

OPEN ACCESS

EDITED BY

Dirk Feldmeyer,
Helmholtz Association of German
Research Centres (HZ), Germany

REVIEWED BY

Jack Waters,
Allen Institute for Brain Science,
United States
Brian Long,
Allen Institute, United States

*CORRESPONDENCE

Melissa A. Herman
✉ melissa.herman@charite.de

[†]These authors have contributed equally
to this work

RECEIVED 27 February 2026

REVISED 28 March 2026

ACCEPTED 31 March 2026

PUBLISHED 05 May 2026

CITATION

Yang S, Tyagi S, Rosenmund C and
Herman MA (2026) Identity crisis:
exploring the boundaries of cell type
identification in the age of single-cell
transcriptomics.
Front. Cell. Neurosci. 20:1819116.
doi: 10.3389/fncel.2026.1819116

COPYRIGHT

© 2026 Yang, Tyagi, Rosenmund and
Herman. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication
in this journal is cited, in accordance
with accepted academic practice. No
use, distribution or reproduction is
permitted which does not comply with
these terms.

Identity crisis: exploring the boundaries of cell type identification in the age of single-cell transcriptomics

Seulkee Yang[†], Sudeeksha Tyagi[†], Christian Rosenmund and
Melissa A. Herman*

Institute for Neurophysiology, Charité Universitätsmedizin—Berlin, Berlin, Germany

The rise of single-cell transcriptomics and comprehensive reference atlases promised a unifying molecular framework to classify cell identity. Yet transcriptomic identities are often interpreted outside the environmental contexts in which they arise. Here, we analyzed primary cortical cultures, which lack native tissue architecture, to compare their transcriptional profiles to multiple *in vivo* mouse cortical reference datasets. We found that while core molecular signatures for major neuronal subclasses are largely preserved *in vitro*, the loss of *in vivo* structure triggers high transcriptional divergence associated with metabolic and physiological state. We also identified clusters that consistently show low confidence in the classification tool. These ambiguous populations express incomplete canonical marker profiles resulting from a lack of structural cues necessary for full maturation. These observations suggest that while transcriptomic reference frameworks capture major aspects of neuronal identity, their interpretation can become less certain when cells are profiled outside their native environment. Our findings highlight the importance of considering environmental context when interpreting transcriptome-based cell type annotations and provide a resource for understanding how neuronal transcriptional programs are reshaped *in vitro*.

KEYWORDS

single-cell transcriptomics, cell identity classification, cortical neuronal identity, *in vitro* vs *in vivo*, primary neuronal cell culture

The scientific necessity of cell type classification

Cells are the fundamental unit of life. To better understand their roles at the organismal level, cell types are differentiated through structural and functional characteristics. In neuroscience, this cell type classification has historically relied on disentangling cell identity across various dimensions, such as morphological, anatomical, functional, and developmental lineage-based criteria (Zeng and Sanes, 2017).

Early cell type identification relied on morphological and anatomical classification of neurons. For instance, the illustrations of Ramón y Cajal revealed a substantial neuronal diversity across species and brain regions (Garcia-Lopez et al., 2010). Yet, electrophysiology and connectomics techniques were critical for providing functional dimensions to cell type classification (Cole and Curtis, 1939; Hodgkin and Huxley, 1939; Nassi et al., 2015; Kornfeld and Denk, 2018). Electrophysiology provides a framework for classifying neurons based on their biophysical attributes, such as membrane conductance, firing thresholds, and temporal firing patterns (Contreras, 2004), while connectomics expands this perspective by examining how neurons

integrate and interface at the circuit level. Mapping the brain's network architecture has marked a significant development in the field of neuroscience, with comprehensive connectomic datasets providing structural and functional insights into neural circuits (Sporns, 2011; Oh et al., 2014; Turner et al., 2022; Liu et al., 2025).

Spatial context and developmental lineage add another dimension to cell type classification. Cells occupying specific anatomical layers or regions often adopt specialized roles (Greig et al., 2013; Holguera and Desplan, 2018). In addition, lineage-tracing studies have shown that cells arising from distinct progenitor zones follow divergent differentiation paths, acquiring different structural and functional traits (Kretschmar and Watt, 2012; Ma et al., 2018). In this landscape, as our understanding of neuronal identity becomes increasingly multifaceted, it is important to consolidate these distinct properties of the cells into their identity. It has become evident that these different features each capture disparate facets of neuronal identity, without necessarily converging into a singular, stable definition of 'cell type'.

