omics: opportunities and challenges. *Nat. Rev. Genet.* 21:410–27

Wagner DE, Weinreb C, Collins ZM, Briggs JA, Megason SG, Klein AM. 2018. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* 360(6392):981–87

Walens A, Lin J, Damrauer JS, McKinney B, Lupo R, et al. 2020. Adaptation and selection shape clonal evolution of tumors during residual disease and recurrence. *Nat. Commun.* 11:5017

Walsh C, Cepko CL. 1992. Widespread dispersion of neuronal clones across functional regions of the cerebral cortex. *Science* 255(5043):434–40

Wang R, Zhang R, Khodaveridan A, Yosef N. 2021. Theoretical guarantees for phylogeny inference from single-cell lineage tracing. bioRxiv 10.1101/2021.11.21.469464. **https://doi.org/10.1101/2021.11.21.469464**

Wang S, Herriges MJ, Hurley K, Kotton DN, Klein AM. 2022. CoSpar identifies early cell fate biases from single-cell transcriptomic and lineage information. *Nat. Biotechnol.* 40:1066–77

Weinreb C, Rodriguez-Fraticelli A, Camargo FD, Klein AM. 2020. Lineage tracing on transcriptional landscapes links state to fate during differentiation. *Science* 367(6479):aaw3381

Weissman TA, Pan YA. 2015. Brainbow: new resources and emerging biological applications for multicolor genetic labeling and analysis. *Genetics* 199(2):293–306

Williams MJ, Werner B, Heide T, Curtis C, Barnes C, et al. 2018. Quantification of subclonal selection in cancer from bulk sequencing data. *Nat. Genet.* 50(6):895–903

Winters IP, Chiou S, Paulk NK, McFarland CD, Lalgudi PV, et al. 2017. Multiplexed in vivo homology-directed repair and tumor barcoding enables parallel quantification of Kras variant oncogenicity. *Nat. Commun.* 8:2053

Woodworth MB, Girskis KM, Walsh CA. 2017. Building a lineage from single cells: genetic techniques for cell lineage tracking. *Nat. Rev. Genet.* 18(4):230–44

Wroblewska A, Dhainaut M, Ben-Zvi B, Rose SA, Park ES, et al. 2018. Protein barcodes enable high-dimensional single-cell CRISPR screens. *Cell* 175(4):1141–55.e16

Yachida S, Jones S, Bozic I, Antal T, Leary R, et al. 2010. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* 467(7319):1114–17

Yang D, Jones MG, Naranjo S, Rideout WM 3rd, Min KH, et al. 2022. Lineage tracing reveals the phylodynamics, plasticity, and paths of tumor evolution. *Cell* 185(11):1905–23.e25

Yang J, Antin P, Berx G, Blanpain C, Brabletz T, et al. 2020. Guidelines and definitions for research on epithelial-mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* 21(6):341–52

Yao Z, Mich JK, Ku S, Menon V, Krostag A, et al. 2017. A single-cell roadmap of lineage bifurcation in human ESC models of embryonic brain development. *Cell Stem Cell* 20:120–34

Yates LR, Knappskog S, Wedge D, Ramery JHR, Gonzalez S, et al. 2017. Genomic evolution of breast cancer metastasis and relapse. *Cancer Cell* 32(2):169–84.e7

Yu M, Bardia A, Wittner BS, Sott SL, Smas ME, et al. 2013. Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science* 339(6119):580–84

Zafar H, Lin C, Bar-Joseph Z. 2020. Single-cell lineage tracing by integrating CRISPR-Cas9 mutations with transcriptomic data. *Nat. Commun.* 11:3055

Zhang W, Bado IL, Hu J, Wan Y, Wu L, et al. 2021. The bone microenvironment invigorates metastatic seeds for further dissemination. *Cell* 184(9):2471–86.e20