Title
Perspectives on defining cell types in the brain

Authors
Mukamel, Eran A
Ngai, John

Publication Date
2019-06-01

DOI
10.1016/j.conb.2018.11.007

Peer reviewed

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Neurobiology

Perspectives on defining cell types in the brain

Eran A Mukamel[1] and John Ngai[2]

The diversity of brain cell types was one of the earliest observations in modern neuroscience and continues to be one of the central concerns of current neuroscience research. Despite impressive recent progress, including single cell transcriptome and epigenome profiling as well as anatomical methods, we still lack a complete census or taxonomy of brain cell types. We argue this is due partly to the conceptual difficulty in defining a cell type. By considering the biological drivers of cell identity, such as networks of genes and gene regulatory elements, we propose a definition of cell type that emphasizes self-stabilizing regulation. We explore the predictions and hypotheses that arise from this definition. Integration of data from multiple modalities, including molecular profiling of genes and gene products, epigenetic landscape, cellular morphology, connectivity, and physiology, will be essential for a meaningful and broadly useful definition of brain cell types.

Addresses
[1] Department of Cognitive Science, University of California, San Diego, CA 92037, United States
[2] Department of Molecular and Cell Biology, Helen Wills Neuroscience Institute, QB3 Functional Genomics Laboratory, University of California, Berkeley, CA 94720, United States

Corresponding authors: Mukamel, Eran A (emukamel@ucsd.edu), Ngai, John (jngai@berkeley.edu)

Current Opinion in Neurobiology 2019, **56**:61–68

This review comes from a themed issue on **Neuronal identity**

Edited by **Sacha Nelson** and **Oliver Hobert**

https://doi.org/10.1016/j.conb.2018.11.007

One of the basic roles of theory in biology is to identify meaningful groupings of individuals. Although descriptive, this work provides an organizing framework and conceptual foundation for hypotheses regarding causal mechanisms and organizing principles. The classification of organisms by natural historians such as Linnaeus, together with deep investigation of particular cases such as the Galapagos finches, set the stage for Darwin's theoretical paradigm of evolution. Similarly, a taxonomy of cell types in the brain is ultimately required to understand how neural circuits evolved to underpin complex behaviors. Here we argue that recent developments in high-throughput single cell molecular analysis will enable a new classification of brain cell types that is unprecedented in its completeness (comprising all cells across all brain regions), quantitative precision, and integration of multiple modalities of molecular regulation (e.g. transcriptome, epigenome), anatomy, connectivity and function. There are both conceptual and practical difficulties standing in the way of a comprehensive and accurate neuronal cell type atlas, and a universally accepted and fixed taxonomy may remain an elusive goal. Nevertheless, the impact of a high-quality cell census will be broadly felt across developmental, molecular, and even computational neuroscience.

Neuroscientists have identified vastly different neural circuits across brain regions and species, yet it is difficult to think of a functional circuit in any complex organism that does not involve multiple clearly identifiable, distinct neuronal cell types [1–4]. The striking diversity of neuron types in the mammalian brain suggests that complex behavior relies not only on expanded cell number but also on an increasing functional specialization that allows particular neuron types to play-specific information processing roles. If that is the case, then information processing in the brain can only be understood by identifying the specialized roles and interactions among brain cell types. Indeed, it is not unreasonable to suppose that the evolutionary processes by which sister cell types duplicate, differentiate and specialize are driven by key functional needs [5], and that understanding the relationships among brain cell types will help to organize our understanding of the information processing functions which these circuits evolved to support.

Cell types as self-stabilizing regulatory programs

Different research communities use the concept of cell type in divergent ways, ranging from highly-specific notions of identifiable single neurons in Drosophila [6,7] to formulations-based on the connectivity [8] or the functional or computational role of a neuron [9]. Although different cell categorizations may be appropriate for different purposes, this diversity of definitions raises the question of whether there is a general, if not universal, notion of cell type that can be useful for a broad range of neuroscientific questions. Indeed, gene expression, cell location, morphology, connectivity, and physiology entail independent measurements but these properties are rarely independent variables.

Here we propose to define a cell type as a self-stabilizing system composed of specific genetic and developmental

processes. We are inspired by Waddington's original concept of canalization, that is the process by which a phenotypic outcome is produced and stabilized in the face of a range of environmental or genetic perturbations [10]. Accordingly, we suggest that each cell type corresponds to a self-stabilizing regulatory program, which acts to maintain and restore the cell type-specific program of gene expression (see Sidebar). We focus on self-stabilizing gene regulation, that is interactions among cell-intrinsic factors (principally genes and epigenetic marks) that form recurrent functional networks with feedback loops that preserve their structure. It is likely that, in some cases, structural factors (e.g. laminar position or connectivity) or physiological interactions (e.g. thalamic innervation) could be important for maintaining cellular identity as well [11]. In contrast with the core cell type-defining features of a cell type, other downstream properties, such as the expression of effector genes [12], may vary over time or in response to extrinsic signals but will generally revert to a canonical pattern induced by the self-stabilizing program.