In view of these complexities, transcriptomics emerged as a promising strategy to construct a unifying molecular framework for cell type classification (Trapnell, 2015; Poulin et al., 2016; Yuste et al., 2020; Zhang et al., 2023). By unravelling the molecular code of cells that orchestrates different developmental, spatial, morphological, and functional aspects, transcriptomics offers a bottom-up outlook into cell identity. This was significantly expedited by the advent of single-cell transcriptomic profiling, first demonstrated by Tang et al. (2009), which presented a strategy for analyzing and classifying individual cells by applying their most basic defining properties.

The promise and limit of transcriptomic cell identity

Single-cell RNA sequencing methods yield a substantial and extensive dataset (Kharchenko, 2021). An inevitable repercussion of this scale is the curse of dimensionality and loss of interpretability (Imoto et al., 2022). To address this, statistical frameworks and dimensionality reduction methods emerged as essential tools to prune noisy data and extract a meaningful subset of biologically relevant information (Van der Maaten and Hinton, 2008; McInnes et al., 2018). Cell type annotation is performed on these reduced-dimensional embeddings, traditionally through the identification of marker genes (Pasquini et al., 2021). However, this manual approach is inherently subjective and labor-intensive. To address these limitations, a growing number of automated cell type annotation methods have emerged (Li et al., 2025). However, despite their efficiency, the reliability and interpretability of such automated annotations, especially in contexts where the reference may not capture the full diversity or novelty of the query data, remain areas of active investigation.

With the advancement and increase in the accessibility of single-cell transcriptomic techniques, a multitude of studies now aim at deciphering cell identity through transcriptomic profiles of neurons in different brain regions. In the primary visual cortex, for instance, from 2016 to 2018, the number of transcriptionally identified glutamatergic neuron subtypes rose from 19 to 33 (Tasic et al., 2016; Tasic et al., 2018). These studies highlight the fast-evolving nature of the field; however, they also allude to a certain

disarray in our current understanding of transcriptomic cell types. While common cell types have been identified across different regions, to what extent is the molecular identity of a neuron anchored to its functional and anatomical architecture? In fact, it has been observed that transcriptomic clusters can also fail to distinguish neurons with distinct projection patterns or, therefore, seemingly distinct roles (Mao and Staiger, 2024; Kim et al., 2020).

Transcriptomic profiling has become a widely applied framework for classifying cells and identifying novel cell types. However, single-cell transcriptomes capture both relatively stable cell-type-associated genes and context-dependent state genes reflecting metabolic activity, stress responses, or environmental adaptation. Changes in the balance between these programs may influence how confidently cells can be assigned to reference-defined cell types. In this article, we explore the nuances of transcriptomic cell type classification using single-nucleus RNA sequencing (snRNA-seq) data from dissociated primary cortical cultures. This *in vitro* system imparts a unique outlook into the transcriptomic profiles of cells stripped of anatomical architecture and spatial cues. When a cell is dissociated from its natural environment and grown in a dish, it experiences altered signaling landscapes, connectivity, and metabolic conditions. These changes raise an important question—how much of the transcriptomic identities defined *in vivo* remain detectable under such conditions? By comparing cultured cortical neurons to *in vivo* reference atlases across multiple developmental stages, we examine how well transcriptome-based cell-type annotations translate across environments. Thereby, we aim to assess which aspects of neuronal transcriptional identity remain robust *in vitro* and which features diverge under an altered environmental context.

Global molecular signatures are preserved in cultured cortical neurons

To examine the influence of spatial and network contexts on neuronal identity, we profiled cortical neurons cultured *in vitro* using snRNA-seq. The primary cultures were generated from cortices dissected from early postnatal days (P) 0–2 pups and grown *in vitro* for 14 days before being processed for snRNA-seq. Following quality control, we generated four separate *in vivo* reference models using publicly available snRNA-seq datasets from mouse visual cortex at P8, P14, P17, and P21 (Cheng et al., 2022). These references were selected due to their comparable sequencing resolution and span of developmental timepoints, which was imperative not only to understand cell-type composition but to establish the maturity and developmental background of cultured neurons. To evaluate the robustness of this annotation framework independently of our dataset, we first performed a self-validation using *in vivo* reference cells with known 'actual' labels. Label transfer showed high agreement with the original annotations, with a strong confusion matrix performance, high F1 scores, and true positive rates at the subclass level (Supplementary Figure S1A). Type-level annotation was also feasible, although performance was modestly reduced relative to subclass-level mapping (Supplementary Figure S1B).

