

REVIEW SUMMARY

NEURODEVELOPMENT

Looking at neurodevelopment through a big data lens

James Briscoe* and Oscar Marin*

BACKGROUND: The formation of the nervous system represents an astonishing feat of self-organization that is compromised in neuropsychiatric conditions such as autism and schizophrenia. Despite impressive progress in neuroscience over the past decades, our understanding of how billions of neurons come together to form the nervous system and enable function and behavior is in its infancy, especially when it comes to the human brain. However, the field is at a turning point. The introduction of new technologies that produce large volumes of high-resolution measurements—big data—has the potential to revolutionize the study of brain development.

ADVANCES: A foundation of developmental neuroscience is the detailed and systematic description of the nervous system. New methods are documenting the cellular composition and organization of neural tissue with ever-increasing resolution. The development of high-throughput and automated microscopy methods is charting the connectivity of thousands of neurons, delin-

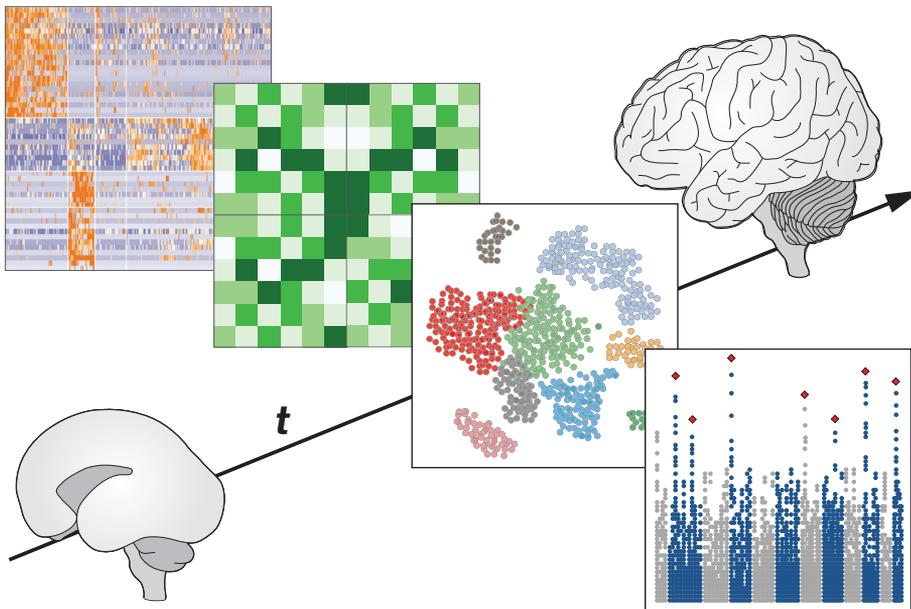
eating the structure of whole regions of the nervous system. Technology is also emerging for the large-scale analysis of the activity of entire brain regions. But probably the most obvious impact of high-throughput techniques is in the development of single-cell molecular profiling. A variety of approaches are being used to produce genome-wide molecular surveys with single-cell resolution from adult and embryonic tissue. In particular, transcriptomic analyses of thousands to millions of cells are providing an unprecedented molecular characterization of the brain, revealing previously unrecognized cell types, allowing interspecies comparisons, and suggesting mechanisms that account for the developmental origin of the diversity and function of neural cell types.

Most neuropsychiatric disorders have a prominent heritable component and arise from the altered developmental processes during the formation of the nervous system. Recent advances in human genetics are beginning to shed light on the genetic architecture of these disorders and suggest how genetic varia-

tion confers susceptibility to disease. Large-scale sequencing studies have revealed causes that range from large-effect heterozygous mutations to highly polygenic conditions. In addition, the contribution of de novo somatic mutations to neurodevelopmental diseases is being recognized. Nevertheless, progressing from genetic findings to underlying biological mechanisms has proved challenging, not least because in many cases identifying the cells relevant to a disease process has been difficult. In this context, a convergence between neurogenetics and developmental neurobiology, driven by the increased resolution of the molecular and genetic assays, is likely to improve our understanding of the origin of neurological disorders and provide insight into basic developmental mechanisms.

Whereas new molecular and genomic tools contribute to the identification of plausible neurobiological mechanisms, methods based on the directed development of pluripotent stem cells offer experimental access to developing human neural tissue to test hypotheses. Rapid progress is being made in the development of techniques that produce specific neural cell types or more complex mixtures of cell types that mimic the development of specific regions of the central nervous system. Questions remain about the accuracy of these in vitro models, and validation and refinement continue. Notwithstanding this uncertainty, the potential to study the etiology of neurological disorders in human neural tissue is already providing important insights.

OUTLOOK: New perspectives are emerging on long-standing questions about the ontogeny, composition, and function of the nervous system. They are addressing fundamental conceptual questions, such as what constitutes a cell type, and revealing biological mechanisms responsible for neurological disorders. The comparison of nervous system development between multiple individuals could conceivably identify individual variation in, for example, neural connectivity patterns that underpin behavioral individuality and may enable the investigation of the complex biological mechanisms underlying individuality. Certainly, integrating data from anatomical, developmental, genetic, and molecular studies has the potential to link cellular processes to functional and behavioral consequences. This strategy would provide fundamental insight and offer a new vision to the field. ■



Big data approaches in developmental neurobiology. New technologies that produce large volumes of high-resolution measurements are documenting gene expression, connectivity, and function in the developing brain with an unprecedented level of detail. In combination with large-scale genetic studies, big data approaches are transforming our ability to interrogate the developing brain and identify causal mechanisms for its associated disorders.

The list of author affiliations is available in the full article online.

*Corresponding author. Email: james.briscoe@crick.ac.uk (J.B.); oscar.marin@kcl.ac.uk (O.M.)

Cite this article as J. Briscoe, O. Marin, *Science* 369, eaaz8627 (2020). DOI: 10.1126/science.aaz8627

S READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.aaz8627>

REVIEW

NEURODEVELOPMENT

Looking at neurodevelopment through a big data lens

James Briscoe^{1*} and Oscar Marín^{2,3*}

The formation of the human brain, which contains nearly 100 billion neurons making an average of 1000 connections each, represents an astonishing feat of self-organization. Despite impressive progress, our understanding of how neurons form the nervous system and enable function is very fragmentary, especially for the human brain. New technologies that produce large volumes of high-resolution measurements—big data—are now being brought to bear on this problem. Single-cell molecular profiling methods allow the exploration of neural diversity with increasing spatial and temporal resolution. Advances in human genetics are shedding light on the genetic architecture of neurodevelopmental disorders, and new approaches are revealing plausible neurobiological mechanisms underlying these conditions. Here, we review the opportunities and challenges of integrating large-scale genomics and genetics for the study of brain development.

A foundation of neuroscience is the description of the nervous system. In recent years, new methods are documenting gene expression in human and animal models with ever-increasing resolution in space and time, and technology is emerging that allows the structural and functional dissection of neural circuits. This is beginning to provide an unprecedented molecular characterization of the brain. At the same time, high-throughput genome sequencing and large-scale genetic studies are revealing how genetic variation confers susceptibility to neuropsychiatric disorders, and methods based on pluripotent stem cells hold the potential to offer insight into the neurobiological basis of these conditions by providing experimental access to the developing human nervous system. The use of big data approaches is therefore transforming our ability to document how cell diversity emerges in the nervous system, with implications for our understanding of neurodevelopmental disorders. A key challenge is to integrate molecular, cellular, and genomic data. For this, appropriate developmental models will be needed so that emerging genetic findings yield mechanistic insight into the neurobiological basis of neuropsychiatric disorders.

Documenting the component parts

The morphological identification of neuronal types and their connectivity has a long tradition, dating back to Ramón y Cajal. The development of high-throughput light and electron microscopy methods and the increasing automation of image analysis indicate that this field has entered the big data era, and progress

continues apace. Three-dimensional electron microscopy reconstructions of neuronal connectivity in the *Drosophila melanogaster* brain and regions of the mouse brain comprehensively chart the location and connectivity of hundreds to thousands of neurons (Fig. 1) (1, 2). This is providing insight into the logic of nervous system wiring. For example, a reconstruction of the *D. melanogaster* larval nervous system revealed an architecture in which multiple sensory inputs converge to enable the rapid selection of escape responses (3). The speed and accuracy with which these massive datasets can be reconstructed are being improved by applying the latest machine learning methods (4). A challenge now is to link detailed maps of morphology and connectivity to molecular identity, function, and developmental origin.

To complement these structural plans of nervous systems, ways to scale physiological methods that assay the activity of neurons in situ are being developed. Advances in semiconductor technology and fabrication have enabled engineering of high-density silicon probes that allow the simultaneous recording from hundreds of neurons at fine spatial and temporal scales (5). Complementary to this are optical imaging approaches that take advantage of voltage or calcium indicators for in vivo imaging of the activity of hundreds and thousands of neurons (6–8). This has already allowed whole-brain imaging of optically accessible zebrafish larvae (9). Combining functional imaging with genetic markers for specific cell types in the zebrafish spinal cord tracked the assembly of motor circuits from the birth of the first neurons to the emergence of coordinated activity (10). The analysis revealed that motor neurons are the first to acquire physiological activity, and this guides the establishment of patterned activity in the rest of the local circuitry, suggesting

an explanation for why motor neurons differentiate earlier and at a higher rate in the developing spinal cord than the interneurons to which they connect (11). Advances in hardware, fluorophores, and computational methods will broaden the applicability of these approaches.