The concept of dynamical stability is familiar in computational neuroscience, where it forms the basis of attractor networks for memory storage such as the Hopfield model [13]. A network of neurons, interacting via excitatory and inhibitory synapses, can encode multiple attractor states which can be retrieved through appropriate dynamics [14,15]. Similarly, networks of DNA binding transcription factor proteins can regulate their own expression to form stable, recurrent attractors [16•]. An attractor of a dynamical system need not be a static fixed point, but could correspond to a limit cycle with periodically repeating properties as in a mitotic cell type. Importantly, an attractor network provides homeostatic, self-stabilizing interactions that ensure that perturbations which move the state of the network away from an attractor will be compensated to restore the equilibrium. At the same time, such networks afford the opportunity for state changes when a signal or perturbation pushes the network out of the basin of attraction of one attractor and into a different equilibrium; experimentally, this is the basis for recent progress in transdifferentiation and artificial induction of neuronal cell types [17]. Despite the intuitive simplicity of the concept of a stable attractor, dynamical systems theory implies there may be important distinctions among cell types in terms of the type of attractor (e.g. fixed point, limit cycle, or chaotic strange attractor) [18] and the types of transitions (bifurcations) they may undergo [19].

Single cell assays advance the search for cell types

The recent development of high-throughput single cell transcriptomic (RNA-Seq) and epigenomic assays (mC-Seq for DNA methylation and ATAC-Seq for open chromatin) has raised expectations for a new and more detailed empirical analysis of brain cell types [20]. Indeed, pioneering studies, reviewed by [21], have been followed by a dramatic increase in the number of single cell transcriptome studies and in the number of profiled cells per study [22••,23••,24••,25,26•]. An example of the high resolution for fine cell type distinctions is a recent description of 133 cell types in two mouse neocortical regions [26•], a level of complexity that was hardly recognized before the advent of single cell transcriptomics. We expect that these data resources will enable fine-grained analysis of self-stabilizing cell type regulation across the diversity of brain cells.

While our proposal seeks a unifying conceptual basis for cell type, in practice we must acknowledge the limits of objectivity and the inevitability of disagreement: there will always be lumpers and splitters, each with valid empirical data and arguments to support their views. The seemingly simple question, 'How many brain cell types exist?' is more difficult to answer than it is to state. Instead of attempting to directly estimate the number of cell types, we propose to focus on objective empirical criteria that could contribute to meaningful discussion of cell type distinctions. One proposal would follow the tradition of systematics, which classifies species by cross-referencing as many traits or features as possible [27]. According to this view, any cell type classification should simultaneously account for the similarities and differences between cells along multiple dimensions, including gene expression (transcriptome), epigenomic state, anatomy (laminar location, dendritic and axonal morphology), connectivity, as well as electrophysiological properties. To this list, we would add that a comprehensive understanding of a cell type should include an account of its role in processing information in the context of local and distributed circuits. By emphasizing self-stabilizing features, we focus on those cellular features most likely to be central to the functional role of a cell type and therefore to co-vary with many other cellular properties.

Cell type definitions: predictions and testable hypotheses

Our proposal to base the definition of cell type in the concept of a dynamically stable state implies several key empirical predictions. We would predict that some features of a cell type, such as expression levels of core transcription regulating genes and corresponding configurations of epigenetic marks, are critical for stabilizing the cell's identity. These cell type-defining characteristics appear to emerge after the final mitotic division, as the cell enters a developmental ground state [28]. Other features, such as the expression of activity-dependent genes [29], may vary across cells or over time without altering the cell's stable type.

Dynamic gene expression across cells of the same type might be observed by using single cell transcriptome data

to estimate the rate of change of gene expression, or 'RNA velocity' [30•]. Computational and statistical modeling of transcriptomic measurements from a range of neuron types could indicate which transcription factors are the core regulators of cell type identity [7]. Such models might take inspiration from biophysical simulations of neuronal electrical dynamics, which showed that specific conserved electrophysiological behaviors such as central pattern generators can be produced by a variety of combinations of molecular components [31]. Similarly, the core transcriptional regulators of cortical neuron identity were inferred from single cell sequencing data [22••,23••,24••,32,33••]. Such information could provide testable causal hypotheses, for example predicting combinations of transcription factors which may be used for direct reprogramming [17,34].

Perturbative experiments — such as transgenic manipulations — provide a powerful approach for demonstrating the causal role of specific molecular regulators in generating self-stabilizing cell types with distinct functional properties. For example, a network of key transcription factors comprising Ctip2, Fezf2, Satb2 and Sox5 were shown to determine Layer 5/6 projection neuron fates in the developing mouse neocortex [35–37], while Sox6 is critical for specifying interneuron cell fates from spatially distinct progenitor cells [38]. Similarly, a single transcription factor, Brn3, was shown to both establish and maintain neuronal identity of a population of medial habenula neurons [39]. Recent studies suggest that post-transcriptional regulation, including alternative splicing or RNA modifications such as methyl-6-adenosine, play a role in shaping neurons' overt phenotype [40,41]. Given the recent progress in single cell transcriptomic and epigenomic assays that can subdivide neurons into types with increasing granularity, we may soon be able to propose causal experiments that could validate fine cell type distinctions by perturbing their putative (combinations of) regulators. In the interim, it would be useful to distinguish between provisional cell types defined by single cell -omics alone, and verified cell types for which more detailed causal data and mechanisms of stabilization are available.