Traditional clustering and annotation approaches can be limited by dimensionality reduction artifacts, parameter sensitivity, and a lack of standardization. To address these challenges and robustly map our cultured cells to *in vivo* cortical types, we used Popular Vote (popV), an automated cell-type annotation framework that integrates eight different automated annotation methods to assign consensus-based cell-type labels (Ergen et al., 2024; Figure 1A). We evaluated annotation performance at both subclass and finer type resolution. As type-level annotations consistently yielded lower consensus rates than subclass-level annotations, we used subclass labels for downstream comparison (Supplementary Figure S1C). To assess how well *in vitro* neurons mapped onto *in vivo* identities, we defined a cell-level agreement score as the number of annotation models supporting the majority-vote label. We then summarized annotation confidence at the cluster level by calculating the fraction of cells with support from five or fewer models in each Leiden cluster (Traag et al., 2019). Clusters with fewer than 30% of such cells were considered high-consensus. Across all references, a substantial fraction of cultured neurons achieved high-consensus labels. The highest percentage of confidently mapped cells was obtained with the P14 reference (80%), suggesting that cultured neurons most closely resemble the transcriptional states of the P14 reference dataset (Figure 1B).

To quantify the correspondence between *in vitro* and *in vivo* cell identities beyond annotation labels, we computed the Jaccard index based on the overlap of marker genes between matched cell types across all the developmental stages (Figure 1C). Overall, transcriptomic similarity between *in vitro* and *in vivo* neurons remained broadly comparable, with only modest shifts in Jaccard indices across time points. Excitatory subclasses displayed more variability and lower indices (0.07–0.28) except for L6b neuron (0.3–0.36), indicating reduced separation of laminar- and projection-specific programs *in vitro*. In contrast, inhibitory neurons such as parvalbumin (Pvalb), somatostatin (Sst), Lamp5, and vasoactive intestinal peptide (Vip) maintained higher Jaccard indices (0.22–0.4) with a trend toward stronger similarity at later postnatal stages. We also computed Spearman correlations between matched *in vitro* and *in vivo* subclasses. Global transcriptomic similarity was comparable across developmental stages (Supplementary Figure S1D). Consequently, we used P14 as the baseline reference for interpreting cell identity and divergence in subsequent analyses. Low-dimensional embeddings of the cultured dataset showed clear segregation between excitatory and inhibitory populations, with further separation into canonical laminar and subclass-associated identities, broadly mirroring the organization observed in *in vivo* reference datasets (Figure 1D). Through popV annotation, we established that all major neuronal subclasses identified in the *in vivo* references are retained in our *in vitro* dataset.

To characterize neuronal identity with a biophysical and functional perspective, we restricted the analysis to a set of 172 ion channel genes (Reva et al., 2025, GO:0005244). Despite representing a limited functional program, ion channel-based Leiden clustering overlaid on the transcriptome-wide subclass labels recovered a clear separation between inhibitory and excitatory populations and partially resolved major interneuron subclasses (Vip, Sst, Pvalb, Lamp5), indicating that ion channel expression captures partial information related to neuronal subclass-relevant structure (Figure 1E). To quantify how well ion channel-based clustering aligned with transcriptome-wide

identity assignments, we computed the Adjusted Rand Index (ARI) between Leiden clusters derived from ion channel genes only and popV-predicted cell-type labels. Notably, inhibitory (GABAergic) neurons exhibited a markedly higher ARI (0.77) compared to excitatory (glutamatergic) neurons, which showed substantially lower agreement (ARI = 0.12). This observation is consistent with the functional specialization of inhibitory neuronal subtypes, which are defined by distinct firing properties and therefore require tightly regulated and subtype-specific ion channel compositions.

Context-dependent divergence of transcriptomic profiles from *in vitro* primary neurons

While neurons cultured *in vitro* maintain core transcriptional features of neuronal identity, we wanted to examine the transcriptional divergence caused by the reshaping of environmental cues in culture. We quantified transcriptional dispersion per subclass by measuring the correlation distance of each cell to the mean expression profile of its assigned subclass within each biological sample. *In vitro* cells showed broader distance distributions than their *in vivo* counterparts, indicating increased transcriptional dispersion under culture conditions (Figure 2A). It should be noted that the *in vitro* dataset exhibits overall lower sequencing depth, fewer detected genes, and higher mitochondrial read fractions compared to the *in vivo* references (Supplementary Figure S2). While these differences are expected given the distinct nature of primary cultures versus intact tissue, they represent an additional source of variability that may influence reference-based annotation and the interpretation of transcriptional divergence. Comparison of subclass composition revealed substantial shifts between the *in vivo* and *in vitro* populations (Figure 2B). Deep-layer glutamatergic neurons were notably reduced in culture conditions, whereas upper-layer intratelencephalic (IT) pyramidal neurons were observed at higher numbers. In contrast, γ -aminobutyric acid (GABA)-ergic neurons are globally more abundant *in vitro*, with a particularly pronounced expansion of the Pvalb population. To characterize the molecular programs underlying these differences, we identified differentially expressed genes (DEGs) across all neuronal subclasses (Supplementary Figure S3). Gene Ontology analyses revealed a striking separation between pathways preferentially active in cultured versus *in vivo* neurons (Figure 2C). Neurons maintained *in vitro* showed strong enrichment for ribosomal biogenesis, oxidative phosphorylation, and biosynthetic pathways, indicative of increased translational demand and heightened metabolic activity. These signatures likely reflect transcriptional programs associated with cellular state in the artificial culture conditions, including altered nutrient availability, growth factor exposure, and the absence of physiological circuit constraints. In contrast, *in vivo* neurons prominently enriched gene sets associated with synaptic transmission, neurotransmitter release, calcium-dependent exocytosis, and trans-synaptic signaling. These pathways are commonly associated with synaptic communication in intact cortical circuits. Hence, while *in vitro* neurons maintain core

