nature neuroscience

Review article

https://doi.org/10.1038/s41593-024-01827-9

Applying single-cell and single-nucleus genomics to studies of cellular heterogeneity and cell fate transitions in the nervous system

Received: 5 December 2023

Accepted: 22 October 2024

Published online: 3 December 2024

Check for updates

Igor Adameyko[1,2], Trygve Bakken[3], Aparna Bhaduri[4], Chintan Chhatbar[5], Mariella G. Filbin[6], David Gate[7], Hannah Hochgerner[8], Chang Nam Kim[9,10], Jordan Krull[11,12], Gioele La Manno[13], Qingyun Li[14,15], Sten Linnarsson[16], Qin Ma[11,12], Christian Mayer[17], Vilas Menon[18], Patricia Nano[4], Marco Prinz[5,19], Steve Quake[20,21,22], Christopher A. Walsh[23,24,25], Jin Yang[14,15], Omer Ali Bayraktar[26] ✉, Ozgun Gokce[27,28] ✉, Naomi Habib[29] ✉, Genevieve Konopka[30,31] ✉, Shane A. Liddelow[32,33,34,35] ✉ & Tomasz J. Nowakowski[9,10,36,37,38] ✉

Single-cell and single-nucleus genomic approaches can provide unbiased and multimodal insights. Here, we discuss what constitutes a molecular cell atlas and how to leverage single-cell omics data to generate hypotheses and gain insights into cell transitions in development and disease of the nervous system. We share points of reflection on what to consider during study design and implementation as well as limitations and pitfalls.

The nervous system's diverse cell types challenge our understanding of the brain activity that underlies mental experiences, behavior and diseases. Single-cell and single-nucleus genomics is crucial for discovering cell diversity and gaining insights into neurodevelopment, disease and brain evolution. These techniques must be carefully applied to avoid technical and biological artifacts and detect rare cell populations while providing enough depth to uncover diverse biological states. Along with refs. 1 and 2, we outline strategies for creating genomic cell atlases, which are essential for developing new hypotheses and tools for delivering genes and chemicals to specific cells.

Generation of a molecular cell atlas

Successful nervous system atlas generation requires a multiscale, multimodal approach to fully capture the complexity of cell types and states. Although this is true in other organ systems, the cellular and transcriptomic diversity of the nervous system is particularly robust, challenging atlasing endeavors. We present a consensus-driven roadmap for creating detailed brain atlases that reflect the developmental, functional and evolutionary aspects of brain cells in both healthy and diseased conditions, including variations in health status, age, gender and ethnicity (Fig. 1).

Initial single-cell transcriptomic atlases, including large consortium efforts such as the Brain Initiative Cell Census Network[3] and the Human Cell Atlas[4], revolutionized our understanding of brain cellular diversity[5–11]. However, questions remain about the completeness and suitability of omic classifications in depicting functional diversity and their role in brain function. New omic technologies add features and functional axes for grouping cells, posing a challenge for defining cell types. Initial studies provide a preliminary framework for mapping cellular function and its relevance to brain health. However, transcriptional atlases based on dissociated cells alone are insufficient for a comprehensive cell atlas. The complexity of the central nervous system (CNS) requires a multimodal approach to understand how chemical and electrical signals shape brain function. Effective characterization demands spatial resolution[1] and investigation across pathological, developmental and evolutionary aspects. Here we address the starting point of this ambitious journey with omic atlases, identifying opportunities and consensus concepts while addressing areas that require clarification for effective cell typing in future brain atlases, enabling new discoveries.

What should a cell type atlas achieve, and when should a study use the term 'cell atlas'? A detailed cell atlas, which is essential for understanding CNS health and disease, must have single-cell resolution

A full list of affiliations appears at the end of the paper. ✉e-mail: ob5@sanger.ac.uk; ozgun.goekce@ukbonn.de; naomi.habib@mail.huji.ac.il; genevieve.konopka@utsouthwestern.edu; shane.liddelow@nyulangone.org; tomasz.nowakowski@ucsf.edu

https://doi.org/10.1038/s41593-024-01827-9

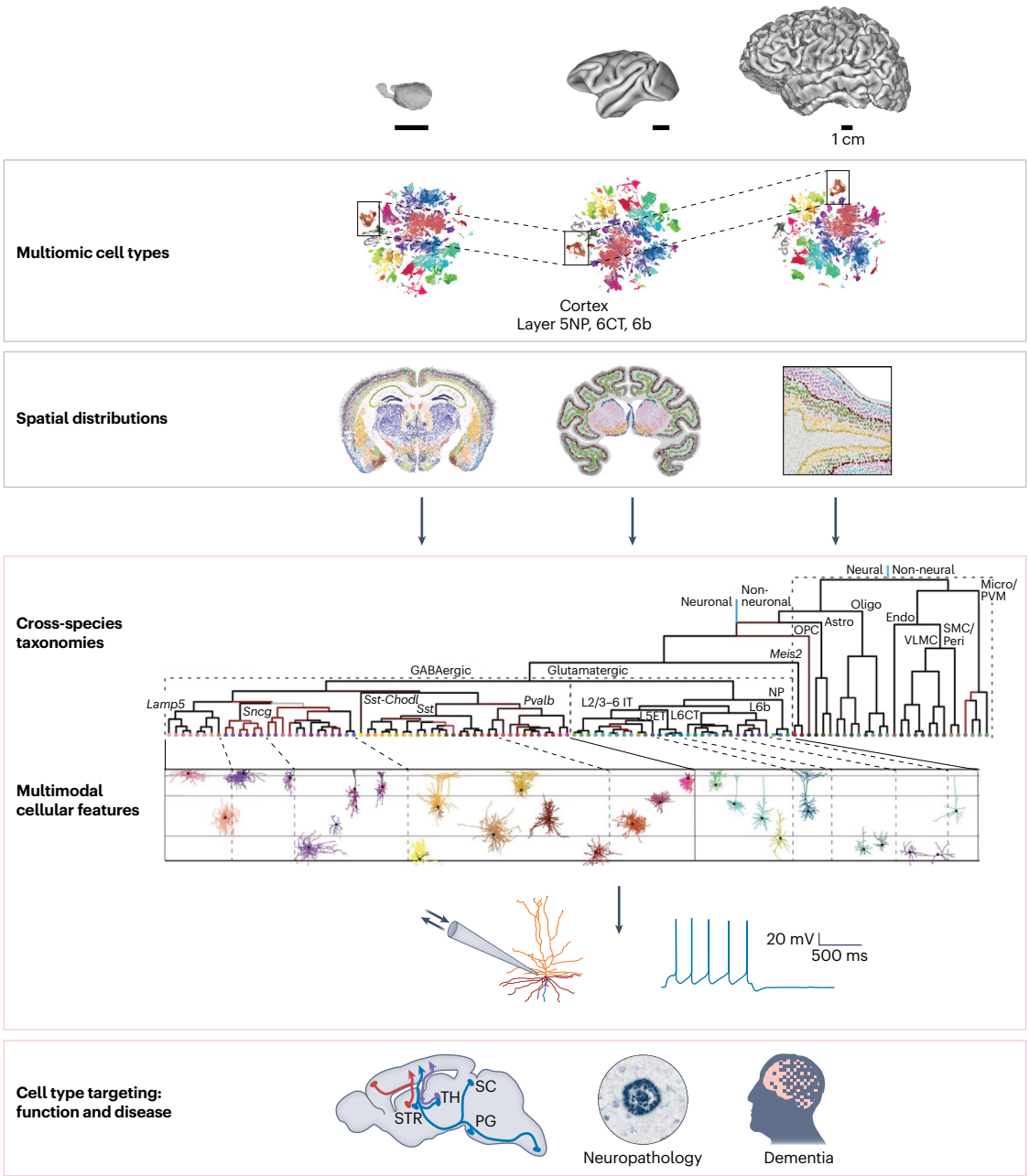

Fig. 1 | Schematics for building a multimodal molecular atlas. Single-cell and single-nucleus profiling of brain tissue from humans and other mammals establishes a foundation for describing cellular diversity. UMAP plots group cells or nuclei by transcriptomic similarity (colors) and align them across species by conserved gene expression. Spatial distributions are described using in situ marker gene measurements. Integrating multiomic and spatial data creates a hierarchical, multimodal cell type atlas. Methods include Patch-seq, long-range neuronal projection tracing and viral tracing. Single-cell genomic data predict cell–cell communication through ligand–receptor pairs. This multimodal description aids understanding of circuit function and dysfunction. Additionally, cell type-specific epigenomic features can be used to develop tools for targeting cell populations for experimental perturbation and potential treatment of brain diseases. Astro, astrocyte; oligo, oligodendrocyte; OPC, oligodendrocyte progenitor cell; PVM, perivascular macrophage; SMC, smooth muscle cell; VLMC, vascular and leptomeningeal cell; peri, pericyte; micro, microglia; NP, near-projecting; L6CT, layer 6 corticothalamic; L5ET, layer 5 extratelencephalic; IT, intratelencephalic; SC, superior colliculus; PG, pontine gray; TH, thalamus; STR, striatum; 5NP, layer 5 near-projecting. Images for cross-species taxonomies and multimodal cellular features adapted from ref. 153, Springer Nature Limited.

to map cell types and states under various conditions. With increasing cell atlas publications, standardizing annotation is crucial. We advocate for a consensus ontogeny-based nomenclature[12], grounded in lineage histories, molecular states and existing literature, which will emphasize reproducibility and data integrity, fostering scientific progress. A robust atlas must demonstrate reproducibility (stability with new data), integrity (accurate representation of all cell types) and high predictive value (defining a cell within a population or as a new subset)[4]. Reproducibility must transcend individual variability, and it is crucial for genetically diverse human cohorts. Creating a brain atlas demands a systematic approach with high-resolution mapping and quality control (Table 1) and should include one or more technologies that can assess a cell's transcriptome, proteome, epigenome and morphology, pinpointing these elements to precise anatomical locations.

Table 1 | Considerations when designing a molecular atlas

Considerations	Key questions	Examples and resources
Scope of the study	What is the focus? Which samples?	Large atlas efforts[4,5,10,11], targeted regions[7,32,37], cell types[15,29], disorders[15,142], cell type evolution[3,24]
Methods	What methods to use?	scRNA-seq[11], snRNA-seq[142], single-cell epigenomics[48,49], multiome[14,143]
Cell sampling	Sex/ancestry/age breakdown? Sample meta-data availability?	Dissociation methods[66], reporting[144]
Experimental design	Multiplexing? Integration with other methods?	Multiplexing for lowering cost[145] or complementary assays[13], spatial methods[20]
Computational analysis	How will the bioinformatician be trained? What are the computational resources? What analysis tools and data-sharing mechanisms will be used?	Current best practices in single-cell analysis[146]

Techniques such as Patch-seq exemplify the methods needed to bridge omics and function–cell phenotypes[13].