Epigenetic regulation and self-stabilizing networks

Self-stabilizing gene regulation can be achieved through direct interaction among transcription factors [42,43], and these connections may be discerned in transcriptomic data through co-expression analysis. However, epigenetic modifications of histone proteins and DNA play important roles in regulating gene expression as well. Therefore, understanding the epigenetic landscape at the single cell level, in parallel with single cell transcriptome analysis, provides an independent evaluation of the cell's molecular identity. Epigenomic data are particularly useful for neurons, which have uniquely abundant non-CG DNA methylation as well as hydroxymethylation [44,45]. These epigenetic marks can have complex consequences, including roles in both repressing and enhancing transcription factor activity [46], which could impact the self-stabilizing interactions among core cell type defining factors. In some cases, the same gene is expressed in multiple neuron types as a consequence of different, cell type-specific epigenetic drivers. The use of unique combinations of enhancers in each cell type can lead to differences in the regulation of gene expression, isoform usage [47,48], or plasticity and activity dependence of expression [49]. Recent advances in single cell bisulfite sequencing [50••,51] and chromatin accessibility profiling [52••,53••,54,55] make it possible to measure these epigenomic signatures in thousands of cells, directly complementing large-scale transcriptomic studies [20,25,51].

Currently, larger datasets and more studies are available for single cell transcriptomes than epigenomes (DNA methylation and open chromatin). Although single cell epigenomic assays do not necessarily resolve a greater number of cell types, they provide an additional layer of regulatory information that is not available from RNA-Seq alone. For example, single cell DNA methylomes and ATAC-Seq data can indicate cell type-specific active enhancer regions [50••,56]. Such epigenomic information is valuable as a stable marker of cell identity that may be less sensitive than gene expression to changes in cell state, including circadian rhythms and activity-dependent gene induction [29]. New methods for joint measurement of epigenomic and transcriptomic information in the same cell could allow linking these complementary measures [57,58]. Given the more recent development of single cell epigenomic methods, we expect continued rapid improvements in these assays to enable recognition of cell types with similar resolution compared with the more mature single cell RNA-Seq methods.

Cell type classification based on these molecular data remains challenging, in part due to the sensitivity of results to different experimental and analytic parameters [29]. Computational methods for combining information across experimental batches and even across data modalities hold promise for reaching broad consensus [59,60]. Methods for statistical cross-validation of cell types from independent datasets are also critical [61•,62].

As rich, high-quality molecular data resources become increasingly available for neurons as well as other cell types (e.g. Human Cell Atlas [63]), we expect increasing opportunities to apply sophisticated machine learning and artificial neural network-based analyses to the challenge of unsupervised and semi-supervised learning of structure in these datasets [64–66]. Directly complementing the transcriptomic and epigenomic data, assays of chromosome conformation (e.g. Hi–C) at the single cell level can identify intra- and interchromosomal

interactions that may represent the 4-dimensional physical manifestation of gene expression regulatory networks in the nucleus [67,68].

Integrating data from multiple modalities: benefits and challenges

Despite the power of high-throughput single cell transcriptomic and epigenomic assays, these modalities likely do not fully reflect all the relevant differences between cell types. For example, transient expression of genes during differentiation can lead to situations in which mature neurons with different projection patterns show indistinguishable transcriptomic profiles [6]. Such distinctions may be evident at the epigenomic level, but in some cases information about anatomy, morphology, connectivity and/or physiology will be needed to fully distinguish neuronal cell types. Indeed, the concept of a self-stabilizing network need not be limited to genes and epigenetic regulators, but could include physiological feedback loops [69].

From these considerations, we argue that there will be substantial benefit from integrating empirical data across multiple modalities, both as a means of cross-validation and to provide greater precision and accuracy in cell type assignments. Methods that can provide spatial context for transcriptomic data via computational analysis [70] and *in situ* sequencing [71] can help connect molecular and anatomical information. Moreover, molecular data can provide a stable platform on which to build and assemble more detailed information from other modalities that are needed to complete a comprehensive characterization of cell types. For example, data that define a cell type at the molecular level can be used to create transgenic lines-based on genes or enhancers that are specifically active in that cell [72,73]. Such tools provide the experimentalist with genetic access, allowing analysis of the cell's functional properties. They would also provide a means of perturbing the cell genetically, optogenetically and chemogenetically to test predictions about the cell's identity and function-based on its transcription and signaling factor networks, connectivity, or physiology. Alternatively, unbiased experimental techniques for simultaneous measurement of physiology and gene expression (e.g. Patch-Seq) may help to link these modalities [74–76]. Such multi-modal data showed, for example, that an apparently continuous distribution of parvalbumin expressing neurons in the dorsal striatum has a corresponding continuous gradient of electrophysiological properties [74].

In contrast to molecular modalities, however, quantitative measurement of anatomy, physiology and connectivity is less straightforward. Analysis of these data requires first defining the key, cell type-defining features and further estimating these from the data, tasks for which no settled consensus is available. For example, optical microscopy can provide information about dendritic and axonal morphology and connectivity, but extracting these parameters is challenging due to the high dimensionality of image data. Here again we expect that sophisticated computational tools, in particular computer vision algorithms such as deep convolutional neural networks (CNNs), will play a key role in objectively quantifying traditional parameters of cellular morphology and connectivity with high-throughput and across the entire brain. Indeed, neural networks have already proved useful for automating and improving the annotation and low-level quantification of anatomical data from high-throughput brain-wide anatomical data [77,78]. These computational approaches may also help to discover new regularities and features of neuronal anatomy that may not be easily extracted and quantified in traditional, manual analyses.