Perhaps the most obvious impact of high-throughput techniques is in the development of single-cell molecular profiling. New methods allow single-cell resolution genome-wide molecular surveys of cells captured from adult and embryonic tissue. A burgeoning number of studies describe the systematic profiling of the transcriptomes of thousands to millions of cells. Broad sampling of the nervous system (12–14) facilitates comparisons of gene expression in neuronal subtypes from different regions of the central nervous system (CNS). This has implicated common transcriptional programs in the maintenance and elaboration of axons and presynaptic terminals in different neuronal subtypes (14). Other studies focus on smaller regions to provide higher-resolution maps (15, 16), revealing that “phenotypic convergence,” in which distinct combinations of transcription factors are used in different cell types to control the expression of similar effector genes such as neurotransmitters, appears to be a widespread phenomenon (17). Detailed analyses of specific CNS regions also enable direct comparisons of cell identities across species (18) and the identification of more cell types than previously recognized, as shown, for instance in the rodent cerebral cortex (15).

There is increasing awareness of the importance of glia in the formation and function of neural circuits. Until recently, the diversity of these cell populations had been less well described, but single-cell RNA sequencing (scRNA-seq) studies are changing this. For instance, it has been shown that mature oligodendrocytes are transcriptionally heterogeneous, and distinct populations are present in different proportions in different brain regions (19). Similarly, multiple molecularly distinct astrocyte types, occupying regionally restricted territories, have been identified (20). Distinct types of astrocytes are associated with specific neural circuits, which suggests unique roles in modulating neuronal activity (21). Moreover, single-cell analyses have revealed specific time- and region-dependent subtypes of microglia (22), brain-resident macrophages, which are critically involved in wiring (23).

Together, these molecular profiling approaches raise the prospect of using the transcriptome to provide a comprehensive classification of cell types. This necessitates linking molecular identity to the location, morphology, connectivity, and function of neurons and glial cells. Conventional methods for

¹The Francis Crick Institute, London NW1 1AT, UK. ²Centre for Developmental Neurobiology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London SE1 1UL, UK. ³MRC Centre for Neurodevelopmental Disorders, King's College London, London SE1 1UL, UK. *Corresponding author. Email: james.briscoe@crick.ac.uk (J.B.); oscar.marin@kcl.ac.uk (O.M.)

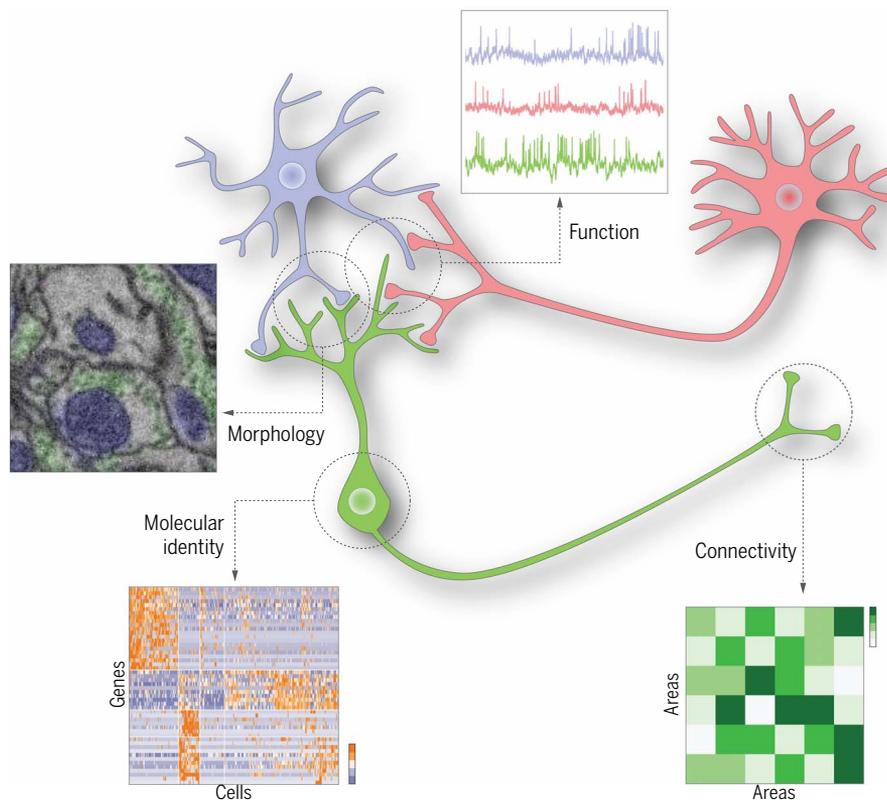


Fig. 1. Big data methods characterizing neuronal identity. Neurons can be classified by morphological, physiological, and molecular criteria and by their connectivity. Microfabricated silicon devices containing dense arrays of electrophysiology probes or genetically encoded voltage or calcium-activity indicators enable monitoring of the activity of hundreds of neurons simultaneously [traces adapted from (8)]. High-throughput electron microscopy and image processing allow the reconstruction of cellular morphology and synapses in regions of the central nervous system [image reproduced from (2)]. Combining the tagging of neurons with unique molecular barcodes and in situ methods to visualize them allows projections and connectivity of thousands of neurons to be mapped. Characterizing individual neurons by methods such as RNA-seq provides molecular catalogs of cell identity in the nervous system [image reproduced from (37), licensed under CC BY 4.0].

mapping gene expression in tissue have been low throughput. Newer approaches based on single-molecule fluorescent in situ hybridization (24), in situ sequencing (25), or spatially resolved capture of mRNA on slides (26) are scaling up efforts to connect the molecular identity of a neuron to its location within the nervous system. Methods that link a neuron's transcriptome to other features such as its connectivity, activity, or function are still in their infancy. For example, "Patch-seq" uses scRNA-seq to assay the transcriptome of cells following patch-clamp recording (27, 28), and "Connect-seq" combines retrograde viral tracing and single-cell transcriptomics to determine the molecular identity of neurons in a particular circuit (29). Moreover, DNA barcoding methods that link synaptic partners have the potential to reveal individual neuronal connectivity, although these do not yet have single-cell resolution (30). Hence, high-throughput analyses linking cell identity to connectivity

and function remain a challenge. Ultimately, however, these methods will be crucial in deciphering how neural circuits form and function in a coordinated manner.

Defining cell types

Alongside the technical challenges, the new approaches prompt conceptual questions about the definition of cell type. The prevailing view is that each type of neuron utilizes a specific set of "functional modules" (e.g., morphology, neurotransmitter, channels, synaptic connectivity, etc.) that collectively define its identity and that are regulated by dedicated transcriptional programs. A cell type is therefore determined and maintained by regulatory programs, or core regulatory complexes (31), acting to govern a cell type-specific program of gene expression (32). This definition is consistent with the idea that recursively linked gene regulatory networks execute specific gene expression programs to define cell identity (33).

But is knowing the transcriptome of a neuron sufficient to define its identity and predict its morphology, connectivity, and function? A comparison of the morphological, connectivity, and functional classification of the 302 neurons of *Caenorhabditis elegans* with their gene expression profiles revealed a strong correlation between molecular and anatomical definitions (34). Most of the identified neuronal classes were distinguishable by the specific combination of transcription factors they expressed. However, there are complications. The analysis revealed a hierarchical structure to neuronal identity in which some neuronal classes could be further partitioned into subclasses based on differences in subsets of gene expression and synaptic partners. A hierarchical organization with subdivisions of increasingly fine-grained subtypes is also apparent in nervous systems that are more complex than that of *C. elegans*. Perhaps the best-characterized example is the muscle-innervating motor neurons of the vertebrate spinal cord (35). These are segregated into discrete columns, each of which projects to distinct locations in the periphery. Each column is then further divided into motor pools that target a single muscle; within a motor pool, alpha and gamma neurons, innervating muscle fibers or spindles, respectively, can be further distinguished. Molecular differences, notably in the combination of transcription factors expressed, correlate with these anatomical and functional distinctions and support a hierarchical Linnaean-like taxonomy for cell type classification (35). This is consistent with the view that cell types can be stratified to various levels of granularity: the extent of which then becomes a question of the purpose of the classification.

Nevertheless, functionally and morphologically distinct neurons located in different regions of the nervous system, or even within the same region of the adult brain, can appear transcriptionally similar (Fig. 2). This raises the question of what the limits of the molecular definition of cell type identity are. To address this, datasets systematically assessing the correspondence between neuronal morphology, function, and gene expression will provide insight. Developmental history is also likely to be an important feature. For example, although differentiating *Drosophila* olfactory projection neurons have gene expression differences that establish their different projection patterns, these transcriptomic differences disappear in mature neurons despite continued differences in innervation patterns and morphology (36). Hence an understanding and analysis of ontogeny is crucial for developing a principled and comprehensive means to define cell type identity.

An evolutionary explanation has been proposed to define cell types (37). In this view, a new cell type arises in evolution from an existing cell

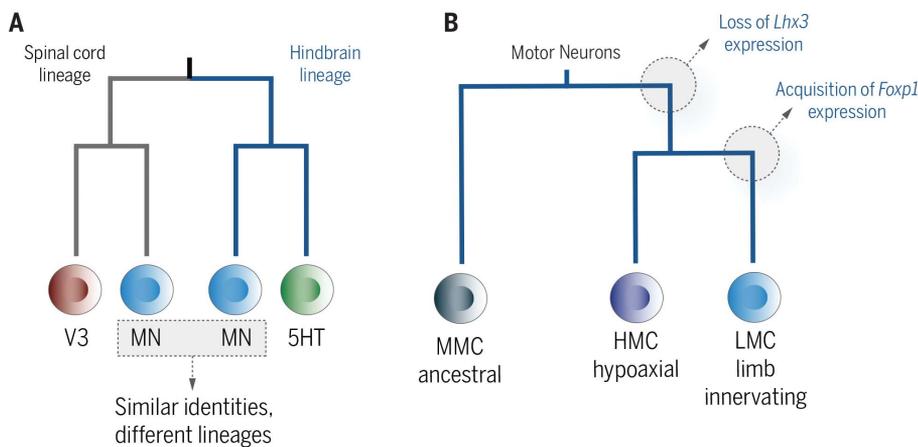


Fig. 2. Developmental and evolutionary lineages of neuronal subtype identity. (A) Example of a developmental lineage. Motor neurons (MN) are generated in both the spinal cord and hindbrain from progenitors with characteristic pedigrees that also include specific interneurons (V3) and serotonergic neurons (5HT), respectively. Despite the similarity of spinal cord and hindbrain MNs, they derive from distinct developmental lineages. (B) Example of the proposed evolutionary diversification of a neuronal subtype. Ancestral motor neurons are proposed to have had a medial motor column (MMC) identity; loss of expression of *LHX3* in a subset of these resulted in the acquisition of hypaxial motor column (HMC) identity, and then subsequent co-option of *FOXP1* gene expression produced lateral motor column (LMC) motor neurons that are responsible for innervating the limbs in tetrapods.