A cell atlas should encompass cell type diversity and capture molecular and functional traits, starting with single-cell transcriptomics. These atlases reveal how cells react to development, aging, disease and sex differences, and they aid in experimental design. Cross-species genomic atlases reveal conserved cell types and pathways, although species-specific findings need extensive comparisons. Artificial intelligence models further enhance cross-modal mapping. While RNA profiling has advanced brain mapping, it is essential to integrate proteins, epigenetics and morphology, as RNA alone cannot capture all features of cell variability. Techniques like cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) improve the identification of cell subsets[14]. A comprehensive atlas establishes a baseline of healthy cell type states for robust disease condition studies, aiding in discovering mechanisms, predictive models and drug targets.

Hypothesis generation from single-cell data

Single-cell and single-nucleus genomics data are more than an inventory of cell types; they generate new hypotheses tested with other neuroscience methods, as any novel claims require follow-up experiments[2]. Single-cell and single-nucleus genomics have revolutionized the study of the nervous system by pinpointing specific cells and molecular pathways and contributing new research questions, such as the roles of glial and immune cells and cell type-specific vulnerabilities in aging, neurodegeneration and psychiatric diseases[15–20]. A hypothesis-generating approach contrasts with traditional methods but facilitates hypothesis-driven research. Historically, studying changes that occur in brain tissue was challenging owing to its complexity and heterogeneity. However, single-cell and single-nucleus RNA sequencing (scRNA-seq and snRNA-seq) technologies now allow large-scale comparisons of brain cell states, providing the resolution required to dissect cell type-specific responses to perturbation[21–23]. These studies show unique molecular signatures for brain cell types, which are useful for classification and function cues[24]. The high-throughput nature of scRNA-seq and snRNA-seq enables rapid assessment of genome-wide gene expression and regulation in all cell types. Thus, alterations in behavior, development or disease can be profiled at single-cell resolution to build hypotheses about the key genes and/or cell types that represent each altered state. When combined with cell type state-specific perturbations, scRNA-seq and snRNA-seq provide information that leads to many novel hypotheses.

Cell transitions in health and disease

A cell transition refers to changes in a cell's molecular and/or functional state[25]. In the nervous system, cell transitions can occur physiologically, for instance, during development and memory formation, or in response to disease, infection or trauma[26–30]. Changes in gene expression patterns indicate shifts in cell state or responses to external stimuli. For example, in Alzheimer's disease, scRNA-seq and snRNA-seq showed changes in microglial gene expression, leading to immune system activation in the brain, which suggests that there might be therapeutic potential in manipulating microglial transitions[15,31]. In Parkinson's disease, snRNA-seq of the substantia nigra revealed gene expression changes in mitochondrial function, oxidative stress and inflammatory processes[32]. In Huntington's disease, snRNA-seq identified gene expression changes in striatal neurons, revealing alterations in cell cycle regulation, transcriptional regulation and synaptic function. scRNA-seq and snRNA-seq generate hypotheses about cell transitions in response to injury or stress[33].

Appropriate methodologies for analyzing different types of biological variation are discussed in ref. 1. In brief, discrete models (for example, cell clustering) applied to nervous system scRNA-seq and snRNA-seq datasets revealed a staggering diversity of cell subtypes, especially neurons. The first step is resolving unique cell types with accurate annotation, often using unsupervised clustering methods. A major challenge is determining the optimal number of cell clusters and their precise annotation. This typically involves a search across the parameter space, including factors such as the number of cells, genes used for clustering and clustering thresholds. These decisions require rigorous assessment and validation to avoid issues such as overclustering or misclassification (as elaborated in ref. 2). An approach to cell type annotation should be dynamic, adapting to experimental parameters and leveraging a priori knowledge to supervise the process. This flexibility accommodates the complexity and diversity of cellular states and lineages in different studies. Annotations rely on prior observations, necessitating careful evidence selection. Automated annotation methods using enriched gene sets from prior literature can be misleading. Manual inspection and careful gene comparisons often yield the clearest results, provided there is no bias[1,2]. Beyond discrete clusters, cells show continuous variations due to developmental history, location and stimulus response, which are essential for nervous system homeostasis. Glial cells, for instance, flexibly alter their identities, forming a gene expression continuum[29,34–37]. This necessitates analytical methods that account for continuous cellular changes beyond discrete cell types.

The large scRNA-seq and snRNA-seq atlases of the mouse, marmoset and human brain across developmental time points[1,2,24,38,39] are invaluable for studying how cell fates change across developmental and disease states and during evolution. For example, scRNA-seq and snRNA-seq approaches have enhanced understanding of cortical arealization by revealing area-specific transcriptional signatures throughout development and differentiation trajectories[40–42]. The data from these atlases further suggest that many CNS cell types exist along a transcriptional continuum: distinct cell clusters can be defined in the adult and developing brain, but the boundaries between these cell subsets are often unclear and reflect continuous transitions of cell transcriptional programs, especially in development and disease.

As cells often do not progress in perfect synchrony, a typical scRNA-seq and snRNA-seq dataset may capture intermediate states along with stable clusters, which allows trajectory inference or pseudotime analysis in silico[43]. This information can generate hypotheses about cell state progression, cell end fate and gene regulatory mechanisms underlying lineage bifurcations. As available tools may yield vastly different paths for the same dataset, choosing the appropriate tools should be a primary consideration, as discussed further in this Review and refs. 1,2.

Inference of regulatory mechanisms driving cell state transitions

Subtle transcriptional changes are driven by signaling changes and can reveal nuanced disease mechanisms. A promising future direction is dissecting how these tightly controlled transcriptomic differences might underlie the emergence of autism spectrum disorder (ASD)[44], schizophrenia[45] and other complex disorders[46,47]. Unlike bulk transcriptomic profiles, scRNA-seq and snRNA-seq approaches may provide more insight into how disruptive gene expression leads to phenotypes, especially with the wealth of neurotypical scRNA-seq and snRNA-seq transcriptomic reference profiles. Yet transcriptomic atlases alone are insufficient to reveal mechanisms that underlie developmental and disease contexts. This can be addressed in part by single-nucleus epigenomics, which profiles the gene regulatory elements that can describe cell states. For example, epigenomic area-specific changes in intermediate neuronal progenitors precede transcriptomic changes[48], and ASD-linked mutations influence the binding of transcription factors to key gene regulatory regions[49]. Multiomic approaches can therefore comprehensively build the gene regulatory networks that confer cell fates, and such approaches are discussed in ref. 1.

The gene regulatory networks and models generated from single-cell multiomic profiles generate hypotheses that can be tested using perturbation platforms[50,51]. Perturbation approaches build on the growing CRISPR-based, Perturb-seq[52] pipeline to activate or downregulate target genes and evaluate the resulting gene expression changes in the same cell. These approaches have been conducted in model systems, such as induced pluripotent stem cell-derived neuronal cultures and human stem cell-derived cortical organoids. For example, a Perturb-seq screen of ASD-related genes in the developing mouse brain suggested that perturbation of these genes impacts the function of both neuronal and glial cell types, with several ASD-related genes impacting a common set of gene networks[50,51]. Another example is a cell type-specific Perturb-seq screen in the brain, which was applied to interrogate microglia–astrocyte crosstalk in vivo in a multiple sclerosis preclinical mouse model[53]. Perturb-seq screens can therefore refine our understanding of the canonical gene signatures and cell types that are affected by complex neurological disorders, showing promise as a tool to identify and test mechanisms for therapeutic interventions. Such approaches are discussed in ref. 2.

Inference of cellular fate decisions and multipotency

Cell atlases raise many questions about how cells make decisions in binary- or multiple-choice situations, such as lineage development, fate selection, transition to pathological versus healthy states or phenotype switching. Transcriptional changes can be approximated as a trajectory from scRNA-seq and snRNA-seq data, with bifurcations representing moments when cells choose a path[54]. The period before the bifurcation (tipping point) is when extrinsic and intrinsic factors compete during a cell's decision making[54,55]. External signals or epigenetic priming are examples of biasing factors. For instance, neural crest progenitors exhibit competing co-activating programs before bifurcations. Single-cell and single-nucleus data can be used to determine which transcripts regulate cell type proportions before bifurcations[56]. Computational tools can analyze active regulons in pseudotime to identify transcription factors that correlate with specific future choices[56], but such predictions of causality during cellular decision making must be experimentally validated[57]. Regulon analysis can be combined with clonal analysis using DNA recombination or virally delivered barcodes to link pre-bifurcation states to future cell fates[58].

Addressing the biasing and decision-making mechanisms is important, because the integrity, homeostasis and functionality of a given tissue depend not only on the existence of specific cell types but also on the correct proportions of those cell types and their proper positioning. Perturbations of balances of biasing factors in disease or experimental setups might cause abnormalities during tissue development, self-renewal, regeneration and healing. Thus, single-cell genomics will increase understanding of how multipotency works in specific stem or progenitor cells according to individual or collective multipotency models[59].

Microenvironment and non-cell-autonomous contributors to gene expression

Given the functional impact of coordinated cellular activities in a complex tissue such as the nervous system, it is pertinent to ask what coordinated activities between unique cell types could explain a biological phenomenon. The first step is resolving unique cell types with accurate annotation, often using unsupervised clustering methods. An important issue is determining the number of cell clusters and their specific annotation, which typically involves a search across the parameter space and should involve assessments and validations to avoid overclustering and misclassifications[2]. An approach to cell type annotation should be dynamic, adapting to experimental parameters and leveraging a priori knowledge to supervise the process. This flexibility accommodates the complexity and diversity of cellular states and lineages in different studies. Annotations rely on prior observations, necessitating careful evidence selection. Automated annotation methods using enriched gene sets from prior literature can be misleading. Manual inspection and careful gene comparisons often yield the clearest results, provided there is no bias[1,2].

Physiological processes often involve communities of cells changing their abundance, connectivity or activity, which can be predicted using scRNA-seq and snRNA-seq data by observing coordinated changes in cellular abundance[27]. Additionally, cell–cell connectivity at the single-neuron level, both locally and between brain areas, can be studied using scRNA-seq and snRNA-seq[2]. Furthermore, cell–cell communication can be predicted using scRNA-seq and snRNA-seq data, presuming coupled abundance between known ligand–receptor pairs (measured through metrics such as correlation and mutual information) is associated with activity levels. However, RNA abundance plays only a stoichiometric role in signaling, and the inherent noise in scRNA-seq and snRNA-seq data risks false positives[1,2,60], so it is imperative to validate notable interactions from the results of single-cell and single-nucleus approaches using standard laboratory assays to ensure accuracy.