Outlook

The challenges of measuring and integrating data about neuronal cell types across multiple modalities are large and exceed the capabilities of any one research group. Moreover, the value of a cell type atlas is directly tied to its broad acceptance and utilization by the neuroscience community. For these reasons, collaboration is critical, and collaborative consortia, including the NIH BRAIN Initiative Cell Census Network (BICCN) to which we contribute, aim to build consensus through joint development of cell type resources. Building on the experience of the previous BRAIN Initiative Cell Census Consortium [20], BICCN members contribute a broad range of expertise and techniques, including high throughput single cell transcriptomic and epigenomic assays which will form the basis for a comprehensive survey of molecular cell types in the mouse brain. In parallel, complementary data from human and non-human primates, though less comprehensive, will enable comparative and evolutionary perspectives on mammalian brain cell types [79,80]. Importantly, the BICCN is also investing in large-scale anatomy, morphology, connectivity, as well as physiological investigations. The long-term, aspirational goal is to integrate information from all of these modalities to provide a comprehensive taxonomy of brain cell types. Although this collaborative and coordinated strategy is critical, organizing a multi-site, multi-platform analysis does bring challenges of coordination, consistency in experimental protocols, analysis, metadata and data formats, and effective distribution of data to the larger scientific community. In addition to the BICCN, other consortia such as the HCA and the Human Biomolecular Atlas Program (HuBMAP) are also attempting to address these challenges and to build large-scale cell type resources [63,81]. These consortia represent experiments in organizing neuroscience research, and the sociological experience and lessons of these endeavors will be valuable dividends of the projects.

Here we have outlined a conceptual perspective emphasizing the self-stabilizing nature of cell types, which may help to guide otherwise subjective debates about the appropriate division between cell types. As Darwin recognized, it is healthy for lumpers and splitters to challenge each other over the empirical and theoretical validity of their frameworks. Their arguments will not go away, but with new molecular and functional data and more data-driven theoretical constructs we look forward to a productive discourse in the years ahead.

SIDEBAR

Self-stabilizing regulation of olfactory neuron identity

Our proposed definition of cell type emphasizes self-stabilizing regulation. Here we explore the relationship between feedback, stability and deterministic versus stochastic processes. Feedback is a key feature of recurrent networks, that is systems in which each element causally affects, and is reciprocally affected by, activity in other elements of the network. Perhaps the simplest paradigmatic example from computational neuroscience is the interaction between excitatory and inhibitory neural populations, as in the classic Wilson–Cowan model [82]. Here, increased spiking activity in a population of excitatory neurons stimulates increased activity in locally connected inhibitory neurons, which in turn leads to inhibition of the excitatory neurons. Depending on synaptic weights, such a system can have different attractors, that is static or periodically varying (oscillating) trajectories that are stable against perturbations [83]. Negative feedback between the two neural populations ensures that any external or internal stimulus, such as afferent input from another neural population, may temporarily move the network away from the attractor but will ultimately be reversed. This basic model explains the dynamic stability of neural circuits which maintain activity levels within a physiologically normal range, and it can explain the breakdown in equilibrium and the generation of epileptic activity following the loss of inhibition.

Complex networks with more than two elements can stabilize multiple stable attractors, that is alternative configurations which are each stable against limited perturbations. A simple type of attractor network appears to govern the choice of a single odorant receptor (OR) out of >1000 OR genes during olfactory sensory neuron differentiation [84]. Here, the initial expression of an OR gene results from stochastic demethylation of overlying repressive histone methylation marks. Expression of a functional OR leads to expression of Adcy3, which in turn promotes the OR expression (positive feedback) and also prevents derepression of other OR genes by inhibiting histone demethylation (negative feedback) [85••]. This 'epigenetic trap' mechanism ensures that only a single OR is expressed in each mature olfactory sensory neuron, and it further enables a cell that fails to express a functional OR to select a new OR for activation. This simple network shows how thousands of distinct attractor states, each corresponding to expression of a single OR, can be encoded and stabilized through epigenetic feedback.

Our concept of dynamic stability should be distinguished from the question of deterministic versus stochastic regulation of gene expression in single cells [86]. A self-stabilizing system can have stochastic state transitions, as in the stochastic selection of a single OR gene during olfactory sensory neuron differentiation. Moreover, random fluctuations in gene expression, for example due to transcriptional bursting, can cause ongoing stochastic differences between the molecular state of individual cells. However, the self-stabilizing dynamics of the core regulators of cell identity serve as a buffer that will prevent stochastic fluctuations from altering the critical, self-reinforcing pillars of the cell's machinery.

Conflict of interest statement

Nothing declared.