type by the gain, loss, or co-option of a new core regulatory complex—the set of transcriptional regulators that lead to the dedicated gene expression program that specifies the cell type (31) (Fig. 2). This implies a bifurcating hierarchy of cell types and emphasizes the importance of ontogeny in defining cell identity. Although this offers an elegant framework for cell type identity, it is possible that molecular demarcations between cell types are indistinct or that dynamic changes in gene expression, in response to activity or environmental cues, blur cell type definitions. Cross-species comparisons should provide insight into conserved identities and pinpoint the developmental and phylogenetic points at which changes in a core regulatory complex result in the emergence of a new neuronal subtype (37). Technical constraints will need to be overcome. Sequencing depth and sensitivity are currently limited, which may lead to the systematic loss of information, particularly of genes expressed at low levels. For example, downsampling data from the mouse cortex from 1,000,000 to 100,000 mapped reads per cell reduces the ability to separate neuronal classes (15). Interspecies comparisons could be confounded if sampling differences are conflated with ascertainment of identity. Conversely, current transcriptomic methods are destructive and result in static snapshots of gene expression that do not account for transient changes in gene expression. This could lead to the agglomeration of distinct cell types or the superfluous division of similar neurons

into distinct subtypes. It remains a possibility, however, that the absence of clear divisions between neuronal subtype identities is a reality and might reflect a fundamental property of nervous systems. Addressing this question will require further experimental and conceptual investigation, and this is of more than abstract interest because identifying neuronal subtypes is crucial for understanding the cellular basis of neuropsychiatric disorders.

The ontogeny of neuronal diversity

Although much progress in defining cell type diversity had been made with conventional molecular, genetic, and developmental biology studies, single-cell transcriptomics is having an impact. For instance, analysis of the developing spinal cord revealed a previously overlooked temporal program diversifying neuronal identity that operates alongside the spatial program (38). Similarly, the transcriptional similarities of embryonic and adult cortical γ -aminobutyric acid (GABA)-releasing interneurons suggest that their fate is established early during development by transcriptional programs that then unfold over the course of several weeks (39).

There have been notable successes with the computational inference of differentiation trajectories from scRNA-seq datasets. At the heart of these methods is the idea that asynchrony within a population of cells means that some cells are further along a differentiation pathway than others and the trajectory of this pathway can be reconstructed by identifying and

ordering cells using the differences in gene expression (40). For example, data from the developing human cortex has been used to order the transition from radial glia cells through intermediate progenitors to maturing glutamatergic neurons. This has led to the suggestion that cell fate decisions occur prior to S phase of the cell cycle (41), although recent data indicate that cell fate can also be modulated shortly after mitosis (42). Moreover, widespread “multilineage priming,” in which a progenitor cell coexpresses genes that are specific to each of its molecularly distinct daughters, has been documented in *C. elegans* embryos (43).

However, inferring trajectories is not always possible, particularly when there are abrupt changes in the transcriptional program of differentiating cells. This can occur, for example, during the transition from a proliferating progenitor cell to a postmitotic neuron. Moreover, the multiplicity of cell types within any sample of neural tissue, which might have arisen from distinct lineages or a multiple-branching differentiation pathway (44), further confounds these methods.

Approaches that address some of these limitations are being developed. These include computational methods that predict gene expression changes from splicing information (45) and techniques that provide a temporal signature to facilitate the ordering of gene expression changes (46). In addition, high-throughput methods are being devised to allow lineage reconstruction. Conventional approaches to lineage reconstruction have relied on methods that indelibly mark a single cell and its progeny with a tracer that can later be visualized. The advent of single-cell profiling has facilitated the design of a new generation of techniques that rely on the introduction of molecular barcodes (47). Approaches include transposon-based libraries encoding a transcript harboring a random sequence that acts as a unique identifier, and systems based on the CRISPR-Cas9-mediated introduction of short insertions or deletions of variable length and position that act as heritable genetic changes. This latter strategy has revealed that although most individual embryonic progenitors generate multiple cell types in the zebrafish brain, these tend to remain in relatively restricted spatial domains (48).

In addition to experimentally introduced genetic barcodes, spontaneous somatic mutations provide information about lineage relationships between cells. Somatic mutations in mitochondrial DNA have been used as clonal markers to infer cellular relationships within the hematopoietic system (49, 50). Advances in single-cell genome sequencing have allowed the identification of naturally occurring somatic mutations in neurons, including new retrotransposition events, which occur at a rate of >1 mutation per cell division (51). These can also

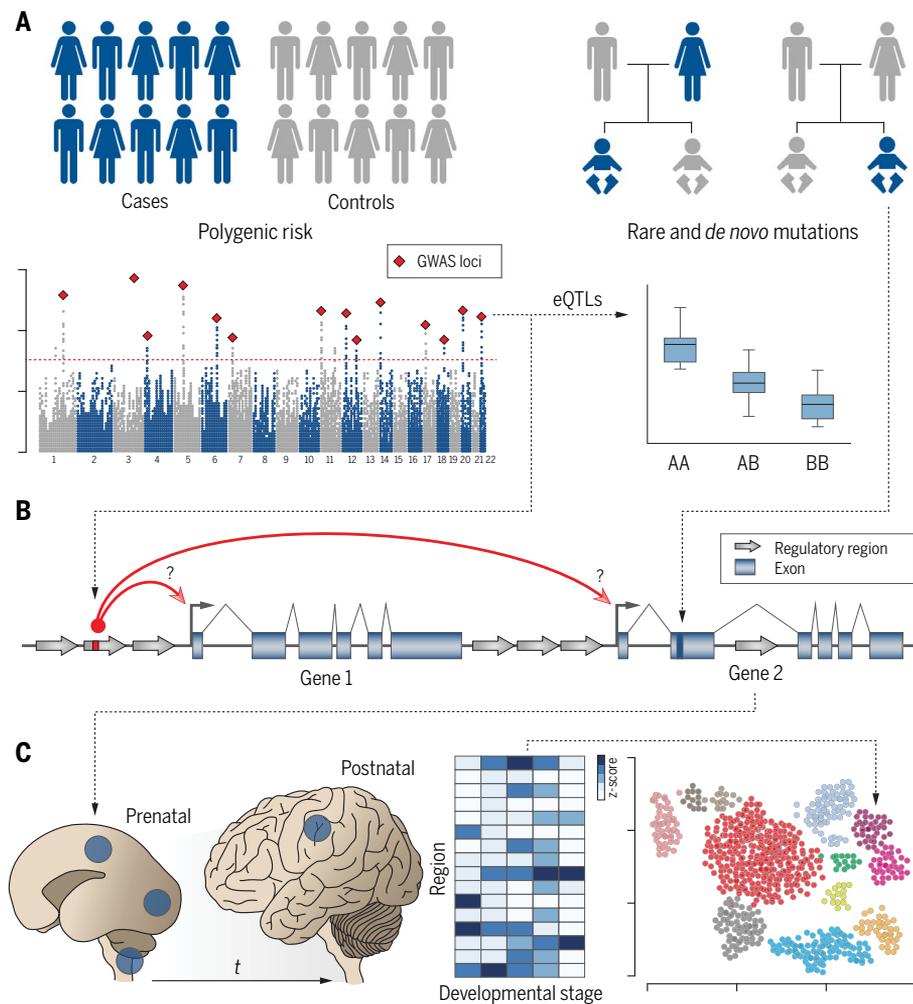


Fig. 3. Neurobiological interpretation of disease-associated gene variants. (A) Schematic representations of polygenic risk and rare and de novo coding mutations. The Manhattan plot indicates hypothetical loci (red rhomboids) reaching genome-wide significance (red dashed line). (B) Functional annotation of noncoding variants involves uncovering regulatory effects and identifying target gene(s) (question marks). Gene targets can also be inferred through computational methods such as eQTL. For coding variants, annotation is primarily based on impact to the amino acid sequence. (C) Single-cell genomic approaches have accelerated the identification of relevant cell types expressing disease-associated genes during brain development.

provide information about lineage relationships (52, 53). Exploiting endogenous mutations raises the possibility of inferring lineage relationships in the human brain, but this remains unproven, and the throughput, fidelity, and, most notably, cost of these methods will need to be improved if they are to become a practical reality.

It is apparent that in some cases, neurons of the same class can have distinct lineage histories. The transcriptomes of the six sets of IL-1 and IL-2 neurons in *C. elegans*, which arise from distinct lineages, gradually converge to the same molecular signature (43). In vertebrates, the somatic motor neurons that reside in the hindbrain have a distinct developmental history from those in the spinal cord but attain similar postmitotic identities (54). Likewise, oligodendrocytes through-

out the CNS transiently converge on a similar transcriptional identity despite spatially, temporally, and molecularly heterogeneous origins (55). Whether there are common features to the convergence of developmental trajectories in each of these cases needs to be investigated.