Genomic landscape

Genetic variants are driving variations in gene expression across individuals yet might be challenging to detect, as these changes are often cell type specific. Single-cell and single-nucleus transcriptomic technologies are sequencing platforms, not just gene-counting tools. An advantage over bulk sequencing is the direct association between variants and the cells containing them, adding substantial statistical power to identify somatic variants, including coding mutations, indels, gene fusions and copy number changes[61–64]. However, scRNA-seq and snRNA-seq face dropout issues and allelic exclusion, leading to high false negative rates. By contrast, single-cell assay for transposase-accessible chromatin with sequencing (scATAC-seq) profiles DNA mutations directly but is limited to sequencing of open chromatin regions. Single-cell DNA sequencing (scDNA-seq) offers a workaround but is less mainstream. Single-cell and single-nucleus analysis also provides a built-in germline control, improving variant detection power. It is crucial to accept null results in the analyses of scDNA-seq, as adjusting filtration parameters can exponentially increase called variants with no association with the studied condition.

Challenges of sampling the developing brain and spinal cord

In addition to dissecting cellular heterogeneity in snapshots of time points of tissue collections, coupling sequencing with the time axis allows the reconstruction of dynamic processes, such as cell

Table 2 | Examples of what types of questions can be addressed with scRNA-seq and snRNA-seq approaches

Questions that can be addressed using single-cell genomics	Data and technique	Key experimental considerations	Types of answers/results that can be obtained	Examples
How does disease status impact the milieu of cell states?	sc/snRNA-seq, scATAC-seq	Sampling power, cell population size, cell purification method	Disease-induced cell dysregulation	147–149
What regulatory activities can be inferred from sequenced cells? What are the primary factors controlling cell identity and activity?	sc/snRNA-seq, scATAC-seq, Perturb-seq	Data resolution (unimodal vs. multimodal data)	GRN, regulatory protein activity	48,49,57
What developmental switches lead to the observed phenotype?	sc/snRNA-seq, scATAC-seq	GRN/pseudotime methods, data resolution, cell purification	Differentiation checkpoint molecules	54
What is the cellular context of the sequenced tissue? What coordinated cell activities could be occurring?	sc/snRNA-seq, scATAC-seq	Cell isolation, cell annotation references	Cellular profile of tissue, cell–cell communication	15,142,147
What is the spectrum of mosaic somatic variants in a cell type or disease?	scDNA-seq, sc/snRNA-seq, single-cell targeted capture sequencing	Amplification protocol, depth of coverage	Somatic variants, cell phylogenies	117,150,151
What unique cell types are present in the sequenced sample?	sc/snRNA-seq, scATAC-seq, single-cell multiomics	Sequencing depth and resolution, cell population size	Rare cell identification, lineage trajectories	21,65,152

GRN, gene regulatory networks.

differentiation or state changes in response to external stimuli. This is particularly meaningful in developmental or disease contexts. The cellular heterogeneity may be attributed to distinct precursors, which can be traced upstream or downstream to build a lineage tree.

De novo discovery of rare precursor populations by scRNA-seq and snRNA-seq is possible if performed at sufficient throughput. However, throughput, depth and cost form the triple constraint to capturing or characterizing putative rare cells[1]. It may be necessary to go through the steps of initial unbiased sequencing, enrichment based on novel markers and/or deep sequencing of enriched cells to reveal the molecular properties of these cells. Modifying sampling to account for rare cell types is particularly pertinent given the heterogeneity of cell types in the nervous system, which varies substantially from the tens of subtypes of principal neurons in the cortex to the fewer subtypes of oligodendrocytes or ependymal cells. The label-free method mining rare cells by sequencing (MIRACL-seq) allows efficient profiling of rare cell types from a complex tissue by overloading droplets with single nuclei or cells[65]. Contaminating cells in the same droplets with the rare cells can be computationally removed. This technique is particularly useful for certain human samples when surface antibody or genetic labeling is not feasible. Other approaches to improve the power of low-abundance or captured cells in atlas experiments include sequencing large numbers of cells or nuclei (which can be cost prohibitive), pre-enriching for cell types of interest or removing cells not of interest (both of which require reporter lines or specific antibody labeling). These enriching or depleting approaches have been particularly useful in the study of microglia[66], astrocytes[26,67] and endothelial cells[68], which are routinely undersampled in scRNA-seq and snRNA-seq atlases.

Considerations in study design

Technical and experimental design of scRNA-seq and snRNA-seq experiments requires thoughtful consideration. Here, we cover some general points, but the parameters for each experiment should be based on the aim of the study (Table 2). Technical considerations include how the tissue is processed and the choice of technology. For example, it may be reasonable to exclude cells that are not of interest either experimentally via cell selection or computationally. The experimenters and data analysts need to be cognizant of how the samples were processed and also provide this information in published studies.

A critical aspect of tissue processing is dissociation artifacts, which can substantially impact the fidelity of single-cell data, potentially distorting the bona fide in vivo transcriptional states of cells. Previous studies have highlighted the prevalence and impact of dissociation-induced transcriptional changes, underscoring the importance of employing best-practice protocols that mitigate these effects[66,69–71].

Computational tools to ensure statistical robustness and power for scRNA-seq and snRNA-seq data's unique challenges and informing other aspects of the study design are discussed in ref. 1. Sequencing depth, crucial to identify rare cell types, should be considered based on the biological question, with multiplexing methods minimizing costs and reducing technical batch effects. Single-cell and single-nucleus experimental design requires appropriate biological replicates and controls to facilitate exploration of previously inaccessible questions. For rare samples (hard-to-acquire species, rare diseases, human samples), replicates with similar features and/or demographics may be challenging, but leveraging other published datasets (bulk RNA-seq, scRNA-seq or snRNA-seq) can account for gaps from replicate numbers that do not permit robust statistical approaches. Cost considerations often make scRNA-seq and snRNA-seq approaches more of a hypothesis-generating tool, followed by validation using orthogonal methods[2]. While these limitations are not ideal, they are essential considerations when choosing an scRNA-seq and/or snRNA-seq experiment, as universal guidelines cannot account for all experimental scenarios.

Common pitfalls in study design and data interpretation

While we have covered experimental design, proper data interpretation is equally important, especially for hypothesis-generating experiments. A common misinterpretation involves cell types of low abundance. Examining enough of such cells depends on sample size, which is particularly relevant for rare disease or species samples. Gene expression in rare cells needs replication (owing to gene dropout), and independent verification of their proportions should be carried out. Confirmatory experiments should use methods that keep tissue intact, such as multiplexed in situ hybridization or spatial transcriptomics with single-cell or single-nucleus resolution. Another pitfall is cell type annotation. While similar published datasets or brain atlases are useful for comparison, each dataset should be annotated within the experimental design context. Sample selection (for example, CNS region, age, disease status, genetic perturbation, species) could alter the presence of specific cell types and states. The investigator should

https://doi.org/10.1038/s41593-024-01827-9

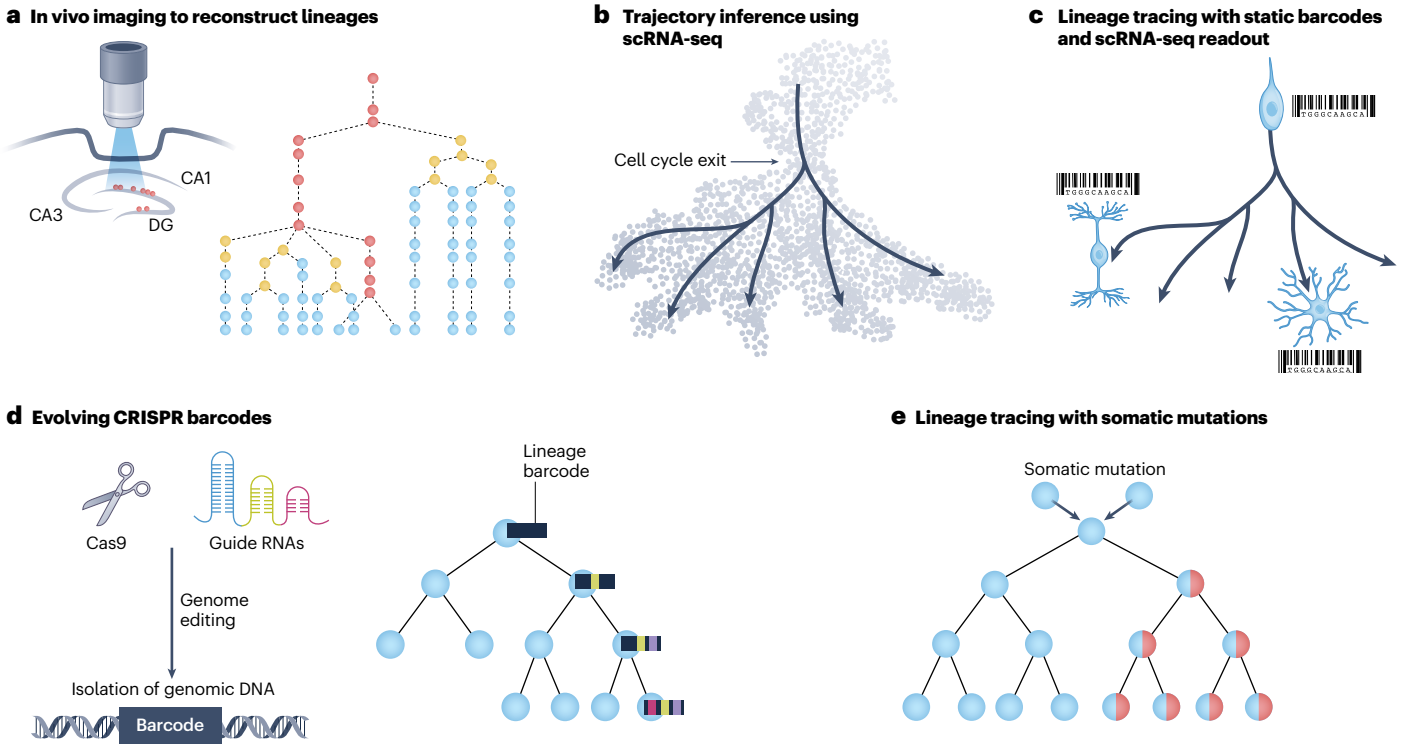

Fig. 2 | Methods for establishing temporal relationships between cells. **a**, Time-lapse microscopy has been the cornerstone technique for tracing cellular processes in vitro and in vivo. Illustrated is an example of lineage tracing performed in vivo in the mouse hippocampus. **b**, Single-cell and single-nucleus transcriptomic datasets reconstruct cellular relationships in gene expression space and can be used to predict developmental transitions. **c**, Molecular barcoding using complex libraries of static barcodes can be used to establish clonal lineage relationships between cells. **d**, CRISPR-based editing of the genome over time can introduce mutations that serve as evolvable barcodes. **e**, Retrospective lineage tracing using somatic mutations can be used in vivo without additional manipulation. CA, cornu ammonis; DG, dentate gyrus. Panel **a**, adapted with permission from ref. 154, AAAS; **b**, adapted from ref. 83, Springer Nature Limited.

assign and correct cell annotations using known biological information. Finally, CNS cells are unique in RNA transport and processing, yet many experiments use nuclei because of technical constraints. Some messenger RNA processing and transport properties may not be captured in nuclei. Profiling gene expression at synapses or dendrites at single-cell resolution can address concerns about missing RNA at synapses[72,73].