Acknowledgments

We thank our BICCN colleagues for ongoing discussions and Dr. Bosiljka Tasic and Dr. Chongyuan Luo for comments on the manuscript. The authors were supported in part by NIH BRAIN Initiative grants U19 MH114830 (J.N.) and U19 MH114831 (E.A.M.).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Llinás RR: **The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function**. *Science* 1988, **242**:1654-1664.
2. Kepecs A, Fishell G: **Interneuron cell types are fit to function**. *Nature* 2014, **505**:318-326.
3. MacNeil MA, Masland RH: **Extreme diversity among amacrine cells: implications for function**. *Neuron* 1998, **20**:971-982.
4. Yu JY, Kanai MI, Demir E, Jefferis GSXE, Dickson BJ: **Cellular organization of the neural circuit that drives Drosophila courtship behavior**. *Curr Biol* 2010, **20**:1602-1614.
5. Arendt D, Musser JM, Baker CVH, Bergman A, Cepko C, Erwin DH, Pavlicev M, Schlosser G, Widder S, Laubichler MD *et al.*: **The origin and evolution of cell types**. *Nat Rev Genet* 2016, **17**:744-757.
6. Li H, Horns F, Wu B, Xie Q, Li J, Li T, Luginbuhl DJ, Quake SR, Luo L: **Classifying Drosophila olfactory projection neuron subtypes by single-cell RNA sequencing**. *Cell* 2017, **171**:1206-1220.e22.
7. Konstantinides N, Kapuralin K, Fadil C, Barboza L, Satija R, Desplan C: **Phenotypic convergence: distinct transcription factors regulate common terminal features**. *Cell* 2018, **174**:622-635.e13.
8. Seung HS, Sümbül U: **Neuronal cell types and connectivity: lessons from the retina**. *Neuron* 2014, **83**:1262-1272.
9. Sanes JR, Masland RH: **The types of retinal ganglion cells: current status and implications for neuronal classification**. *Annu Rev Neurosci* 2015, **38**:221-246.
10. Waddington CH: *The Strategy of The Genes*. Routledge; 2014.

11. Zhou X, Franklin RA, Adler M, Jacox JB, Bailis W, Shyer JA, Flavell RA, Mayo A, Alon U, Medzhitov R: **Circuit design features of a stable two-cell system**. *Cell* 2018, **172**:744-757.e17.

12. Hobert O: **Regulation of terminal differentiation programs in the nervous system**. *Annu Rev Cell Dev Biol* 2011, **27**:681-696.

13. Hopfield JJ: **Neural networks and physical systems with emergent collective computational abilities**. *Proc Natl Acad Sci U S A* 1982, **79**:2554-2558.

14. Wills TJ, Lever C, Cacucci F, Burgess N, O'Keefe J: **Attractor dynamics in the hippocampal representation of the local environment**. *Science* 2005, **308**:873-876.

15. Amit DJ, Gutfreund H, Sompolinsky H: **Spin-glass models of neural networks**. *Spin Glass Theory and Beyond*. World Scientific; 1986:416-427.

16. • Lang AH, Li H, Collins JJ, Mehta P: **Epigenetic landscapes explain partially reprogrammed cells and identify key reprogramming genes**. *PLoS Comput Biol* 2014, **10**:e1003734.
A model inspired by statistical physics explains how epigenetic landscapes form from the interaction and positive feedback of transcription factor genes.

17. Tsunemoto R, Lee S, Szűcs A, Chubukov P, Sokolova I, Blanchard JW, Eade KT, Bruggemann J, Wu C, Torkamani A *et al.*: **Diverse reprogramming codes for neuronal identity**. *Nature* 2018, **557**:375-380.

18. Strogatz SH: *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. CRC Press; 2018.

19. Ferrell JE Jr: **Bistability, bifurcations, and Waddington's epigenetic landscape**. *Curr Biol* 2012, **22**:R458-466.

20. Ecker JR, Geschwind DH, Kriegstein AR, Ngai J, Osten P, Polioudakis D, Regev A, Sestan N, Wickersham IR, Zeng H: **The BRAIN initiative cell census consortium: lessons learned toward generating a comprehensive brain cell atlas**. *Neuron* 2017, **96**:542-557.

21. Johnson MB, Walsh CA: **Cerebral cortical neuron diversity and development at single-cell resolution**. *Curr Opin Neurobiol* 2017, **42**:9-16.

22. •• Zeisel A, Hochgerner H, Lönnerberg P, Johnsson A, Memic F, van der Zwan J, Häring M, Braun E, Borm LE, La Manno G *et al.*: **Molecular architecture of the mouse nervous system**. *Cell* 2018, **174**:999-1014.e22.
This study provides one of the largest and most comprehensive single cell transcriptome datasets for the mouse central and peripheral nervous system, including 265 curated clusters of neuronal and non-neuronal cells. The authors provide a taxonomy, inferred spatial distribution of neuron types, and analyze the major transcription factors driving the diversity of cell types throughout the body. The data and analysis can be further explored through a useful computational interface, mousebrain.org.

23. •• Saunders A, Macosko EZ, Wysoker A, Goldman M, Krienen FM, de Rivera H, Bien E, Baum M, Bortolin L, Wang S *et al.*: **Molecular diversity and specializations among the cells of the adult mouse brain**. *Cell* 2018, **174**:1015-1030.e16.
This study used droplet based sequencing to produce a large dataset with comprehensive coverage of 9 cortical and subcortical regions of the adult mouse brain. The authors provide an online resource, dropviz.org, for browsing and analyzing the data.

24. •• Mayer C, Hafemeister C, Bandler RC, Machold R, Brito RB, Jaglin X, Allaway K, Butler A, Fishell G, Satija R: **Developmental diversification of cortical inhibitory interneurons**. *Nature* 2018, **555**:457-462.
By measuring single cell transcriptomes from the progenitors of interneurons in the ganglionic eminences during gestation, this study traces the molecular history of cortical inhibitory neurons. The data indicate that the distinct adult cell types derived from the medial, central and lateral ganglionic eminences arise from highly similar precursors which begin to differentiate upon becoming post-mitotic.

25. Hochgerner H, Zeisel A, Lönnerberg P, Linnarsson S: **Conserved properties of dentate gyrus neurogenesis across postnatal development revealed by single-cell RNA sequencing**. *Nat Neurosci* 2018, **21**:290-299.