It is also unclear whether the genomic regulatory landscapes that produce the transcriptomes of equivalent neuronal classes originating from distinct lineages are similar. The application of techniques such as single-cell ATAC-seq (assay for transposase-accessible chromatin using sequencing) and CUT&Tag (cleavage under targets and tagmentation) that profile the regulatory genome of individual cells could address this question (56, 57). Identifying causal relationships between the regulatory genome and gene expression requires

overcoming the technical challenges of combining multiple assays on single cells (58, 59). Beyond insight into the mechanisms generating particular cell types, determining the developmental history of otherwise seemingly equivalent neuronal subtypes might explain why certain subsets of neurons are affected in particular disorders. Moreover, combining knowledge of cell lineage, developmental trajectory, and molecular mechanism has the potential to provide unprecedented insight into the central developmental neurobiology question of how neuronal diversity arises in the nervous system. The hope is that systematic analysis between different regions of the nervous system and different species might reveal rules that explain the underlying logic to the acquisition and maintenance of neuronal identity. This is of course only the first step in the assembly of functional neuronal circuits, and the role of stochastic and activity-dependent mechanisms will need to be investigated alongside the genetic programs specifying cell type identity.

The genetics of neurodevelopmental disorders

In parallel to the explosion of data from molecular approaches, the field of neuropsychiatric genetics has seen spectacular advances over the past decade. Most neuropsychiatric disorders arise from the alteration of normal developmental trajectories and have a prominent heritable component, irrespective of the age at which they are clinically diagnosed (60). The genetic architecture of neurodevelopmental disorders is diverse, ranging from an abundance of large-effect heterozygous mutations in autism spectrum disorder (ASD) (61) to highly polygenic in schizophrenia, involving the simultaneous contribution of multiple alleles with small effects (62). Many specific genetic associations are shared between multiple disorders with the same genetic variant responsible for different disorders in different individuals (63). In addition, the contribution of de novo somatic mutations to neurodevelopmental diseases is increasingly recognized (64).

Understanding the etiology of neuropsychiatric disorders has been difficult because in most cases there is no obvious neuropathology, and so the underlying biological mechanisms and the cells involved in the disease process are largely unknown (Fig. 3). In this context, combining genomics and transcriptomics may get us closer to identifying the cellular substrates of neuropsychiatric conditions such as ASD, in which large-effect, likely gene-disrupting coding mutations are relatively common. Bulk transcriptomics of the developing human brain revealed that ASD risk genes are abundantly expressed in certain brain regions (prefrontal and primary motor cortices, striatum, cerebellum, and medial dorsal nucleus of the thalamus) particularly during early and mid-fetal periods (65, 66). Now, single-cell resolution

transcriptome atlases of these brain regions are allowing the identification of the specific cell types in which candidate genes are expressed (67–69). Thus, the combination of advanced genomics and transcriptomics may get us closer to defining the cellular points of convergence in neurodevelopmental disorders such as ASD (70).

Identifying the relevant cell types or specific brain development events in neuropsychiatric conditions characterized by highly polygenic risk will be much more challenging. Genome-wide association studies (GWASs) have identified hundreds of common genetic variants—often in the form of single-nucleotide polymorphisms—that are statistically associated with increased risk in schizophrenia (62). Most variants do not directly identify the cause of the biological effect (Fig. 3). The linkage disequilibrium structure of the human genome means that any GWAS variant usually has many hundreds of nearby variants, requiring fine mapping to define the causal polymorphism. Even once identified, there is often difficulty associating the causal variant with a gene and a function, as sequence differences most frequently occur in noncoding regions. Most common variants are far away from the nearest known gene or are located in non-protein-coding regions of the genome (71). Further hampering this effort is that much of the noncoding genome is evolutionarily divergent between humans and other species. Although initial attempts are encouraging (72), confidently associating specific risk genes to specific genes and cellular populations in the developing brain is an enormous challenge. The unequivocal association of common risk variants to specific genes will require precise mapping of the interactions between regulatory elements and genes.

Considering the complexity of transcriptional regulation in vertebrates and the three-dimensional structure of the genome, big data approaches are necessary to identify causal disease genes from GWASs. For example, many disease-associated variants seem to be enriched in predicted transcriptional regulatory regions, known as cis-regulatory elements (CREs). Computational methods are beginning to use epigenomic data to prioritize candidate causal variants. Encouragingly, cell and tissue-specific epigenomic analyses (chromatin accessibility, transcription factor binding, histone marks) suggest that common variants for a particular disease are particularly enriched in CREs active in disease-relevant cell types (73). Moreover, assays that identify chromatin loops and the interaction between gene-coding regions and putative regulatory elements further aid the interpretation of GWAS-identified alleles (74).

Extending these studies to single-cell resolution, to identify regulatory element usage in defined cell types, will enhance the interpretation of regions identified by GWASs (Fig. 3).

In recent years, expression quantitative trait loci (eQTL) analyses based on the integration of genetic information with bulk RNA-seq data have identified downstream expression changes caused by disease-associated genes (75). Because transcriptional alterations are often cell type-specific, the opportunity now arises to use scRNA-seq to map eQTL across different cell types and developmental stages relevant for the disease process. Although such approaches are still in their early stages, proof-of-concept studies have demonstrated the feasibility of using scRNA-seq data for eQTL and gene regulatory network analyses (76). Notwithstanding the difficulties, it has been proposed that risk variants for schizophrenia are particularly abundant among genes that regulate the development and function of the synapse. The functional analysis of complement component 4 (C4), a protein that is abnormally increased in the brain of schizophrenia patients (77), represents one of the best examples so far linking common genetic variation to a neurodevelopmental mechanism in schizophrenia. C4 is encoded in humans by multiple structurally diverse alleles, and common variation increases schizophrenia risk by increasing the amounts of a specific C4 form. In mice, C4 regulates synapse elimination during development (77), which implies that this important developmental process—known as synaptic pruning—is abnormally enhanced in the brain of schizophrenia patients. It should be emphasized, however, that the effect size of common variation in humans is very small (i.e., C4 which has the largest of the small effect sizes, is just one of hundreds of gene variants that collectively contribute a modest proportion of the overall disease risk), so it is difficult to infer the actual role of defective synaptic pruning in the disease process in the absence of independent functional validation of other risk variants or rare mutations associated with schizophrenia.

One area where our understanding remains poor is the role of somatic mutations in neurodevelopmental disorders. Neurodevelopmental disorders had been assumed to be caused by inherited or de novo germline mutations. However, next-generation sequencing and single-cell sequencing technologies have revealed that the rate of somatic mutations is particularly high during neurogenesis (53) and may contribute to neurodevelopmental diseases. Indeed, whole-exome sequencing has found that somatic mutations contribute to ~5% of ASD risk in families in which only a single individual has ASD (78). Further analyses that relate genomic alterations to transcriptomic changes and any abnormalities in cell function or brain development will be necessary to understand the contribution of somatic mutations to neurodevelopmental disorders.

Together, these observations highlight how a convergence between developmental neuro-

biology and neurogenetics, driven by the increased resolution of the molecular and genetic assays, is likely to improve our understanding of both basic developmental mechanisms and the origin of neurological disease.

Modeling human brain development

The identification of genes and cells associated with neurodevelopmental disorders is only the beginning of a tortuous path to link gene variation with specific neurobiological functions. Experimental approaches are essential to identify the in vivo function of candidate genes. Unfortunately, whereas classical loss-of-function approaches might be useful to unravel the function of gene-disrupting coding mutations, knockouts are less likely to recapitulate the functionally subtle polymorphisms that are common in complex neurodevelopmental diseases, and animal models do not always recapitulate the social and cognitive deficits associated with human disorders. Nevertheless, large-scale phenotypic analyses of animal models carrying loss-of-function mutations in genes associated with common risk variants might be useful to identify functional convergence during brain development (79).

Cell culture-based assays represent a tractable and scalable approach for determining the function of particular genes and potential regulatory variants (Fig. 4). For example, human neurons derived from induced pluripotent stem cells (iPSCs) have been used to map gene regulatory element interactions in vitro (80). Protocols for generating specific neuronal subtypes (81) allow the landscape of epigenomic regulation to be explored in relatively well-characterized, specific types of neurons; the disadvantage is that it remains to be determined whether the right type of cells are being interrogated, leading to a circular problem. High-resolution molecular and genomic data from single-cell analyses provide a way to benchmark the similarity of in vitro-derived neurons to their in vivo counterparts (82, 83). More generally, neurons induced from patient-derived iPSCs or control iPSCs harboring engineered mutations can be used in cell biology and functional assays. For example, motor neurons derived from individuals with amyotrophic lateral sclerosis (ALS) caused by SOD1 mutations have been shown to display neurofilament aggregation and endoplasmic reticulum stress (84, 85), while glutamatergic cortical neurons differentiated from iPSCs derived from individuals with SHANK3 haploinsufficiency, associated with ASD, exhibited decreased glutamatergic neurotransmission that was corrected by reintroducing SHANK3 expression (86). Deriving neurons from iPSCs obtained from patients might currently be the only amenable way to investigate the functional consequences of highly polygenic risk traits (87). In addition, neurons derived from

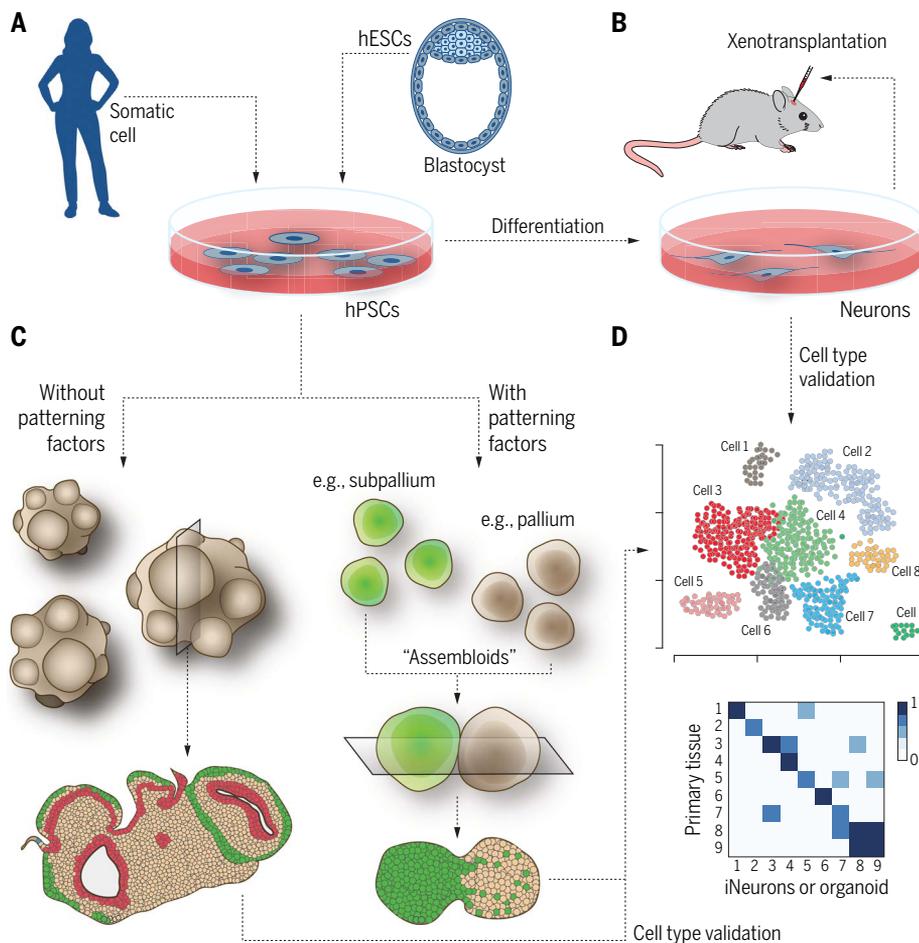


Fig. 4. Modeling the impact of disease-associated gene variants in human brain development.