Data integration and batch correction add complexity to interpreting scRNA-seq and snRNA-seq atlases. Integrating datasets from different sources enhances sample size and statistical power, enriching biological insights[1,2]. Selecting the best method for data integration and batch correction requires an understanding of techniques such as canonical correlation analysis[74], mutual nearest neighbors or latent space projections, and optimizing them for the specific dataset[75,76], as discussed in ref. 1. Benchmarking studies on batch correction and other analyses are critical[77,78] and should include impartial computational metrics and thorough biological validation. With ongoing enhancements and integration of scRNA-seq and snRNA-seq tools and approaches, predicting all potential pitfalls is impossible. However, neuroscientists should consider their data's output with respect to new hypotheses: can the experimental results, especially identification of novel cell types and proportions, be supported independently? Could a novel cell type, state or proportion change be driven by technical artifacts or incorrect data annotation? We advocate for a common sense approach to interpreting results that deviate from current knowledge about cell types or states in the tissue of interest. Single-cell and single-nucleus genomics can provide important insights into the nervous system, but it is crucial to consider critical factors related to design, implementation and analysis as detailed throughout this piece.

Enhancing the accessibility of scRNA-seq and snRNA-seq atlases requires user-friendly platforms. The Allen Brain Cell Atlas and CELLxGENE offer web-based interfaces for exploring scRNA-seq and snRNA-seq data. The Allen Brain Cell Atlas also visualizes multimodal scRNA-seq and snRNA-seq data, aiding in understanding brain complexity. Proper data sharing is crucial (see ref. 1 for more on this). However, ultimately, we should aim to enhance coding literacy among biologists and enable them to easily mine the datasets; artificial intelligence-based coding support for code generation and explanation for noncomputational biologists should be improved and promoted.

Mapping cellular transitions in development and disease

Transcriptomic cell states emerge in early development, but can be altered under disease conditions. The ability to determine temporal relationships represents a critical feature of scRNA-seq and snRNA-seq experiments that is necessary to address fundamental questions in neuroscience (Fig. 2).

Lineage tracing is an important experimental approach that connects information about cells' history, state and fate[79,80]. In its simplest form, lineage tracing constructs a lineage tree providing a path to all the possible fates. However, without knowledge about the genetic program(s) that might steer the heterogeneity and dynamics of cell fate, lineage information has little benefit. Lineage tracing combines the power of interrogating cell states and lineage histories to allow reconstruction of the lineage tree, which is inclusive of clonal relationships between divergent fates and their critical regulators on the developmental time scale.

One of the aims of lineage tracing is to find molecular switch(s) that could modify cell state or cell division dynamics during development or even target diseased cells by modifying their transcriptional and/or epigenetic signatures. Below, we describe three approaches to map cell transitions and discuss two applications to understand disease states.

Among the three major approaches, computational inference has been the most extensively explored owing to its broad applicability to any scRNA-seq or snRNA-seq dataset. The main advantage here is the minimal data requirements, mainly spliced or unspliced gene counts, which, for example, have been used to reconstruct mouse hippocampal development from original RNA velocity work[81]. Prospective lineage tracing is the newest methodology gaining traction with recent tools such as CellTagging, LARRY, STICR, and so on[58,82–84], uncovering novel shared lineages of cortical excitatory and inhibitory neurons. Retrospective tracing is among the most powerful tools available and allows higher resolution than static prospective barcoding. However, there are limitations on implementation of retrospective lineage tracing, such as requirements of CRISPR system-based tracing (delivery and toxicity of double-stranded breaks) or high sequencing depth requirements of somatic tracing using single-cell whole-genome sequencing.

Computational inference of cell lineage

Developmental processes involve sequential and progressive changes in gene expression driven by gene regulatory programs[85,86]. Pseudotime reconstruction infers temporal relationships from scRNA-seq and snRNA-seq transcriptomes[86,87]. After dimensionality reduction of scRNA-seq and snRNA-seq datasets, cells along the trajectory show closer gene expression. These methods predict novel lineage relationships and identify molecular drivers of cell fate transitions. Pseudotemporal ordering predicts cell differentiation and state trajectories. Combined with multi-time point profiling, it highlights broad lineage changes over time. Popular methods use minimum spanning trees in Monocle 2 or 3, with lower-dimensional embedding (*t*-distributed stochastic neighbor embedding, uniform manifold approximation and projection (UMAP), principal-component analysis) to compute nearest neighbors or alternative approaches such as Palantir based on diffusion maps to orders cells along a pseudotime over a low-dimensional phenotypic manifold[88]. RNA velocity-based models such as velocyto, scVelo, dynamo and CellRank fit splicing kinetics on spliced and unspliced RNA, when possible, improving with time-resolved metabolic RNA labeling[81,89–91]. RNA velocity modeling is enhanced by single-cell long-read sequencing, which provides more complete splicing information[81,92]. However, newer trajectory inference algorithms integrate novel splicing models with multiview pseudotime and similarity analyses, marking a considerable advance in the field. The integration of pseudotemporal inference with time course sampling enables the prediction of dynamic relationships and the identification of cell state transitions[43]. Pseudotemporal inference, paired with time course sampling or further advanced methods such as Zman-seq for recording transcriptomic dynamics[28], predicts dynamic relationships and identifies cell state transitions. It can identify novel driver genes and gene regulatory network elements, especially with multiomics (scRNA-seq, snRNA-seq and scATAC-seq)[91,93,94].

Challenges in inference of cell trajectories

It is important to underscore that cell atlases depend heavily on assumptions made during computational analyses, such as the validity of distance in the reduced dimensional space. Incorrect assumptions of cell type and/or state similarity based on their reduced projections can lead to erroneous conclusions regarding trajectories of cellular change. The underlying dimensional reduction greatly influences conclusions, and misuse of batch correction techniques such as harmony and Seurat can inject substantial bias into the analysis[75,95]. Future computational methods may need to take a more holistic view of the sources of variation in multidimensional and multimodal space.

When transient states are unstable, such as during rapid differentiation, analyses can falsely suggest cell states, leading to incorrect lineage conclusions. False positive results can increase with dimensional reduction techniques, where parameter tweaks force cellular embeddings with different distance constraints, making populations appear close in lower-dimensional space, which leads to transitions being computed between unrelated populations.

Determining the directionality of states or identifying root and terminal states is a challenge with pseudotemporal methods. RNA velocity-based approaches struggle to distinguish mature cell states, especially with short-read sequencing data, leading to arbitrary transitions, and moreover do not fit nuclear data well. Supervised methods such as Monocle 2 and 3 and CellRank classify root and terminal states before pseudotemporal calculation, providing better control[91]. These unsupervised models often fail to differentiate velocity estimates from transcriptional noise. While constraining paths and directionality leads to accurate interpretations, prior knowledge is unfeasible in less-understood systems. The best practice is to validate novel trajectories using lineage tracing. Orthogonal validation is crucial for generating cell state atlases[2]. Low-abundance cells are often underpowered for interpreting terminal states, or transient differentiation states are incorrectly attributed as cell states. Minimal orthogonal validation through visualization (for example, in situ, high-resolution proteomics, spatial transcriptomics) can overcome shortcomings and should be required for all atlas development studies.

Prospective lineage tracing

Trajectory estimation as described above is unable to directly link the developmental relationship between individual cells and cannot be used to investigate whether the clonal progeny of a progenitor follow the same developmental pathway or branch off into multiple different pathways. By labeling an individual progenitor at an early time point and tracking the cell states of its clonal progeny at a later time point, lineage tracing is a viable approach to overcome this limitation.

Clonal labeling to trace developmental lineages can be achieved, for example, by using retroviral vectors[96–101] or transposase systems[83,102,103] that stably integrate genes encoding histochemical or fluorescent markers into the genome of a mitotic progenitor. This allows the marker gene to be passed on to the clonal progeny during cell division, enabling the labeling of clonally related neurons. Lineage tracing using marker genes has two major confounding factors: lumping errors (clustered cells that are not clonal) and splitting errors (dispersed cells that are clonal but are not recognized as such).

As an alternative to marker genes, diverse DNA barcodes can be used to trace clonal relationships. Static cellular barcodes[97,104–106] and evolving cellular barcodes[102,107–109] can tag clonal relationships more accurately and allow parallel tracking of thousands of lineages. Static lineage barcodes refer to unique DNA sequences that are stably integrated into a cell's genome, facilitating the identification of clonal progeny through sequencing. Evolving cellular barcodes enable the tracking of cell lineages across multiple generations by inducing mutations in the barcodes over time. This is achieved by using CRISPR–Cas9 genome editing to create repeated mutations within a DNA barcode that accumulate over many cell divisions.

By combining barcoding techniques with transcriptome analysis[83,84,108,110], it is possible to reveal developmental relationships between neuronal types and states. Furthermore, the integration of cellular lineage barcodes with trajectory inference methods[80,111,112] or genetic perturbation[113–115] holds great potential to improve predictions of fate probabilities and to identify key regulatory genes that drive developmental fate decisions.

Despite its potential, the accuracy of lineage tracing through scRNA-seq and/or snRNA-seq is still limited by the incomplete recovery of clones, which may result in the underestimation of true clonal diversity. This is because current in vivo methods combining barcode

lineage tracing with scRNA-seq and/or snRNA-seq suffer from partial lineage recovery for several reasons, in particular, cell loss during tissue dissociation, sorting and droplet loading. These steps can lead to biased sampling of cells, resulting in an underestimation of true clonal diversity. Recent developments in lineage-tracing methods that recover barcodes in intact tissue hold promise for more complete lineage recovery[110].