26. • Tasic B, Yao Z, Graybuck LT, Smith KA, Nguyen TN, Bertagnolli D, Goldy J, Garren E, Economo MN, Viswanathan S *et al.*: **Shared and distinct transcriptomic cell types across neocortical areas**. *Nature* 2018, **563**:72-78.
A high resolution, fine-grained analysis identified 133 cell types in two mouse cortical regions using single cell transcriptomes and retrograde labeling of cells with defined projections. This study suggests a high level of complexity in terms of the number of cortical cell types. Glutamatergic neurons are particularly diverse, with distinct molecular signatures in the two cortical regions, compared with more uniform GABAergic populations.

27. Zeng H, Sanes JR: **Neuronal cell-type classification: challenges, opportunities and the path forward**. *Nat Rev Neurosci* 2017, **18**:530-546.

28. Fishell G, Heintz N: **The neuron identity problem: form meets function**. *Neuron* 2013, **80**:602-612.

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

30. • La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastriti ME, Lönnerberg P, Furlan A *et al.*: **RNA velocity of single cells**. *Nature* 2018, **560**:494-498.
By comparing the relative abundance of RNA fragments from mature (spliced) vs. immature (intron-containing) transcripts, this study shows that single cell transcriptome data can be used to estimate the rate and direction of transcriptional change in single cells. This information, termed RNA velocity, has the potential to provide a dynamic view of cell types and their state changes on the timescale of several hours.

31. Marder E, Goaillard J-M: **Variability, compensation and homeostasis in neuron and network function**. *Nat Rev Neurosci* 2006, **7**:563-574.

32. Paul A, Crow M, Raudales R, He M, Gillis J, Huang ZJ: **Transcriptional architecture of synaptic communication delineates GABAergic neuron identity**. *Cell* 2017, **171**:522-539.e20.

33. •• Tasic B, Menon V, Nguyen TN, Kim TK, Jarsky T, Yao Z, Levi B, Gray LT, Sorensen SA, Dolbeare T *et al.*: **Adult mouse cortical cell taxonomy revealed by single cell transcriptomics**. *Nat Neurosci* 2016, **19**:335-346.
A pioneering study of brain cell types in the mouse primary visual cortex using single cell transcriptomes, demonstrating the power of molecular data for high-resolution cell type identification.

34. Rackham OJL, Firas J, Fang H, Oates ME, Holmes ML, Knaupp AS, FANTOM Consortium, Suzuki H, Nefzger CM, Daub CO *et al.*: **A predictive computational framework for direct reprogramming between human cell types**. *Nat Genet* 2016, **48**:331-335.

35. Leone DP, Srinivasan K, Chen B, Alcamo E, McConnell SK: **The determination of projection neuron identity in the developing cerebral cortex**. *Curr Opin Neurobiol* 2008, **18**:28-35.

36. Lai T, Jabaudon D, Molyneaux BJ, Azim E, Arlotta P, Menezes JRL, Macklis JD: **SOX5 controls the sequential generation of distinct corticofugal neuron subtypes**. *Neuron* 2008, **57**:232-247.

37. Custo Greig LF, Woodworth MB, Galazo MJ, Padmanabhan H, Macklis JD: **Molecular logic of neocortical projection neuron specification, development and diversity**. *Nat Rev Neurosci* 2013, **14**:755-769.

38. Azim E, Jabaudon D, Fame RM, Macklis JD: **SOX6 controls dorsal progenitor identity and interneuron diversity during neocortical development**. *Nat Neurosci* 2009, **12**:1238-1247.

39. Serrano-Saiz E, Leyva-Díaz E, De La Cruz E, Hobert O: **BRN3-type POU homeobox genes maintain the identity of mature postmitotic neurons in nematodes and mice**. *Curr Biol* 2018, **28** 2813–2823.e2.

40. Wamsley B, Jaglin XH, Favuzzi E, Quattrocolo G, Nigro MJ, Yusuf N, Khodadadi-Jamayran A, Rudy B, Fishell G: **Rbfox1 mediates cell-type-specific splicing in cortical interneurons**. *Neuron* 2018, **100** 846–859.e7.

41. Frye M, Harada BT, Behm M, He C: **RNA modifications modulate gene expression during development**. *Science* 2018, **361**:1346-1349.

42. Amati B, Land H: **Myc—Max—Mad: a transcription factor network controlling cell cycle progression, differentiation and death**. *Curr Opin Genet Dev* 1994, **4**:102-108.

43. Ding B, Cave JW, Dobner PR, Mullikin-Kilpatrick D, Bartzokis M, Zhu H, Chow C-W, Gronostajski RM, Kilpatrick DL: **Reciprocal autoregulation by NFI occupancy and ETV1 promotes the developmental expression of dendrite-synapse genes in cerebellar granule neurons**. *Mol Biol Cell* 2016, **27**:1488-1499.

44. Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, Lucero J, Huang Y, Dwork AJ, Schultz MD *et al.*: **Global epigenomic reconfiguration during mammalian brain development**. *Science* 2013, **341** 1237905–1237905.

45. Kriaucionis S, Heintz N: **The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain**. *Science* 2009, **324**:929-930.

46. Luo C, Hajkova P, Ecker JR: **Dynamic DNA methylation: In the right place at the right time**. *Science* 2018, **361**:1336-1340.

47. Fuccillo MV, Földy C, Gökce Ö, Rothwell PE, Sun GL, Malenka RC, Südhof TC: **Single-cell mRNA profiling reveals cell-type-specific expression of neurexin isoforms**. *Neuron* 2015, **87**:326-340.