(A) Human pluripotent stem cells (hPSCs) can be derived from somatic cells or embryonic stem cells (hESCs) and differentiated into distinct classes of neurons and glial cells. (B) The impact of disease-associated gene variants in neural development and function can be studied in vitro or through xenotransplantation in rodents. (C) hPSCs can also be used to derive brain organoids, with or without the use of patterning methods to direct the differentiation of specific classes of cells. Patterned brain organoids can be fused into “assembloids” to recreate regional features of brain development. (D) Single-cell genomic approaches can be used to benchmark the generation of relevant cell types in vitro by comparing transcriptional identities with those found in primary tissue.

human pluripotent stem cells integrate into functional circuits when transplanted into the mouse brain (88–90). Xenotransplantation is therefore a plausible methodology to study the formation and plasticity of neural circuits involving human neurons carrying specific mutations linked to neurodevelopmental disorders. For example, transplanted pyramidal projection neurons derived from individuals with Down syndrome exhibit substantial differences in synapse turnover in vivo compared with neurons derived from control iPSCs (91).

Brain organoids are also emerging as a powerful alternative to investigate human brain development in vitro (92). These three-dimensional neural tissues grown in culture from human stem cells have the potential to overcome some of the limitations of two-

dimensional culture systems, as they seem to recapitulate at least some of the anatomical and cytoarchitectural characteristics of specific brain regions in vitro (Fig. 4). The fusion of distinct brain region-specific organoids into more complex “assembloids” has been adopted as an approach that bypasses the difficulties of simultaneously patterning different brain regions in vitro. These allow the analysis of long-range neuronal migration across brain regions and the formation of major axonal tracts (93–97).

One issue that limits the use of brain organoids is that protocols for their generation have been prone to variable results. In this context, big data approaches are being used to validate organoids. Recent single-cell transcriptomic studies suggest that brain-region-specific organoids can consistently generate a diversity

of cell types (98, 99). Whether this is fully representative of primary tissue remains in contention, and it has been suggested that the activation of endoplasmic reticulum stress pathways might compromise cell type specification, at least in some types of organoids (100). In spite of these caveats, organoids allow the exploration of early events in human brain development. For instance, organoids derived from a patient with microcephaly, caused by a mutation in the cell cycle-related gene *CDK5RAP2*, have a premature depletion of the progenitor pool and are smaller than those derived from controls (101). Conversely, organoids engineered to carrying a mutation in the *PTEN* gene, which is found in patients with macrocephaly, exhibit a marked increase in proliferative cells, which results in abnormal overgrowth (102). These findings provide a foundation for the use of human brain organoids to investigate the mechanisms underlying polygenic and idiopathic developmental disorders. For example, transcriptional analysis of organoids derived from ASD patients with macrocephaly revealed consistent alterations in programs of gene expression associated with cell proliferation and differentiation of inhibitory neurons (103). These defects seem to be largely caused by the abnormal up-regulation in the expression of *FOXG1*, a gene crucial in forebrain patterning whose mutation causes an atypical form of Rett syndrome (104). Coordinated electrical activity has been observed in brain organoids (105–107), raising the prospect of modeling a range of neurological disorders and screening neuromodulatory drugs (107). Robust assays and rigorous controls will be crucial in experiments involving low-impact genes to overcome the inherent variability of organoids and differences between cell lines. The pleiotropy of many genetic variants implicated in neuropsychiatric disorders is an additional complication, but it might lead to unifying pathophysiological explanations by providing insight into the underlying cellular or molecular mechanisms.

Organoids have emerged as a powerful way to explore early brain development in humans, but their validation depends on our ability to systematically assess how similar they are to the primary tissue. Developing standardized protocols, quality controls, and analytical approaches will support this goal. At the moment, this comparison has been exclusively based on single-cell transcriptomic approaches, but we should ideally be able to integrate data across multiple levels of analysis (molecular, morphology, physiology) to better define the limitations of this approach for understanding how neural circuits are established in humans under normal circumstances and in disease.

Outlook

New technology and ever-larger and higher-resolution datasets are providing new perspectives

on long-standing questions about the ontogeny, composition, and function of the nervous system. Progress will depend not only on the acquisition and analysis of data but on their successful application to address fundamental conceptual questions such as what constitutes a cell type. Similarly, it is unlikely that our understanding of the mechanisms underlying neuropsychiatric conditions will increase by simply identifying additional risk genes. Consequently, methods to address the underlying biological mechanisms must be a priority.

The convergence of developmental neurobiology and neurogenetics is an exciting prospect. Integrating data from anatomical, developmental, genetic, and molecular studies has the potential to link cellular processes to functional and behavioral consequences. The analysis of developmental trajectories using big data approaches may enable the investigation of differences among multiple individuals and ultimately the complex biological mechanisms underlying individuality (9, 48, 108). This would provide fundamental insight and offer a new vision to the field.