Retrospective inference of cell lineage

Beyond the information embedded in RNA profiles, every individual has a lineage map inscribed in somatic mutations that occur with each cell division in the nuclear genome[116–120]. Early cell divisions accumulate about two to three single-nucleotide variants per division, with higher rates during neurogenesis[99,118,121]. Microsatellite mutation rates are higher but undetermined, while copy number variation and transposon insertion occur occasionally, mapping when any two cells diverge from a common ancestor[122–124]. While most spontaneous somatic mutations are functionally silent, some alter gene function and create risk for neuropsychiatric diseases, notably epilepsy[125–128] and ASD[121,129–132].

Proliferation-related mutations, while abundant, occur throughout the genome, typically in non-transcribed sequences, and so are not usually captured in routine scRNA-seq, snRNA-seq or scATAC-seq, because (1) these methods cover only a small fraction of the genome, and (2) they cover the genome at low coverage, making it difficult to call variations in genome sequence[119]. The highly dispersed nature of somatic single-nucleotide variants throughout the genome presently makes it cost prohibitive to measure them at scale, because they show only modest regional enrichment. Therefore, a major technical challenge that prevents the widespread use of somatic mutations as lineage markers is the lack of high-throughput, affordable methods to assay DNA mutations and markers of cell identity, although multiple methods are under development that are likely to improve this situation.

Mitochondrial mutations have also been explored as lineage marks in humans[133–135]. They have the advantage that the mitochondrial genome mutates 100–1,000 times faster than the nuclear genome, but it is also 100,000 times smaller, so that mutations per cell division in the mitochondrial genome are a fraction of the mutations in the nuclear genome. On the other hand, the mitochondrial genome is virtually all open chromatin and so is captured reliably by ATAC-seq almost in its entirety, providing a proof-of-principle system in which some marks of lineage and good marks of cell type can be obtained in a single experiment[133,134]. Recent advances in methods for isolating cells and sequencing mitochondrial DNA promise to improve this approach[133,135].

Cell lineage, defined as the sequence of cell divisions that generates a structure, is likely to be different from the apparent cell lineage inferred by assessing seemingly continuous states of single-cell or single-nucleus RNA gene expression. For example, studies of cell lineage in humans with barcoded viral libraries show both convergence (formation of a single cell type from multiple lineage origins) and divergence (formation of distinct cell types from a common progenitor)[84].

One major limitation for such somatic mutation-based approaches, and an area of technological development, is that lineage and transcriptomic information using scRNA-seq and/or snRNA-seq lacks spatial resolution for in vivo systems[136]. This limitation makes it challenging to reconstruct dynamic processes involving cell migration. In addition, sequential profiling of the same cells over time is another appealing possibility that enables direct comparisons of cell states along a trajectory[1,137] (Box 1).

Tracking cells during disease progression: brain tumors

Studies of brain tumors represent a unique example of single-cell lineage tracing. scRNA-seq and snRNA-seq studies of pediatric and adult brain tumors showed that tumors contain progenitor cell states that reflect the ones found during normal development. For example, scRNA-seq profiling of histone H3 K27M gliomas and adult glioblastoma showed that these tumors are composed of a majority of oligodendrocyte progenitor-like cells that retain partial differentiation toward more differentiated neurodevelopmental cell types. The latest scRNA-seq and snRNA-seq atlases of human brain development, together with advanced computational data integration methods, allow precise comparison of cell states and cell fate across healthy and malignant tissues. This has enabled the discovery of neurodevelopmental programs used by neural stem cells of the normally developing brain that are recruited by multiple types of tumor, including programs that are found in the human but not the mouse developing brain[138,139]. These findings serve as a foundation for re-evaluating the accuracy of experimental models used to study disease states.

Unlike cells in normal development that follow distinct stages of differentiation, the unpredictability of aberrant cell fate transitions

BOX 1

Tracking the dynamics of cell states: CNS myeloid cells as a model

Transcriptional cell states emerge early in development and transition during differentiation, maturation, aging and disease. CNS myeloid cells, including microglia and CNS-associated macrophages, exemplify this paradigm, although their ontogeny and clonal relationships remain unclear[155]. Aging partially explains differences between white and gray matter microglia[37], but whether a single microglial cell can adopt all context-dependent states is unknown. Disease increases microglial turnover[156,157], possibly involving circulating myeloid cells with distinct molecular signatures. Static measurements, however, fail to address these dynamics fully.

For targeted therapy, understanding a cell's past states during bifurcation, before it drives or responds to disease, is crucial. Lineage-tracing technologies have demonstrated their potential in tracking cancer clones through evolution, adaptation and metastasis[158–160]. In CNS diseases, pseudotime trajectories have revealed transcriptional states and key regulatory genes in microglia[155], but experimental validation of intermediate states is lacking.

Current lineage-tracing technologies face technical issues such as off-target effects, barcode expression silencing and stable clone screening. Random barcode insertions offer high diversity and resolution but have higher off-target effects. By contrast, site-directed insertions avoid off-target effects but have lower resolution owing to limited barcodes[80].

Data analysis methods for lineage tracing as applied to scRNA-seq and snRNA-seq are not universal, and therefore harmonizing results across laboratories could be challenging. Spatial contexts are also crucial, and, while spatially resolved lineage-tracing methods exist, they are not widely available. Some cell responses, like those of microglia, are tightly regulated, and evolving barcodes may not capture small microglial populations corresponding to a single state.

Combination barcodes, where a static barcode marks clonal reconstruction and editable barcodes are used for phylogenetic reconstruction[159], could be implemented. CellTag-multi, which is compatible with multiple single-cell modalities, allows recovery of barcodes from single-cell chromatin accessibility assays[143].

BOX 2

Tracking cellular fate in neuroscience research

- Trajectory estimation is a computational technique used in single-cell transcriptomics and epigenomics to infer the developmental pathways or trajectories of cells on the basis of their gene expression and epigenetic profiles.
- Lineage tracing is a technique used to trace the direct clonal progeny of individual progenitors. This technique involves labeling progenitors with heritable markers or tags to understand their fate and behavior during development.
- Fate mapping is a technique used to track a group of cells that share a particular characteristic, such as the expression of a marker gene or a transcription factor during development.

within brain tumors is a key challenge. In addition, scarcity of longitudinal brain tumor samples prevents tracking of cell states over time. Moreover, the timing of cell fate switching is unknown, both under physiologic conditions and under therapeutic pressure (that is, it is unclear whether it takes days, minutes or some other time period). Most of the experiments proposed above are difficult or not feasible in vivo, which is problematic as the tumor microenvironment is likely to be a major contributor to cancer cell fate transitions.

Pseudotime analyses, lineage tracing and multiome analyses can all be used to assess intrinsic and extrinsic factors that regulate cell states and fate transitions in brain tumors during evolution as well as during drug or radiation treatment (Box 2). Using multiomic scRNA-seq and scATAC-seq data to reconstruct gene regulatory networks, as discussed above, could also be applied to identify networks governing cell fate specifications in cancer cells[140], and states that emerge as a consequence of specific genetic mutations can be applied. This can be done directly on patient tissues or in experimental in vitro or in vivo models. Validating predictions from such analyses is by no means straightforward and requires more direct ways of probing tumor cell lineages. For example, lineage tracing of cancer cells in the presence of a perturbation (genetic or pharmacological) or when co-cultured with normal cells of the brain tumor microenvironment can be done in experimental models using methods described above. Alternatively, the plasticity and trajectory of specific cell states can be assessed by sorting cancer cells directly from patients' tumors on the basis of the expression of state-specific cell surface markers and analyzing their composition and trajectory at different time points by subsequent scRNA-seq, snRNA-seq and/or lineage tracing. Tumor cells have increased burden of copy number aberrations and other DNA mutations, and this feature has been used as a basis to model tumor cell evolution[141].

There are two important conceptual challenges in understanding brain tumor growth. First, most cancer cells are stalled in their cell fate specification, which is determined by both intrinsic (underlying oncogenic mutations) and microenvironmental (including normal cells within the tumor microenvironment, such as neurons, astrocytes and microglia) factors. Second, cancer cells retain some degree of developmental plasticity and can differentiate toward more mature cell states. This suggests that cell fate mechanisms are not only dysregulated in cancer but also fuel cancer growth and resistance, allowing cancer cells to switch from treatment-sensitive to treatment-resistant states. Therefore, an understanding of the mechanisms that govern cell fate transitions is necessary to design therapies that effectively target each state and/or target plasticity.

Conclusion and future directions

Here, we summarize the implementation of scRNA-seq and snRNA-seq in neuroscience for constructing atlases, generating hypotheses and understanding cell fate and lineage progression in normal development and disease. As evidenced by our detailed descriptions of pitfalls and challenges, there are still many technical hurdles to overcome. The specific examples that we choose to discuss do not capture the full spectrum of possible applications or challenges but are merely intended to illustrate common use cases. Furthermore, while we strive to present unified concepts in these domains, there is no existing consensus handbook regarding the design, execution and/or interpretation of scRNA-seq and snRNA-seq experiments for CNS samples. Nonetheless, this lack of constraint presents an opportunity for further development, in terms of both tool building and application. For example, harnessing cell fate tools and applying them to postmitotic cells such as neurons could unveil signatures of activity-dependent gene expression. Increasing the scale at which electrophysiological and scRNA-seq and snRNA-seq transcriptomic signatures are measured (for example, Patch-seq) could also enhance our understanding (in some tissues) of how individual brain cells respond to activity or network perturbations. Genetic tools for scalable recording or perturbation of predicted regulatory network hub transcriptional programs in the context of cell lineage analysis would enable systematic testing of predictions from cell atlas studies, and many such technologies are beginning to emerge. Finally, expanding the number of species and developmental time points included in building atlases will also yield more hypotheses about brain development, evolution and disease.