48. Zhang X, Chen MH, Wu X, Kodani A, Fan J, Doan R, Ozawa M, Ma J, Yoshida N, Reiter JF *et al.*: **Cell-type-specific alternative splicing governs cell fate in the developing cerebral cortex**. *Cell* 2016, **166**:1147-1162.e15.

49. Hu P, Fabyanic E, Kwon DY, Tang S, Zhou Z, Wu H: **Dissecting cell-type composition and activity-dependent transcriptional state in mammalian brains by massively parallel single-nucleus RNA-Seq**. *Mol Cell* 2017, **68**:1006-1015.e7.

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

This study applied single cell bisulfite sequencing to obtain more than 6000 DNA methylomes from human and mouse frontal cortical neurons. The data showed that DNA methylation is highly cell type specific, with robust signatures of multiple excitatory and inhibitory neuron subtypes.

51. Luo C, Rivkin A, Zhou J, Sandoval JP, Kurihara L, Lucero J, Castanon R, Nery JR, Pinto-Duarte A, Bui B *et al.*: **Robust single-cell DNA methylome profiling with snmC-seq2**. *Nat Commun* 2018, **9**:3824.

52. •• Preissl S, Fang R, Huang H, Zhao Y, Raviram R, Gorkin DU, Zhang Y, Sos BC, Afzal V, Dickel DE *et al.*: **Single-nucleus analysis of accessible chromatin in developing mouse forebrain reveals cell-type-specific transcriptional regulation**. *Nat Neurosci* 2018, **21**:432-439.

Efficient single nucleus profiling of accessible chromatin identifies thousands of active regions per cell, including promoters and distal enhancers. These epigenomic signatures are cell type specific markers of neuron identity and can be used for unbiased classification.

53. •• Lake BB, Chen S, Sos BC, Fan J, Kaeser GE, Yung YC, Duong TE, Gao D, Chun J, Kharchenko PV *et al.*: **Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain**. *Nat Biotechnol* 2018, **36**:70-80.

By generating single cell transcriptomes and chromatin accessibility profiles in the same tissue, this study begins to integrate transcriptomic and epigenomic information about thousands of brain cells.

54. Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, Steemers FJ, Trapnell C, Shendure J: **Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing**. *Science* 2015, **348**:910-914.

55. Cusanovich DA, Hill AJ, Aghamirzaie D, Daza RM, Pliner HA, Berletch JB, Filippova GN, Huang X, Christiansen L, DeWitt WS *et al.*: **A single-cell atlas of in vivo mammalian chromatin accessibility**. *Cell* 2018 http://dx.doi.org/10.1016/j.cell.2018.06.052.

56. He Y, Gorkin DU, Dickel DE, Nery JR, Castanon RG, Lee AY, Shen Y, Visel A, Pennacchio LA, Ren B *et al.*: **Improved regulatory element prediction based on tissue-specific local epigenomic signatures**. *Proc Natl Acad Sci U S A* 2017, **114**:E1633-E1640.

57. Cao J, Cusanovich DA, Ramani V, Aghamirzaie D, Pliner HA, Hill AJ, Daza RM, McFaline-Figueroa JL, Packer JS, Christiansen L *et al.*: **Joint profiling of chromatin accessibility and gene expression in thousands of single cells**. *Science* 2018, **361**:1380-1385.

58. Angermueller C, Clark SJ, Lee HJ, Macaulay IC, Teng MJ, Hu TX, Krueger F, Smallwood SA, Ponting CP, Voet T *et al.*: **Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity**. *Nat Methods* 2016, **13**:229-232.

59. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R: **Integrating single-cell transcriptomic data across different conditions, technologies, and species**. *Nat Biotechnol* 2018, **36**:411-420.

60. Haghverdi L, Lun ATL, Morgan MD, Marioni JC: **Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors**. *Nat Biotechnol* 2018, **36**:421-427.

61. • Crow M, Paul A, Ballouz S, Huang ZJ, Gillis J: **Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor**. *Nat Commun* 2018, **9**:884.

A statistical method for validating cell types defined by single cell molecular data through cross-validation with independent data. The MetaNeighbor method uses a nearest neighbor graph to find pairs of similar cells in two datasets based on relevant gene sets or other features. The neighbor graphs are then used to attempt to reconstruct a cell type classification in one dataset based on information in the other, neighboring datasets, to assess the robust validity of the cell type.

62. Risso D, Purvis L, Fletcher RB, Das D, Ngai J, Dudoit S, Purdom E: **clusterExperiment and RSEC: a bioconductor package and framework for clustering of single-cell and other large gene expression datasets**. *PLoS Comput Biol* 2018, **14**:e1006378.

63. Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, Bodenmiller B, Campbell P, Carninci P, Clatworthy M *et al.*: **The human cell atlas**. *Elife* 2017, **6**.

64. Lopez R, Regier J, Cole MB, Jordan M, Yosef N: **Bayesian inference for a generative model of transcriptome profiles from single-cell RNA sequencing**. *bioRxiv* 2018 http://dx.doi.org/10.1101/292037.

65. Wang D, Gu J: **VASC: dimension reduction and visualization of single cell RNA sequencing data by deep variational autoencoder**. *bioRxiv* 2017 http://dx.doi.org/10.1101/199315.