REFERENCES AND NOTES

- Zheng *et al.*, A complete electron microscopy volume of the brain of adult *Drosophila melanogaster*. *Cell* **174**, 730–743.e22 (2018). doi: [10.1016/j.cell.2018.06.019](https://doi.org/10.1016/j.cell.2018.06.019); pmid: 30033368
- Motta *et al.*, Dense connectomic reconstruction in layer 4 of the somatosensory cortex. *Science* **366**, eaay3134 (2019). doi: [10.1126/science.aay3134](https://doi.org/10.1126/science.aay3134); pmid: 31649140
- Ohyama *et al.*, A multilevel multimodal circuit enhances action selection in *Drosophila*. *Nature* **520**, 633–639 (2015). doi: [10.1038/nature14297](https://doi.org/10.1038/nature14297); pmid: 25896325
- Kornfeld, W. Denk, Progress and remaining challenges in high-throughput volume electron microscopy. *Curr. Opin. Neurobiol.* **50**, 261–267 (2018). doi: [10.1016/j.conb.2018.04.030](https://doi.org/10.1016/j.conb.2018.04.030); pmid: 29758457
- Jun *et al.*, Fully integrated silicon probes for high-density recording of neural activity. *Nature* **551**, 232–236 (2017). doi: [10.1038/nature24636](https://doi.org/10.1038/nature24636); pmid: 29120427
- Knöpfel, C. Song, Optical voltage imaging in neurons: Moving from technology development to practical tool. *Nat. Rev. Neurosci.* **20**, 719–727 (2019). doi: [10.1038/s41583-019-0231-4](https://doi.org/10.1038/s41583-019-0231-4); pmid: 31705060
- Dana *et al.*, High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nat. Methods* **16**, 649–657 (2019). doi: [10.1038/s41592-019-0435-6](https://doi.org/10.1038/s41592-019-0435-6); pmid: 31209382
- Abdelfattah *et al.*, Bright and photostable chemigenetic indicators for extended in vivo voltage imaging. *Science* **365**, 699–704 (2019). doi: [10.1126/science.aav6416](https://doi.org/10.1126/science.aav6416); pmid: 31371562
- Naumann *et al.*, From whole-brain data to functional circuit models: The zebrafish optomotor response. *Cell* **167**, 947–960.e20 (2016). doi: [10.1016/j.cell.2016.10.019](https://doi.org/10.1016/j.cell.2016.10.019); pmid: 27814522
- Wan *et al.*, Single-cell reconstruction of emerging population activity in an entire developing circuit. *Cell* **179**, 355–372.e23 (2019). doi: [10.1016/j.cell.2019.08.039](https://doi.org/10.1016/j.cell.2019.08.039); pmid: 31564455
- Kicheva *et al.*, Coordination of progenitor specification and growth in mouse and chick spinal cord. *Science* **345**, 1254927 (2014). doi: [10.1126/science.1254927](https://doi.org/10.1126/science.1254927); pmid: 25258086
- Zeisel *et al.*, Molecular architecture of the mouse nervous system. *Cell* **174**, 999–1014.e22 (2018). doi: [10.1016/j.cell.2018.06.021](https://doi.org/10.1016/j.cell.2018.06.021); pmid: 30096314
- Rosenberg *et al.*, Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* **360**, 176–182 (2018). doi: [10.1126/science.aam8999](https://doi.org/10.1126/science.aam8999); pmid: 29545511
- Saunders *et al.*, Molecular diversity and specializations among the cells of the adult mouse brain. *Cell* **174**, 1015–1030.e16 (2018). doi: [10.1016/j.cell.2018.07.028](https://doi.org/10.1016/j.cell.2018.07.028); pmid: 30096299
- Tasic *et al.*, Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018). doi: [10.1038/s41586-018-0654-5](https://doi.org/10.1038/s41586-018-0654-5); pmid: 30382198
- Peng *et al.*, Molecular classification and comparative taxonomics of foveal and peripheral cells in primate retina. *Cell* **176**, 1222–1237.e22 (2019). doi: [10.1016/j.cell.2019.01.004](https://doi.org/10.1016/j.cell.2019.01.004); pmid: 30712875
- Konstantinides *et al.*, Phenotypic convergence: Distinct transcription factors regulate common terminal features. *Cell* **174**, 622–635.e13 (2018). doi: [10.1016/j.cell.2018.05.021](https://doi.org/10.1016/j.cell.2018.05.021); pmid: 29909983
- M. M. Sousa *et al.*, Molecular and cellular reorganization of neural circuits in the human lineage. *Science* **358**, 1027–1032 (2017). doi: [10.1126/science.aan3456](https://doi.org/10.1126/science.aan3456); pmid: 29170230
- Marques *et al.*, Oligodendrocyte heterogeneity in the mouse juvenile and adult central nervous system. *Science* **352**, 1326–1329 (2016). doi: [10.1126/science.aaf6463](https://doi.org/10.1126/science.aaf6463); pmid: 27284195
- Bayraktar *et al.*, Astrocyte layers in the mammalian cerebral cortex revealed by a single-cell in situ transcriptomic map. *Nat. Neurosci.* **23**, 500–509 (2020). doi: [10.1038/s41593-020-0602-1](https://doi.org/10.1038/s41593-020-0602-1); pmid: 32203496
- Chai *et al.*, Neural circuit-specialized astrocytes: Transcriptomic, proteomic, morphological, and functional evidence. *Neuron* **95**, 531–549.e9 (2017). doi: [10.1016/j.neuron.2017.06.029](https://doi.org/10.1016/j.neuron.2017.06.029); pmid: 28712653
- Masuda *et al.*, Spatial and temporal heterogeneity of mouse and human microglia at single-cell resolution. *Nature* **566**, 388–392 (2019). doi: [10.1038/s41586-019-0924-x](https://doi.org/10.1038/s41586-019-0924-x); pmid: 30760929
- Thion, F. Ginhoux, S. Garel, Microglia and early brain development: An intimate journey. *Science* **362**, 185–189 (2018). doi: [10.1126/science.aaf0474](https://doi.org/10.1126/science.aaf0474); pmid: 30309946
- Lubeck, A. F. Coskun, T. Zhiyentayev, M. Ahmad, L. Cai, Single-cell in situ RNA profiling by sequential hybridization. *Nat. Methods* **11**, 360–361 (2014). doi: [10.1038/nmeth.2892](https://doi.org/10.1038/nmeth.2892); pmid: 24681720
- Lee *et al.*, Highly multiplexed subcellular RNA sequencing in situ. *Science* **343**, 1360–1363 (2014). doi: [10.1126/science.1250212](https://doi.org/10.1126/science.1250212); pmid: 24578530
- Rodrigues *et al.*, Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 (2019). doi: [10.1126/science.aaw1219](https://doi.org/10.1126/science.aaw1219); pmid: 30923225
- Cadwell *et al.*, Electrophysiological, transcriptomic and morphological profiling of single neurons using Patch-seq. *Nat. Biotechnol.* **34**, 199–203 (2016). doi: [10.1038/nbt.3445](https://doi.org/10.1038/nbt.3445); pmid: 26689543
- Fuzik *et al.*, Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes. *Nat. Biotechnol.* **34**, 175–183 (2016). doi: [10.1038/nbt.3443](https://doi.org/10.1038/nbt.3443); pmid: 26689544
- Hanchate *et al.*, Connect-seq to superimpose molecular on anatomical neural circuit maps. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 4375–4384 (2020). doi: [10.1073/pnas.1912176117](https://doi.org/10.1073/pnas.1912176117); pmid: 32034095
- Han *et al.*, The logic of single-cell projections from visual cortex. *Nature* **556**, 51–56 (2018). doi: [10.1038/nature26159](https://doi.org/10.1038/nature26159); pmid: 29590093
- Arendt *et al.*, The origin and evolution of cell types. *Nat. Rev. Genet.* **17**, 744–757 (2016). doi: [10.1038/nrg.2016.127](https://doi.org/10.1038/nrg.2016.127); pmid: 27818507
- Mukamel, J. Ngai, Perspectives on defining cell types in the brain. *Curr. Opin. Neurobiol.* **56**, 61–68 (2019). doi: [10.1016/j.conb.2018.11.007](https://doi.org/10.1016/j.conb.2018.11.007); pmid: 30530112
- Davidson, D. H. Erwin, Gene regulatory networks and the evolution of animal body plans. *Science* **311**, 796–800 (2006). doi: [10.1126/science.1113832](https://doi.org/10.1126/science.1113832); pmid: 16469913
- Obert, L. Glenwinkel, J. White, Revisiting neuronal cell type classification in *Caenorhabditis elegans*. *Curr. Biol.* **26**, R1197–R1203 (2016). doi: [10.1016/j.cub.2016.10.027](https://doi.org/10.1016/j.cub.2016.10.027); pmid: 27875702
- Arber, Motor circuits in action: Specification, connectivity, and function. *Neuron* **74**, 975–989 (2012). doi: [10.1016/j.neuron.2012.05.011](https://doi.org/10.1016/j.neuron.2012.05.011); pmid: 22726829
- Li *et al.*, Classifying *Drosophila* olfactory projection neuron subtypes by single-cell RNA sequencing. *Cell* **171**, 1206–1220.e22 (2017). doi: [10.1016/j.cell.2017.10.019](https://doi.org/10.1016/j.cell.2017.10.019); pmid: 29149607
- Krebschull *et al.*, Cerebellar nuclei evolved by repeatedly duplicating a conserved cell type set. *bioRxiv* 20.06.25.170118 [Preprint]. 25 June 2020. doi: [10.1101/2020.06.25.170118](https://doi.org/10.1101/2020.06.25.170118)
- Delle *et al.*, Single cell transcriptomics reveals spatial and temporal dynamics of gene expression in the developing mouse spinal cord. *Development* **146**, dev173807 (2019). doi: [10.1242/dev.173807](https://doi.org/10.1242/dev.173807); pmid: 30846445
- Mi *et al.*, Early emergence of cortical interneuron diversity in the mouse embryo. *Science* **360**, 81–85 (2018). doi: [10.1126/science.aar6821](https://doi.org/10.1126/science.aar6821); pmid: 29472441
- Tritschler *et al.*, Concepts and limitations for learning developmental trajectories from single cell genomics. *Development* **146**, dev170506 (2019). doi: [10.1242/dev.170506](https://doi.org/10.1242/dev.170506); pmid: 31249007
- Polioudakis *et al.*, A single-cell transcriptomic atlas of human neocortical development during mid-gestation. *Neuron* **103**, 785–801.e8 (2019). doi: [10.1016/j.neuron.2019.06.011](https://doi.org/10.1016/j.neuron.2019.06.011); pmid: 31303374
- Iwata, P. Casimir, P. Vanderhaeghen, Mitochondrial dynamics in postmitotic cells regulate neurogenesis. *Science* **369**, 858–862 (2020). doi: [10.1126/science.aba9760](https://doi.org/10.1126/science.aba9760); pmid: 32792401
- S. Packer *et al.*, A lineage-resolved molecular atlas of *C. elegans* embryogenesis at single-cell resolution. *Science* **365**, eaax1971 (2019). doi: [10.1126/science.aax1971](https://doi.org/10.1126/science.aax1971); pmid: 31488706
- Mayer *et al.*, Developmental diversification of cortical inhibitory interneurons. *Nature* **555**, 457–462 (2018). doi: [10.1038/nature25999](https://doi.org/10.1038/nature25999); pmid: 29513653
- La Manno *et al.*, RNA velocity of single cells. *Nature* **560**, 494–498 (2018). doi: [10.1038/s41586-018-0414-6](https://doi.org/10.1038/s41586-018-0414-6); pmid: 30089906
- Telley *et al.*, Sequential transcriptional waves direct the differentiation of newborn neurons in the mouse neocortex. *Science* **351**, 1443–1446 (2016). doi: [10.1126/science.aad8361](https://doi.org/10.1126/science.aad8361); pmid: 26940868
- Woodworth, K. M. Girsakis, C. A. Walsh, Building a lineage from single cells: Genetic techniques for cell lineage tracking. *Nat. Rev. Genet.* **18**, 230–244 (2017). doi: [10.1038/nrg.2016.159](https://doi.org/10.1038/nrg.2016.159); pmid: 28111472
- Raj *et al.*, Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. *Nat. Biotechnol.* **36**, 442–450 (2018). doi: [10.1038/nbt.4103](https://doi.org/10.1038/nbt.4103); pmid: 29608178
- Ludwig *et al.*, Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. *Cell* **176**, 1325–1339.e22 (2019). doi: [10.1016/j.cell.2019.01.022](https://doi.org/10.1016/j.cell.2019.01.022); pmid: 30827679
- Xu *et al.*, Single-cell lineage tracing by endogenous mutations enriched in transposase accessible mitochondrial DNA. *eLife* **8**, e45105 (2019). doi: [10.7554/eLife.45105](https://doi.org/10.7554/eLife.45105); pmid: 30958261
- Ervony, E. Lee, P. J. Park, C. A. Walsh, Resolving rates of mutation in the brain using single-neuron genomics. *eLife* **5**, e12966 (2016). doi: [10.7554/eLife.12966](https://doi.org/10.7554/eLife.12966); pmid: 26901440
- Lozano, A. G. et al., Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science* **359**, 555–559 (2018). doi: [10.1126/science.aao4426](https://doi.org/10.1126/science.aao4426); pmid: 29217584
- Bae *et al.*, Different mutational rates and mechanisms in human cells at prenatation and neurogenesis. *Science* **359**, 550–555 (2018). doi: [10.1126/science.aan8690](https://doi.org/10.1126/science.aan8690); pmid: 29217587
- Metzlis *et al.*, Nervous system regionalization entails axial allocation before neural differentiation. *Cell* **175**, 1105–1118.e17 (2018). doi: [10.1016/j.cell.2018.09.040](https://doi.org/10.1016/j.cell.2018.09.040); pmid: 30343898
- Marques *et al.*, Transcriptional convergence of oligodendrocyte lineage progenitors during development. *Dev. Cell* **46**, 504–517.e7 (2018). doi: [10.1016/j.devcel.2018.07.005](https://doi.org/10.1016/j.devcel.2018.07.005); pmid: 30078729
- Cusanovich *et al.*, The cis-regulatory dynamics of embryonic development at single-cell resolution. *Nature* **555**, 538–542 (2018). doi: [10.1038/nature25981](https://doi.org/10.1038/nature25981); pmid: 29539636
- Kaya-Okur *et al.*, CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* **10**, 1930 (2019). doi: [10.1038/s41467-019-09982-5](https://doi.org/10.1038/s41467-019-09982-5); pmid: 31036827
- Cao *et al.*, Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science* **361**, 1380–1385 (2018). doi: [10.1126/science.aau0730](https://doi.org/10.1126/science.aau0730); pmid: 30166440
- Clark *et al.*, scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat. Commun.* **9**, 781 (2018). doi: [10.1038/s41467-018-03149-4](https://doi.org/10.1038/s41467-018-03149-4); pmid: 29472610
- Marín, Developmental timing and critical windows for the treatment of psychiatric disorders. *Nat. Med.* **22**, 1229–1238 (2016). doi: [10.1038/nm.4225](https://doi.org/10.1038/nm.4225); pmid: 27783067