References

1. Bonev, B. et al. Opportunities and challenges of single-cell and spatially resolved genomics methods for neuroscience discovery. *Nat. Neurosci.* https://doi.org/10.1038/s41593-024-01806-0 (2024).
2. Colonna, M. et al. Implementation and validation of single-cell genomics experiments in neuroscience. *Nat. Neurosci.* https://doi.org/10.1038/s41593-024-01814-0 (2024).
3. Bakken, T. E. et al. Comparative cellular analysis of motor cortex in human, marmoset and mouse. *Nature* **598**, 111–119 (2021).
4. Regev, A. et al. The Human Cell Atlas. *eLife* **6**, e27041 (2017).
5. Ecker, J. R. et al. The BRAIN Initiative Cell Census Consortium: lessons learned toward generating a comprehensive brain cell atlas. *Neuron* **96**, 542–557 (2017).
6. Zeisel, A. et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–1142 (2015).
7. Yao, Z. et al. A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell* **184**, 3222–3241 (2021).
8. Pollen, A. A. et al. Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat. Biotechnol.* **32**, 1053–1058 (2014).
9. Langlieb, J. et al. The molecular cytoarchitecture of the adult mouse brain. *Nature* **624**, 333–342 (2023).
10. Tasic, B. et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19**, 335–346 (2016).
11. Tabula Muris, C. et al. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* **562**, 367–372 (2018).
12. Domcke, S. & Shendure, J. A reference cell tree will serve science better than a reference cell atlas. *Cell* **186**, 1103–1114 (2023).
13. Gouwens, N. W. et al. Integrated morphoelectric and transcriptomic classification of cortical GABAergic cells. *Cell* **183**, 935–953 (2020).
14. Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).

15. Keren-Shaul, H. et al. A unique microglia type associated with restricting development of Alzheimer's disease. *Cell* **169**, 1276–1290 (2017).
16. Gazestani, V. et al. Early Alzheimer's disease pathology in human cortex involves transient cell states. *Cell* **186**, 4438–4453 (2023).
17. Kaya, T. et al. CD8$^+$ T cells induce interferon-responsive oligodendrocytes and microglia in white matter aging. *Nat. Neurosci.* **25**, 1446–1457 (2022).
18. Chen, X. et al. Microglia-mediated T cell infiltration drives neurodegeneration in tauopathy. *Nature* **615**, 668–677 (2023).
19. Nagy, C. et al. Single-nucleus transcriptomics of the prefrontal cortex in major depressive disorder implicates oligodendrocyte precursor cells and excitatory neurons. *Nat. Neurosci.* **23**, 771–781 (2020).
20. Maniatis, S. et al. Spatiotemporal dynamics of molecular pathology in amyotrophic lateral sclerosis. *Science* **364**, 89–93 (2019).
21. Wang, Z., Gerstein, M. & Snyder, M. RNA-seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63 (2009).
22. Squair, J. W. et al. Confronting false discoveries in single-cell differential expression. *Nat. Commun.* **12**, 5692 (2021).
23. Stark, R., Grzelak, M. & Hadfield, J. RNA sequencing: the teenage years. *Nat. Rev. Genet.* **20**, 631–656 (2019).
24. Hodge, R. D. et al. Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
25. MacLean, A. L., Hong, T. & Nie, Q. Exploring intermediate cell states through the lens of single cells. *Curr. Opin. Syst. Biol.* **9**, 32–41 (2018).
26. Sadick, J. S. et al. Astrocytes and oligodendrocytes undergo subtype-specific transcriptional changes in Alzheimer's disease. *Neuron* **110**, 1788–1805 (2022).
27. Cain, A. et al. Multicellular communities are perturbed in the aging human brain and Alzheimer's disease. *Nat. Neurosci.* **26**, 1267–1280 (2023).
28. Kirschenbaum, D. et al. Time-resolved single-cell transcriptomics defines immune trajectories in glioblastoma. *Cell* **187**, 149–165 (2024).
29. Habib, N. et al. Disease-associated astrocytes in Alzheimer's disease and aging. *Nat. Neurosci.* **23**, 701–706 (2020).
30. Sun, W. et al. Spatial transcriptomics reveal neuron–astrocyte synergy in long-term memory. *Nature* **627**, 374–381 (2024).
31. Olah, M. et al. Single cell RNA sequencing of human microglia uncovers a subset associated with Alzheimer's disease. *Nat. Commun.* **11**, 6129 (2020).
32. Kamath, T. et al. Single-cell genomic profiling of human dopamine neurons identifies a population that selectively degenerates in Parkinson's disease. *Nat. Neurosci.* **25**, 588–595 (2022).
33. Malaiya, S. et al. Single-nucleus RNA-seq reveals dysregulation of striatal cell identity due to Huntington's disease mutations. *J. Neurosci.* **41**, 5534–5552 (2021).
34. Stanley, G., Gokce, O., Malenka, R. C., Sudhof, T. C. & Quake, S. R. Continuous and discrete neuron types of the adult murine striatum. *Neuron* **105**, 688–699 (2020).
35. Gokce, O. et al. Cellular taxonomy of the mouse striatum as revealed by single-cell RNA-seq. *Cell Rep.* **16**, 1126–1137 (2016).
36. Hashemiaghdam, A. & Mroczek, M. Microglia heterogeneity and neurodegeneration: the emerging paradigm of the role of immunity in Alzheimer's disease. *J. Neuroimmunol.* **341**, 577185 (2020).
37. Safaiyan, S. et al. White matter aging drives microglial diversity. *Neuron* **109**, 1100–1117 (2021).
38. Siletti, K. et al. Transcriptomic diversity of cell types across the adult human brain. *Science* **382**, eadd7046 (2023).
39. Krienen, F. M. et al. A marmoset brain cell census reveals regional specialization of cellular identities. *Sci. Adv.* **9**, eadk3986 (2023).
40. Nowakowski, T. J. et al. Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* **358**, 1318–1323 (2017).
41. Tasic, B. et al. Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018).
42. Bhaduri, A. et al. An atlas of cortical arealization identifies dynamic molecular signatures. *Nature* **598**, 200–204 (2021).
43. Saelens, W., Cannoodt, R., Todorov, H. & Saeys, Y. A comparison of single-cell trajectory inference methods. *Nat. Biotechnol.* **37**, 547–554 (2019).
44. Quesnel-Vallieres, M., Weatheritt, R. J., Cordes, S. P. & Blencowe, B. J. Autism spectrum disorder: insights into convergent mechanisms from transcriptomics. *Nat. Rev. Genet.* **20**, 51–63 (2019).
45. Yang, Q. et al. Consistent gene signature of schizophrenia identified by a novel feature selection strategy from comprehensive sets of transcriptomic data. *Brief. Bioinform.* **21**, 1058–1068 (2020).
46. Lobentanzer, S., Hanin, G., Klein, J. & Soreq, H. Integrative transcriptomics reveals sexually dimorphic control of the cholinergic/neurokine interface in schizophrenia and bipolar disorder. *Cell Rep.* **29**, 764–777 (2019).
47. Ardesch, D. J., Libedinsky, I., Scholtens, L. H., Wei, Y. & van den Heuvel, M. P. Convergence of brain transcriptomic and neuroimaging patterns in schizophrenia, bipolar disorder, autism spectrum disorder, and major depressive disorder. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **8**, 630–639 (2023).
48. Ziffra, R. S. et al. Single-cell epigenomics reveals mechanisms of human cortical development. *Nature* **598**, 205–213 (2021).
49. Trevino, A. E. et al. Chromatin and gene-regulatory dynamics of the developing human cerebral cortex at single-cell resolution. *Cell* **184**, 5053–5069 (2021).
50. Jin, X. et al. In vivo Perturb-seq reveals neuronal and glial abnormalities associated with autism risk genes. *Science* **370**, eaaz6063 (2020).
51. Fleck, J. S. et al. Inferring and perturbing cell fate regulomes in human brain organoids. *Nature* **621**, 365–372 (2023).
52. Dixit, A. et al. Perturb-seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* **167**, 1853–1866 (2016).
53. Wheeler, M. A. et al. Droplet-based forward genetic screening of astrocyte–microglia cross-talk. *Science* **379**, 1023–1030 (2023).
54. Soldatov, R. et al. Spatiotemporal structure of cell fate decisions in murine neural crest. *Science* **364**, eaas9536 (2019).
55. Kastriti, M. E. et al. Schwann cell precursors represent a neural crest-like state with biased multipotency. *EMBO J.* **41**, e108780 (2022).
56. Faure, L., Soldatov, R., Kharchenko, P. V. & Adameyko, I. scFates: a scalable python package for advanced pseudotime and bifurcation analysis from single-cell data. *Bioinformatics* **39**, btac746 (2023).
57. Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
58. Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F. D. & Klein, A. M. Lineage tracing on transcriptional landscapes links state to fate during differentiation. *Science* **367**, eaaw3381 (2020).
59. Erickson, A. G., Kameneva, P. & Adameyko, I. The transcriptional portraits of the neural crest at the individual cell level. *Semin. Cell Dev. Biol.* **138**, 68–80 (2023).
60. Jin, S. et al. Inference and analysis of cell–cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).
61. Jin, Z. et al. Single-cell gene fusion detection by scFusion. *Nat. Commun.* **13**, 1084 (2022).