66. Grønbech CH, Vording MF, Timshel PN, Sønderby CK, Pers TH, Winther O: **scVAE: variational auto-encoders for single-cell gene expression data**. *bioRxiv* 2018 http://dx.doi.org/10.1101/318295.

67. Ramani V, Deng X, Qiu R, Gunderson KL, Steemers FJ, Disteche CM, Noble WS, Duan Z, Shendure J: **Massively multiplex single-cell Hi-C**. *Nat Methods* 2017, **14**:263-266.

68. Stevens TJ, Lando D, Basu S, Atkinson LP, Cao Y, Lee SF, Leeb M, Wohlfahrt KJ, Boucher W, O'Shaughnessy-Kirwan A *et al.*: **3D structures of individual mammalian genomes studied by single-cell Hi-C**. *Nature* 2017, **544**:59.

69. Mullins C, Fishell G, Tsien RW: **Unifying Views of autism spectrum disorders: a consideration of autoregulatory feedback loops**. *Neuron* 2016, **89**:1131-1156.

70. Satija R, Farrell JA, Gennert D, Schier AF, Regev A: **Spatial reconstruction of single-cell gene expression data**. *Nat Biotechnol* 2015, **33**:495-502.

71. Lein E, Borm LE, Linnarsson S: **The promise of spatial transcriptomics for neuroscience in the era of molecular cell typing**. *Science* 2017, **358**:64-69.

72. Madisen L, Garner AR, Shimaoka D, Chuong AS, Klapoetke NC, Li L, van der Bourg A, Niino Y, Egolf L, Monetti C *et al.*: **Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance**. *Neuron* 2015, **85**:942-958.

73. Huang ZJ, Zeng H: **Genetic approaches to neural circuits in the mouse**. *Annu Rev Neurosci* 2013, **36**:183-215.

74. Muñoz-Manchado AB, Bengtsson Gonzales C, Zeisel A, Munguba H, Bekkouche B, Skene NG, Lönnerberg P, Ryge J,

Harris KD, Linnarsson S *et al.*: **Diversity of interneurons in the dorsal striatum revealed by single-cell RNA sequencing and PatchSeq**. *Cell Rep* 2018, **24**:2179-2190.e7.

75. Cadwell CR, Palasantza A, Jiang X, Berens P, Deng Q, Yilmaz M, Reimer J, Shen S, Bethge M, Tolias KF *et al.*: **Electrophysiological, transcriptomic and morphologic profiling of single neurons using Patch-seq**. *Nat Biotechnol* 2016, **34**:199-203.

76. Fuzik J, Zeisel A, Máté Z, Calvigioni D, Yanagawa Y, Szabó G, Linnarsson S, Harkany T: **Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes**. *Nat Biotechnol* 2016, **34**:175-183.

77. Kim Y, Venkataraju KU, Pradhan K, Mende C, Taranda J, Turaga SC, Arganda-Carreras I, Ng L, Hawrylycz MJ, Rockland KS *et al.*: **Mapping social behavior-induced brain activation at cellular resolution in the mouse**. *Cell Rep* 2015, **10**:292-305.

78. Kim Y, Yang GR, Pradhan K, Venkataraju KU, Bota M, García Del Molino LC, Fitzgerald G, Ram K, He M, Levine JM *et al.*: **Brain-wide maps reveal stereotyped cell-type-based cortical architecture and subcortical sexual dimorphism**. *Cell* 2017, **171**:456-469.e22.

79. Boldog E, Bakken TE, Hodge RD, Novotny M, Aevermann BD, Baka J, Bordé S, Close JL, Diez-Fuertes F, Ding S-L *et al.*: **Transcriptomic and morphophysiological evidence for a specialized human cortical GABAergic cell type**. *Nat Neurosci* 2018, **21**:1185-1195.

80. Hodge RD, Bakken TE, Miller JA, Smith KA, Barkan ER, Graybuck LT, Close JL, Long B, Penn O, Yao Z *et al.*: **Conserved cell types with divergent features between human and mouse cortex**. *bioRxiv* 2018 http://dx.doi.org/10.1101/384826.

81. *The Human BioMolecular Atlas Program - HuBMAP | NIH Common Fund*. 2018 https://commonfund.nih.gov/hubmap.

82. Wilson HR, Cowan JD: **Excitatory and inhibitory interactions in localized populations of model neurons**. *Biophys J* 1972, **12**:1-24.

83. Izhikevich EM: *Dynamical Systems in Neuroscience*. MIT Press; 2007.

84. Dalton RP, Lomvardas S: **Chemosensory receptor specificity and regulation**. *Annu Rev Neurosci* 2015, **38**:331-349.

85. •• Lyons DB, Allen WE, Goh T, Tsai L, Barnea G, Lomvardas S: **An epigenetic trap stabilizes singular olfactory receptor expression**. *Cell* 2013, **154**:325-336.
This paper offers a mechanistic explanation for the expression of a single odorant receptor in each mammalian olfactory sensory neuron, one of the most impressive examples of the regulation of neuronal cell type diversity. A negative feedback between two genes, Adcy3 and Lsd1, ensures that when one odorant receptor is expressed, all of the hundreds of other receptor genes remain epigenetically repressed. The self-stabilizing interaction between gene expression and epigenetic regulation ensures the proper functioning of the sensory neurons.

86. Symmons O, Raj A: **What's luck got to do with it: single cells, multiple fates, and biological nondeterminism**. *Mol Cell* 2016, **62**:788-802.