61. L. M. Iakoucheva, A. R. Muotri, J. Sebat, Getting to the cores of autism. *Cell* **178**, 1287–1298 (2019). doi: [10.1016/j.cell.2019.07.037](https://doi.org/10.1016/j.cell.2019.07.037); pmid: [31491383](https://pubmed.ncbi.nlm.nih.gov/31491383/)
62. I. Giegling et al., Genetics of schizophrenia: A consensus paper of the WFSBP Task Force on Genetics. *World J. Biol. Psychiatry* **18**, 492–505 (2017). doi: [10.1080/15622975.2016.1268715](https://doi.org/10.1080/15622975.2016.1268715); pmid: [28112043](https://pubmed.ncbi.nlm.nih.gov/28112043/)
63. M. C. O'Donovan, M. J. Owen, The implications of the shared genetics of psychiatric disorders. *Nat. Med.* **22**, 1214–1219 (2016). doi: [10.1038/nm.4196](https://doi.org/10.1038/nm.4196); pmid: [27783064](https://pubmed.ncbi.nlm.nih.gov/27783064/)
64. A. M. D'Gama, C. A. Walsh, Somatic mosaicism and neurodevelopmental disease. *Nat. Neurosci.* **21**, 1504–1514 (2018). doi: [10.1038/s41593-018-0257-3](https://doi.org/10.1038/s41593-018-0257-3); pmid: [30349109](https://pubmed.ncbi.nlm.nih.gov/30349109/)
65. N. N. Parikshak et al., Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* **155**, 1008–1021 (2013). doi: [10.1016/j.cell.2013.10.031](https://doi.org/10.1016/j.cell.2013.10.031); pmid: [24267887](https://pubmed.ncbi.nlm.nih.gov/24267887/)
66. A. J. Willsey et al., Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell* **155**, 997–1007 (2013). doi: [10.1016/j.cell.2013.10.020](https://doi.org/10.1016/j.cell.2013.10.020); pmid: [24267886](https://pubmed.ncbi.nlm.nih.gov/24267886/)
67. M. Li et al., Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362**, 685–689 (2018). doi: [10.1126/science.aat7615](https://doi.org/10.1126/science.aat7615); pmid: [30545854](https://pubmed.ncbi.nlm.nih.gov/30545854/)
68. S. Zhong et al., A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature* **555**, 524–528 (2018). doi: [10.1038/nature25980](https://doi.org/10.1038/nature25980); pmid: [29539641](https://pubmed.ncbi.nlm.nih.gov/29539641/)
69. D. S. Lee et al., Simultaneous profiling of 3D genome structure and DNA methylation in single human cells. *Nat. Methods* **16**, 999–1006 (2019). doi: [10.1038/s41592-019-0547-z](https://doi.org/10.1038/s41592-019-0547-z); pmid: [31501549](https://pubmed.ncbi.nlm.nih.gov/31501549/)
70. D. Velmeshev et al., Single-cell genomics identifies cell type-specific molecular changes in autism. *Science* **364**, 685–689 (2019). doi: [10.1126/science.aav8130](https://doi.org/10.1126/science.aav8130); pmid: [31097668](https://pubmed.ncbi.nlm.nih.gov/31097668/)
71. M. T. Maurano et al., Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190–1195 (2012). doi: [10.1126/science.1222794](https://doi.org/10.1126/science.1222794); pmid: [22955828](https://pubmed.ncbi.nlm.nih.gov/22955828/)
72. N. G. Skene et al., Genetic identification of brain cell types underlying schizophrenia. *Nat. Genet.* **50**, 825–833 (2018). doi: [10.1038/s41588-018-0129-5](https://doi.org/10.1038/s41588-018-0129-5); pmid: [29785013](https://pubmed.ncbi.nlm.nih.gov/29785013/)
73. L. de la Torre-Ubieta et al., The dynamic landscape of open chromatin during human cortical neurogenesis. *Cell* **172**, 289–304.e18 (2018). doi: [10.1016/j.cell.2017.12.014](https://doi.org/10.1016/j.cell.2017.12.014); pmid: [29307494](https://pubmed.ncbi.nlm.nih.gov/29307494/)
74. A. Nott et al., Brain cell type-specific enhancer-promoter interactome maps and disease-risk association. *Science* **366**, 1134–1139 (2019). doi: [10.1126/science.aay0793](https://doi.org/10.1126/science.aay0793); pmid: [31727856](https://pubmed.ncbi.nlm.nih.gov/31727856/)
75. D. Wang et al., Comprehensive functional genomic resource and integrative model for the human brain. *Science* **362**, eaat8464 (2018). doi: [10.1126/science.aat8464](https://doi.org/10.1126/science.aat8464); pmid: [30545857](https://pubmed.ncbi.nlm.nih.gov/30545857/)
76. M. G. P. van der Wijst et al., Single-cell RNA sequencing identifies celltype-specific cis-eQTLs and co-expression QTLs. *Nat. Genet.* **50**, 493–497 (2018). doi: [10.1038/s41588-018-0089-9](https://doi.org/10.1038/s41588-018-0089-9); pmid: [29610479](https://pubmed.ncbi.nlm.nih.gov/29610479/)
77. A. Sekar et al., Schizophrenia risk from complex variation of complement component 4. *Nature* **530**, 177–183 (2016). doi: [10.1038/nature16549](https://doi.org/10.1038/nature16549); pmid: [26814963](https://pubmed.ncbi.nlm.nih.gov/26814963/)
78. D. Freed, J. Pevsner, The contribution of mosaic variants to autism spectrum disorder. *PLOS Genet.* **12**, e1006245 (2016). doi: [10.1371/journal.pgen.1006245](https://doi.org/10.1371/journal.pgen.1006245); pmid: [27632392](https://pubmed.ncbi.nlm.nih.gov/27632392/)
79. S. B. Thyme et al., Phenotypic landscape of schizophrenia-associated genes defines candidates and their shared functions. *Cell* **177**, 478–491.e20 (2019). doi: [10.1016/j.cell.2019.01.048](https://doi.org/10.1016/j.cell.2019.01.048); pmid: [30929901](https://pubmed.ncbi.nlm.nih.gov/30929901/)
80. M. Song et al., Mapping cis-regulatory chromatin contacts in neural cells links neuropsychiatric disorder risk variants to target genes. *Nat. Genet.* **51**, 1252–1262 (2019). doi: [10.1038/s41588-019-0472-1](https://doi.org/10.1038/s41588-019-0472-1); pmid: [31367015](https://pubmed.ncbi.nlm.nih.gov/31367015/)
81. Y. Tao, S. C. Zhang, Neural subtype specification from human pluripotent stem cells. *Cell Stem Cell* **19**, 573–586 (2016). doi: [10.1016/j.stem.2016.10.015](https://doi.org/10.1016/j.stem.2016.10.015); pmid: [27814479](https://pubmed.ncbi.nlm.nih.gov/27814479/)
82. B. Treutlein et al., Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature* **534**, 391–395 (2016). doi: [10.1038/nature18323](https://doi.org/10.1038/nature18323); pmid: [27281220](https://pubmed.ncbi.nlm.nih.gov/27281220/)
83. R. Tsunemoto et al., Diverse reprogramming codes for neuronal identity. *Nature* **557**, 375–380 (2018). doi: [10.1038/s41586-018-0103-5](https://doi.org/10.1038/s41586-018-0103-5); pmid: [29743677](https://pubmed.ncbi.nlm.nih.gov/29743677/)
84. E. Kiskinis et al., Pathways disrupted in human ALS motor neurons identified through genetic correction of mutant SOD1. *Cell Stem Cell* **14**, 781–795 (2014). doi: [10.1016/j.stem.2014.03.004](https://doi.org/10.1016/j.stem.2014.03.004); pmid: [24704492](https://pubmed.ncbi.nlm.nih.gov/24704492/)
85. H. Chen et al., Modeling ALS with iPSCs reveals that mutant SOD1 misregulates neurofilament balance in motor neurons. *Cell Stem Cell* **14**, 796–809 (2014). doi: [10.1016/j.stem.2014.02.004](https://doi.org/10.1016/j.stem.2014.02.004); pmid: [24704493](https://pubmed.ncbi.nlm.nih.gov/24704493/)
86. A. Shcheglovitov et al., SHANK3 and IGF1 restore synaptic deficits in neurons from 22q13 deletion syndrome patients. *Nature* **503**, 267–271 (2013). doi: [10.1038/nature12618](https://doi.org/10.1038/nature12618); pmid: [24132240](https://pubmed.ncbi.nlm.nih.gov/24132240/)
87. K. J. Brennand et al., Creating patient-specific neural cells for the in vitro study of brain disorders. *Stem Cell Reports* **5**, 933–945 (2015). doi: [10.1016/j.stemcr.2015.10.011](https://doi.org/10.1016/j.stemcr.2015.10.011); pmid: [26610635](https://pubmed.ncbi.nlm.nih.gov/26610635/)
88. I. Espuny-Camacho et al., Pyramidal neurons derived from human pluripotent stem cells integrate efficiently into mouse brain circuits in vivo. *Neuron* **77**, 440–456 (2013). doi: [10.1016/j.neuron.2012.12.011](https://doi.org/10.1016/j.neuron.2012.12.011); pmid: [23395372](https://pubmed.ncbi.nlm.nih.gov/23395372/)
89. A. M. Maroof et al., Directed differentiation and functional maturation of cortical interneurons from human embryonic stem cells. *Cell Stem Cell* **12**, 559–572 (2013). doi: [10.1016/j.stem.2013.04.008](https://doi.org/10.1016/j.stem.2013.04.008); pmid: [23642365](https://pubmed.ncbi.nlm.nih.gov/23642365/)
90. D. Linaro et al., Xenotransplanted human cortical neurons reveal species-specific development and functional integration into mouse visual circuits. *Neuron* **104**, 972–986.e6 (2019). doi: [10.1016/j.neuron.2019.10.002](https://doi.org/10.1016/j.neuron.2019.10.002); pmid: [31761708](https://pubmed.ncbi.nlm.nih.gov/31761708/)
91. R. Real et al., In vivo modeling of human neuron dynamics and Down syndrome. *Science* **362**, eaau1810 (2018). doi: [10.1126/science.aau1810](https://doi.org/10.1126/science.aau1810); pmid: [30309905](https://pubmed.ncbi.nlm.nih.gov/30309905/)
92. S. P. Pasca, The rise of three-dimensional human brain cultures. *Nature* **553**, 437–445 (2018). doi: [10.1038/nature25032](https://doi.org/10.1038/nature25032); pmid: [29364288](https://pubmed.ncbi.nlm.nih.gov/29364288/)
93. F. Birey et al., Assembly of functionally integrated human forebrain spheroids. *Nature* **545**, 54–59 (2017). doi: [10.1038/nature22330](https://doi.org/10.1038/nature22330); pmid: [28445465](https://pubmed.ncbi.nlm.nih.gov/28445465/)
94. J. A. Bagley, D. Reumann, S. Bian, J. Lévi-Strauss, J. A. Knoblich, Fused cerebral organoids model interactions between brain regions. *Nat. Methods* **14**, 743–751 (2017). doi: [10.1038/nmeth.4304](https://doi.org/10.1038/nmeth.4304); pmid: [28504681](https://pubmed.ncbi.nlm.nih.gov/28504681/)
95. Y. Xiang et al., hESC-derived thalamic organoids form reciprocal projections when fused with cortical organoids. *Cell Stem Cell* **24**, 487–497.e7 (2019). doi: [10.1016/j.stem.2018.12.015](https://doi.org/10.1016/j.stem.2018.12.015); pmid: [30799279](https://pubmed.ncbi.nlm.nih.gov/30799279/)
96. Y. Xiang et al., Fusion of regionally specified hPSC-derived organoids models human brain development and interneuron migration. *Cell Stem Cell* **21**, 383–398.e7 (2017). doi: [10.1016/j.stem.2017.07.007](https://doi.org/10.1016/j.stem.2017.07.007); pmid: [28757360](https://pubmed.ncbi.nlm.nih.gov/28757360/)
97. S. L. Giandomenico et al., Cerebral organoids at the air-liquid interface generate diverse nerve tracts with functional output. *Nat. Neurosci.* **22**, 669–679 (2019). doi: [10.1038/s41593-019-0350-2](https://doi.org/10.1038/s41593-019-0350-2); pmid: [30886407](https://pubmed.ncbi.nlm.nih.gov/30886407/)
98. S. J. Yoon et al., Reliability of human cortical organoid generation. *Nat. Methods* **16**, 75–78 (2019). doi: [10.1038/s41592-018-0255-0](https://doi.org/10.1038/s41592-018-0255-0); pmid: [30573846](https://pubmed.ncbi.nlm.nih.gov/30573846/)
99. S. Velasco et al., Individual brain organoids reproducibly form cell diversity of the human cerebral cortex. *Nature* **570**, 523–527 (2019). doi: [10.1038/s41586-019-1289-x](https://doi.org/10.1038/s41586-019-1289-x); pmid: [31680907](https://pubmed.ncbi.nlm.nih.gov/31680907/)
100. A. Bhaduri et al., Cell stress in cortical organoids impairs molecular subtype specification. *Nature* **578**, 142–148 (2020). doi: [10.1038/s41586-020-1962-0](https://doi.org/10.1038/s41586-020-1962-0); pmid: [31996853](https://pubmed.ncbi.nlm.nih.gov/31996853/)
101. M. A. Lancaster et al., Cerebral organoids model human brain development and microcephaly. *Nature* **501**, 373–379 (2013). doi: [10.1038/nature12517](https://doi.org/10.1038/nature12517); pmid: [23995685](https://pubmed.ncbi.nlm.nih.gov/23995685/)
102. Y. Li et al., Induction of expansion and folding in human cerebral organoids. *Cell Stem Cell* **20**, 385–396.e3 (2017). doi: [10.1016/j.stem.2016.11.017](https://doi.org/10.1016/j.stem.2016.11.017); pmid: [28041895](https://pubmed.ncbi.nlm.nih.gov/28041895/)
103. J. Mariani et al., FOXG1-dependent dysregulation of GABA/glutamate neuron differentiation in autism spectrum disorders. *Cell* **162**, 375–390 (2015). doi: [10.1016/j.cell.2015.06.034](https://doi.org/10.1016/j.cell.2015.06.034); pmid: [26186191](https://pubmed.ncbi.nlm.nih.gov/26186191/)
104. F. Ariani et al., FOXG1 is responsible for the congenital variant of Rett syndrome. *Am. J. Hum. Genet.* **83**, 89–93 (2008). doi: [10.1016/j.ajhg.2008.05.015](https://doi.org/10.1016/j.ajhg.2008.05.015); pmid: [18571142](https://pubmed.ncbi.nlm.nih.gov/18571142/)
105. G. Quadrato et al., Cell diversity and network dynamics in photosensitive human brain organoids. *Nature* **545**, 48–53 (2017). doi: [10.1038/nature22047](https://doi.org/10.1038/nature22047); pmid: [28445462](https://pubmed.ncbi.nlm.nih.gov/28445462/)
106. A. Trumpp, M. J. Depew, J. L. Rubenstein, J. M. Bishop, G. R. Martin, Cre-mediated gene inactivation demonstrates that FGFR3 is required for cell survival and patterning of the first branchial arch. *Genes Dev.* **13**, 3136–3148 (1999). doi: [10.1101/gad.13.23.3136](https://doi.org/10.1101/gad.13.23.3136); pmid: [10601039](https://pubmed.ncbi.nlm.nih.gov/10601039/)
107. R. A. Samarasinghe et al., Identification of neural oscillations and epileptiform changes in human brain organoids. *bioRxiv* 820183 [Preprint]. 28 October 2019. doi: [10.1101/820183](https://doi.org/10.1101/820183)
108. G. A. Linneweber et al., A neurodevelopmental origin of behavioral individuality in the *Drosophila* visual system. *Science* **367**, 1112–1119 (2020). doi: [10.1126/science.aaw7182](https://doi.org/10.1126/science.aaw7182); pmid: [32139539](https://pubmed.ncbi.nlm.nih.gov/32139539/)