62. Patel, A. P. et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396–1401 (2014).
63. Gao, R. et al. Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. *Nat. Biotechnol.* **39**, 599–608 (2021).
64. Petti, A. A. et al. A general approach for detecting expressed mutations in AML cells using single cell RNA-sequencing. *Nat. Commun.* **10**, 3660 (2019).
65. Drokhlyansky, E. et al. The human and mouse enteric nervous system at single-cell resolution. *Cell* **182**, 1606–1622 (2020).
66. Marsh, S. E. et al. Dissection of artifactual and confounding glial signatures by single-cell sequencing of mouse and human brain. *Nat. Neurosci.* **25**, 306–316 (2022).
67. Hasel, P., Rose, I. V. L., Sadick, J. S., Kim, R. D. & Liddelow, S. A. Neuroinflammatory astrocyte subtypes in the mouse brain. *Nat. Neurosci.* **24**, 1475–1487 (2021).
68. Yang, A. C. et al. A human brain vascular atlas reveals diverse mediators of Alzheimer's risk. *Nature* **603**, 885–892 (2022).
69. Wu, Y. E., Pan, L., Zuo, Y., Li, X. & Hong, W. Detecting activated cell populations using single-cell RNA-seq. *Neuron* **96**, 313–329 (2017).
70. Van Hove, H. et al. A single-cell atlas of mouse brain macrophages reveals unique transcriptional identities shaped by ontogeny and tissue environment. *Nat. Neurosci.* **22**, 1021–1035 (2019).
71. Liu, L. et al. Dissociation of microdissected mouse brain tissue for artifact free single-cell RNA sequencing. *STAR Protoc.* **2**, 100590 (2021).
72. Niu, M. et al. Droplet-based transcriptome profiling of individual synapses. *Nat. Biotechnol.* **41**, 1332–1344 (2023).
73. Perez, J. D. et al. Subcellular sequencing of single neurons reveals the dendritic transcriptome of GABAergic interneurons. *eLife* **10**, e63092 (2021).
74. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 (2019).
75. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
76. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nat. Methods* **15**, 1053–1058 (2018).
77. Luecken, M. D. et al. Benchmarking atlas-level data integration in single-cell genomics. *Nat. Methods* **19**, 41–50 (2022).
78. Tran, H. T. N. et al. A benchmark of batch-effect correction methods for single-cell RNA sequencing data. *Genome Biol.* **21**, 12 (2020).
79. VanHorn, S. & Morris, S. A. Next-generation lineage tracing and fate mapping to interrogate development. *Dev. Cell* **56**, 7–21 (2021).
80. Wagner, D. E. & Klein, A. M. Lineage tracing meets single-cell omics: opportunities and challenges. *Nat. Rev. Genet.* **21**, 410–427 (2020).
81. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
82. Biddy, B. A. et al. Single-cell mapping of lineage and identity in direct reprogramming. *Nature* **564**, 219–224 (2018).
83. Bandler, R. C. et al. Single-cell delineation of lineage and genetic identity in the mouse brain. *Nature* **601**, 404–409 (2022).
84. Delgado, R. N. et al. Individual human cortical progenitors can produce excitatory and inhibitory neurons. *Nature* **601**, 397–403 (2022).
85. Martik, M. L., Lyons, D. C. & McClay, D. R. Developmental gene regulatory networks in sea urchins and what we can learn from them. *F1000Res.* **5**, F1000 (2016).
86. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
87. Durruthy-Durruthy, R. & Heller, S. Applications for single cell trajectory analysis in inner ear development and regeneration. *Cell Tissue Res.* **361**, 49–57 (2015).
88. Setty, M. et al. Characterization of cell fate probabilities in single-cell data with Palantir. *Nat. Biotechnol.* **37**, 451–460 (2019).
89. Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 (2020).
90. Qiu, X. et al. Mapping transcriptomic vector fields of single cells. *Cell* **185**, 690–711 (2022).
91. Lange, M. et al. CellRank for directed single-cell fate mapping. *Nat. Methods* **19**, 159–170 (2022).
92. Bergen, V., Soldatov, R. A., Kharchenko, P. V. & Theis, F. J. RNA velocity—current challenges and future perspectives. *Mol. Syst. Biol.* **17**, e10282 (2021).
93. Lynch, A. W. et al. MIRA: joint regulatory modeling of multimodal expression and chromatin accessibility in single cells. *Nat. Methods* **19**, 1097–1108 (2022).
94. Bravo Gonzalez-Blas, C. et al. SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. *Nat. Methods* **20**, 1355–1367 (2023).
95. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
96. Price, J., Turner, D. & Cepko, C. Lineage analysis in the vertebrate nervous system by retrovirus-mediated gene transfer. *Proc. Natl Acad. Sci. USA* **84**, 156–160 (1987).
97. Walsh, C. & Cepko, C. L. Widespread dispersion of neuronal clones across functional regions of the cerebral cortex. *Science* **255**, 434–440 (1992).
98. Noctor, S. C., Flint, A. C., Weissman, T. A., Dammerman, R. S. & Kriegstein, A. R. Neurons derived from radial glial cells establish radial units in neocortex. *Nature* **409**, 714–720 (2001).
99. Yu, Y. C. et al. Preferential electrical coupling regulates neocortical lineage-dependent microcircuit assembly. *Nature* **486**, 113–117 (2012).
100. Li, Y. et al. Clonally related visual cortical neurons show similar stimulus feature selectivity. *Nature* **486**, 118–121 (2012).
101. Ciceri, G. et al. Lineage-specific laminar organization of cortical GABAergic interneurons. *Nat. Neurosci.* **16**, 1199–1210 (2013).
102. Kalhor, R. et al. Developmental barcoding of whole mouse via homing CRISPR. *Science* **361**, eaat9804 (2018).
103. Wagner, D. E. et al. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **360**, 981–987 (2018).
104. Golden, J. A., Fields-Berry, S. C. & Cepko, C. L. Construction and characterization of a highly complex retroviral library for lineage analysis. *Proc. Natl Acad. Sci. USA* **92**, 5704–5708 (1995).
105. Harwell, C. C. et al. Wide dispersion and diversity of clonally related inhibitory interneurons. *Neuron* **87**, 999–1007 (2015).
106. Mayer, C. et al. Clonally related forebrain interneurons disperse broadly across both functional areas and structural boundaries. *Neuron* **87**, 989–998 (2015).
107. Alemany, A., Florescu, M., Baron, C. S., Peterson-Maduro, J. & van Oudenaarden, A. Whole-organism clone tracing using single-cell sequencing. *Nature* **556**, 108–112 (2018).
108. Leeper, K. et al. Lineage barcoding in mice with homing CRISPR. *Nat. Protoc.* **16**, 2088–2108 (2021).
109. McKenna, A. et al. Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science* **353**, aaf7907 (2016).

110. Ratz, M. et al. Clonal relations in the mouse brain revealed by single-cell and spatial transcriptomics. *Nat. Neurosci.* **25**, 285–294 (2022).
111. Forrow, A. & Schiebinger, G. LineageOT is a unified framework for lineage tracing and trajectory inference. *Nat. Commun.* **12**, 4940 (2021).
112. Lange, M. et al. Mapping lineage-traced cells across time points with moslin. *Genome Biol.* **25**, 277 (2024).
113. Esk, C. et al. A human tissue screen identifies a regulator of ER secretion as a brain-size determinant. *Science* **370**, 935–941 (2020).
114. He, Z. et al. Lineage recording in human cerebral organoids. *Nat. Methods* **19**, 90–99 (2022).
115. Dvoretskova, E. et al. Spatial enhancer activation influences inhibitory neuron identity during mouse embryonic development. *Nat. Neurosci.* **27**, 862–872 (2024).
116. Fasching, L. et al. Early developmental asymmetries in cell lineage trees in living individuals. *Science* **371**, 1245–1248 (2021).
117. Lodato, M. A. et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science* **350**, 94–98 (2015).
118. Bae, T. et al. Different mutational rates and mechanisms in human cells at pregastrulation and neurogenesis. *Science* **359**, 550–555 (2018).
119. Bizzotto, S. et al. Landmarks of human embryonic development inscribed in somatic mutations. *Science* **371**, 1249–1253 (2021).
120. Coorens, T. H. H. et al. Extensive phylogenies of human development inferred from somatic mutations. *Nature* **597**, 387–392 (2021).
121. Rodin, R. E. et al. The landscape of somatic mutation in cerebral cortex of autistic and neurotypical individuals revealed by ultra-deep whole-genome sequencing. *Nat. Neurosci.* **24**, 176–185 (2021).
122. Erwin, J. A. et al. L1-associated genomic regions are deleted in somatic cells of the healthy human brain. *Nat. Neurosci.* **19**, 1583–1591 (2016).
123. Evrony, G. D. et al. Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* **151**, 483–496 (2012).
124. Evrony, G. D. et al. Cell lineage analysis in human brain using endogenous retroelements. *Neuron* **85**, 49–59 (2015).
125. Baldassari, S. et al. Dissecting the genetic basis of focal cortical dysplasia: a large cohort study. *Acta Neuropathol.* **138**, 885–900 (2019).
126. Chung, C. et al. Comprehensive multi-omic profiling of somatic mutations in malformations of cortical development. *Nat. Genet.* **55**, 209–220 (2023).
127. Kim, J. H. et al. Ultra-low level somatic mutations and structural variations in focal cortical dysplasia type II. *Ann. Neurol.* **93**, 1082–1093 (2023).
128. Khoshkhoo, S. et al. Contribution of somatic Ras/Raf/mitogen-activated protein kinase variants in the hippocampus in drug-resistant mesial temporal lobe epilepsy. *JAMA Neurol.* **80**, 578–587 (2023).
129. D'Gama, A. M. et al. Targeted DNA sequencing from autism spectrum disorder brains implicates multiple genetic mechanisms. *Neuron* **88**, 910–917 (2015).
130. Freed, D. & Pevsner, J. The contribution of mosaic variants to autism spectrum disorder. *PLoS Genet.* **12**, e1006245 (2016).
131. Krupp, D. R. et al. Exonic mosaic mutations contribute risk for autism spectrum disorder. *Am. J. Hum. Genet.* **101**, 369–390 (2017).
132. Lim, E. T. et al. Rates, distribution and implications of postzygotic mosaic mutations in autism spectrum disorder. *Nat. Neurosci.* **20**, 1217–1224 (2017).
133. Lareau, C. A. et al. Mitochondrial single-cell ATAC-seq for high-throughput multi-omic detection of mitochondrial genotypes and chromatin accessibility. *Nat. Protoc.* **18**, 1416–1440 (2023).
134. Ludwig, L. S. et al. Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. *Cell* **176**, 1325–1339 (2019).
135. Miller, T. E. et al. Mitochondrial variant enrichment from high-throughput single-cell RNA sequencing resolves clonal populations. *Nat. Biotechnol.* **40**, 1030–1034 (2022).
136. Frieda, K. L. et al. Synthetic recording and in situ readout of lineage information in single cells. *Nature* **541**, 107–111 (2017).
137. Chen, W. et al. Live-seq enables temporal transcriptomic recording of single cells. *Nature* **608**, 733–740 (2022).
138. Smith, K. S. et al. Unified rhombic lip origins of group 3 and group 4 medulloblastoma. *Nature* **609**, 1012–1020 (2022).
139. Bhaduri, A. et al. Outer radial glia-like cancer stem cells contribute to heterogeneity of glioblastoma. *Cell Stem Cell* **26**, 48–63 (2020).
140. Kanton, S. et al. Organoid single-cell genomic atlas uncovers human-specific features of brain development. *Nature* **574**, 418–422 (2019).
141. Davis, A., Gao, R. & Navin, N. Tumor evolution: linear, branching, neutral or punctuated? *Biochim. Biophys. Acta Rev. Cancer* **1867**, 151–161 (2017).
142. Mathys, H. et al. Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332–337 (2019).
143. Jindal, K. et al. Single-cell lineage capture across genomic modalities with CellTag-multi reveals fate-specific gene regulatory changes. *Nat. Biotechnol.* **42**, 946–959 (2024).
144. Fullgrabe, A. et al. Guidelines for reporting single-cell RNA-seq experiments. *Nat. Biotechnol.* **38**, 1384–1386 (2020).
145. Qiu, C. et al. A single-cell time-lapse of mouse prenatal development from gastrula to birth. *Nature* **626**, 1084–1093 (2024).
146. Luecken, M. D. & Theis, F. J. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol. Syst. Biol.* **15**, e8746 (2019).
147. Arneson, D. et al. Single cell molecular alterations reveal target cells and pathways of concussive brain injury. *Nat. Commun.* **9**, 3894 (2018).
148. Grubman, A. et al. A single-cell atlas of entorhinal cortex from individuals with Alzheimer's disease reveals cell-type-specific gene expression regulation. *Nat. Neurosci.* **22**, 2087–2097 (2019).
149. Lau, S. F., Cao, H., Fu, A. K. Y. & Ip, N. Y. Single-nucleus transcriptome analysis reveals dysregulation of angiogenic endothelial cells and neuroprotective glia in Alzheimer's disease. *Proc. Natl Acad. Sci. USA* **117**, 25800–25809 (2020).
150. Miller, M. B. et al. Somatic genomic changes in single Alzheimer's disease neurons. *Nature* **604**, 714–722 (2022).
151. Townsend, S. E. et al. Single-nuclei transcriptomics enable detection of somatic variants in patient brain tissue. *Sci. Rep.* **13**, 527 (2023).
152. Zhou, Y. et al. Molecular landscapes of human hippocampal immature neurons across lifespan. *Nature* **607**, 527–533 (2022).
153. BRAIN Initiative Cell Census Network (BICCN), A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature* **598**, 86–102 (2021).
154. Pilz, G. A. et al. Live imaging of neurogenesis in the adult mouse hippocampus. *Science* **359**, 658–662 (2018).
155. Paolicelli, R. C. et al. Microglia states and nomenclature: a field at its crossroads. *Neuron* **110**, 3458–3483 (2022).
156. Chhatbar, C. et al. Type I interferon receptor signaling of neurons and astrocytes regulates microglia activation during viral encephalitis. *Cell Rep.* **25**, 118–129 (2018).

157. Masuda, T. et al. Spatial and temporal heterogeneity of mouse and human microglia at single-cell resolution. *Nature* **566**, 388–392 (2019).
158. Quinn, J. J. et al. Single-cell lineages reveal the rates, routes, and drivers of metastasis in cancer xenografts. *Science* **371**, eabc1944 (2021).
159. Simeonov, K. P. et al. Single-cell lineage tracing of metastatic cancer reveals selection of hybrid EMT states. *Cancer Cell* **39**, 1150–1162 (2021).
160. Aivazidis, A. et al. Model-based inference of RNA velocity modules improves cell fate prediction. Preprint at *bioRxiv* https://doi.org/10.1101/2023.08.03.551650 (2023).

Acknowledgements

We thank R. Awatramani, H. S. Bateup, B. Bonev, C. R. Cadwell, E. Caglayan, G. Castelo-Branco, F. Chen, J. L. Chen, S. Codeluppi, R. Corces, J. Fan, J. Gillis, G. Green, M. Heiman, K. Harris, F. Inoue, M. Kampmann, M. Kellis, F. Krienen, M. Monje, M. R. O'Dea, R. Patani, A. Pollen, A. Levine, M. Lotfollahi, C. Luo, E. Macosko, S. Marsh, K. R. Maynard, M. Nitzan, F. J. Quintana, V. Ramani, R. Satija, M. Scavuzzo, L. Schirmer, M. Schmitz, Y. Shen, S. Sloan, N. Sun, P. Tesar, F. Theis, J. Tollkuhn, M. A. Tosches, M. E. Urbanek, X. Wang, J. D. Welch, J. Werner, H. Zeng, B. Tasic and T. Kaya for insightful feedback on this work and J. Kuhl for the illustrations in Fig. 2. G.K. is a Jon Heighten Scholar in Autism Research and holds the Townsend Distinguished Chair in Research on Autism Spectrum Disorders at UT Southwestern. G.K. was partially supported by the James S. McDonnell Foundation 21st Century Science Initiative in Understanding Human Cognition Scholar Award (220020467), the Simons Foundation for Autism Research Award (947591), and NIH grants (NS126143, HG011641, MH126481, MH134809, NS115821). D.G. was partially supported by NIH funding (R01AG078713). S.A.L. is supported by the NIH (R01EY033353), the Carol and Gene Ludwig Family Foundation, the Collaborative Center for XDP through Massachusetts General Hospital and the Cure Alzheimer's Fund. Q.M. is a professor and the Chief of the Bioinformatics and Computational Biology Section in the Department of Biomedical Informatics at Ohio State University. This work was partially supported by awards R01GM131399, R21HG012482 and U54AG075931 from the National Institutes of Health. This work was also supported by the Pelotonia Institute of Immuno-Oncology. I.A. is a professor and the Department Chair at the Department of Neuroimmunology at the Center for Brain Research at the Medical University of Vienna, Austria and is a group leader at the Karolinska Institutet in Stockholm, Sweden. I.A. was supported by the Swedish Research Council, the Knut and Alice Wallenberg Foundation, ERC Synergy Grant KILL-OR-DIFFERENTIATE (856529, ERC-2019-SyG), the Austrian Science Fund (FWF) and EMBO Young Investigator Grants, an Alex's Lemonade Stand Foundation Crazy 8 Initiative Award, the Paradifference Foundation and Cancerfonden. Q.L. is an assistant professor of Neuroscience and Genetics, and J.Y. is a postdoctoral scholar at Washington University in St. Louis School of Medicine. This work was partially supported by the Whitehall Foundation (2021-08-003), a NARSAD Young Investigator Grant from the Brain and Behavior Research Foundation and NIH grant R01AG078512 to Q.L. and the German Research Foundation ImmunoSensation (EXC2151-390873048 to O.G.). This work was partially supported by awards R00NS111731 to A.B. and T32GM008243 and a CIRM-BSCRC Postdoctoral Fellowship to P.N., the Esther A. and Joseph Klingenstein Fund (to T.J.N.), the Shurl and Kay Curci Foundation (T.J.N.) and the Sontag Foundation (to T.J.N.) and a gift from the William K. Bowes Jr Foundation. T.J.N. is a New York Stem Cell Foundation Robertson Neuroscience Investigator. C.N.K. is supported by the UCSF Discovery Fellowship. This work was supported by the NIH Director's New Innovator Award 1DP2 NS127705-01 to M.G.F. T.B. thanks the founder of the Allen Institute, Paul G. Allen, for his vision, encouragement and support. The work was supported by grants from an NDCN Chan Zuckerberg Initiative grant (to O.G.); C.A.W. was supported by the NINDS (NS03245) and the Allen Discovery Center program, a Paul G. Allen Frontiers Group-advised program of the Paul G. Allen Family Foundation. C.A.W. is an investigator at the Howard Hughes Medical Institute.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Omer Ali Bayraktar, Ozgun Gokce, Naomi Habib, Genevieve Konopka, Shane A. Liddelow or Tomasz J. Nowakowski.

Peer review information *Nature Neuroscience* thanks Konstantin Khodosevich and Kiavash Movahedi for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

[1]Department of Neuroimmunology, Center for Brain Research, Medical University of Vienna, Vienna, Austria. [2]Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden. [3]Allen Institute for Brain Science, Seattle, WA, USA. [4]Department of Biological Chemistry, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA. [5]Institute of Neuropathology, Medical Faculty, University of Freiburg, Freiburg, Germany. [6]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston Children's Hospital, Boston, MA, USA. [7]The Ken and Ruth Davee Department of Neurology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. [8]Faculty of Biotechnology and Food Engineering, Technion Israel Institute of Technology, Haifa, Israel. [9]Department of Neurological Surgery, University of California San Francisco, San Francisco, CA, USA. [10]Department of Anatomy, University of California San Francisco, San Francisco, CA, USA. [11]Department of Biomedical Informatics, College of Medicine, the Ohio State University, Columbus, OH, USA. [12]Pelotonia Institute for Immuno-Oncology, the James Comprehensive Cancer Center, the Ohio State University, Columbus, OH, USA. [13]Laboratory of Neurodevelopmental Systems Biology, Brain Mind Institute, School of Life Sciences, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. [14]Department of Neuroscience, Washington University in St. Louis School of Medicine, St. Louis, MO, USA. [15]Department of Genetics, Washington University in St. Louis School of Medicine, St. Louis, MO, USA. [16]Division of Molecular Neurobiology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden. [17]Max Planck Institute for Biological Intelligence, Martinsried, Germany. [18]Department of Neurology, Center for Translational and Computational Neuroimmunology, Columbia University, New York, NY, USA. [19]Signalling Research Centres BIOSS and CIBSS, University of Freiburg, Freiburg, Germany. [20]Department of Bioengineering, Stanford University, Stanford, CA, USA. [21]Department of Applied Physics, Stanford University, Stanford, CA, USA. [22]Chan Zuckerberg Biohub, San Francisco, CA, USA.

[23]Division of Genetics and Genomics, Manton Center for Orphan Disease, Boston Children's Hospital, Boston, MA, USA. [24]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [25]Howard Hughes Medical Institute, Boston Children's Hospital, Boston, MA, USA. [26]Wellcome Sanger Institute, Hinxton, UK. [27]Department of Old Age Psychiatry and Cognitive Disorders, University Hospital Bonn, Bonn, Germany. [28]German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany. [29]The Edmond and Lily Safra Center for Brain Sciences, the Hebrew University of Jerusalem, Jerusalem, Israel. [30]Department of Neuroscience, UT Southwestern Medical Center, Dallas, TX, USA. [31]Peter O'Donnell Jr. Brain Institute, UT Southwestern Medical Center, Dallas, TX, USA. [32]Neuroscience Institute, NYU Grossman School of Medicine, New York, NY, USA. [33]Department of Neuroscience and Physiology, NYU Grossman School of Medicine, New York, NY, USA. [34]Parekh Center for Interdisciplinary Neurology, NYU Grossman School of Medicine, New York, NY, USA. [35]Department of Ophthalmology, NYU Grossman School of Medicine, New York, NY, USA. [36]Department of Psychiatry and Behavioral Sciences, University of California San Francisco, San Francisco, CA, USA. [37]Weill Institute for Neurosciences, University of California San Francisco, San Francisco, CA, USA. [38]The Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, University of California San Francisco, San Francisco, CA, USA. ✉e-mail: ob5@sanger.ac.uk; ozgun.goekce@ukbonn.de; naomi.habib@mail.huji.ac.il; genevieve.konopka@utsouthwestern.edu; shane.liddelow@nyulangone.org; tomasz.nowakowski@ucsf.edu