ACKNOWLEDGMENTS

We thank B. Berninger, S.J. Butler, C. Houart, and A.F. Schier for feedback on an earlier version of this manuscript, and members of the Briscoe and Marin laboratories for stimulating discussions and ideas. We apologize to colleagues whose work is not cited in this review because of space limitations. **Funding:** Our research on this topic is supported by grants from the European Research Council (ERC-2017-AdG 787355) and Wellcome Trust (215556/Z/19/Z) to O.M.; by the Francis Crick Institute, which receives its core funding from Cancer Research UK, the UK Medical Research Council, and Wellcome Trust (all under FC001051); and by the European Research Council (ERC-2020-AdG 742138) to J.B. **Competing interests:** None declared.

10.1126/science.aaz8627

Looking at neurodevelopment through a big data lens

James Briscoe and Oscar Marin

Science **369** (6510), eaaz8627.
DOI: 10.1126/science.aaz8627

Insights from big data

The billions of neurons that make up the adult brain are organized into domains and circuits during development. High-resolution measurements such as those enabled by single-cell molecular profiling have revealed unexpected cellular diversity. Genomic tools are lending insight into mechanisms behind neurodevelopmental disorders. Briscoe and Marèn review the insights gained as big data analyses are applied to neurodevelopmental questions.

Science, this issue p. eaaz8627

ARTICLE TOOLS

<http://science.sciencemag.org/content/369/6510/eaaz8627>

RELATED CONTENT

<http://stm.sciencemag.org/content/scitransmed/10/452/eaar4338.full>
<http://stm.sciencemag.org/content/scitransmed/12/544/eaam8572.full>
<http://stm.sciencemag.org/content/scitransmed/10/461/eaam8434.full>
<http://stm.sciencemag.org/content/scitransmed/11/494/eaao0498.full>

REFERENCES

This article cites 106 articles, 28 of which you can access for free
<http://science.sciencemag.org/content/369/6510/eaaz8627#BIBL>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2020 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works