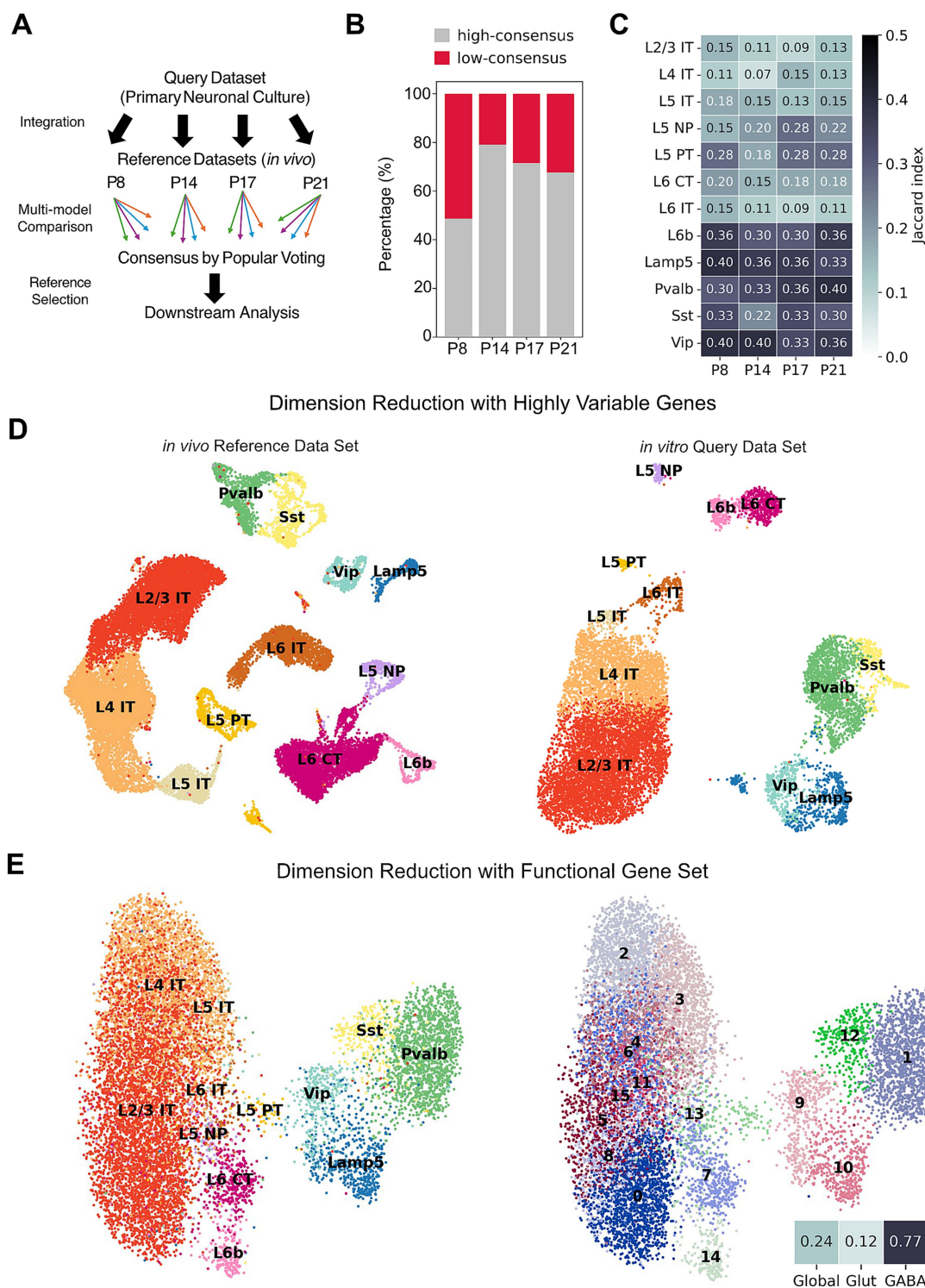
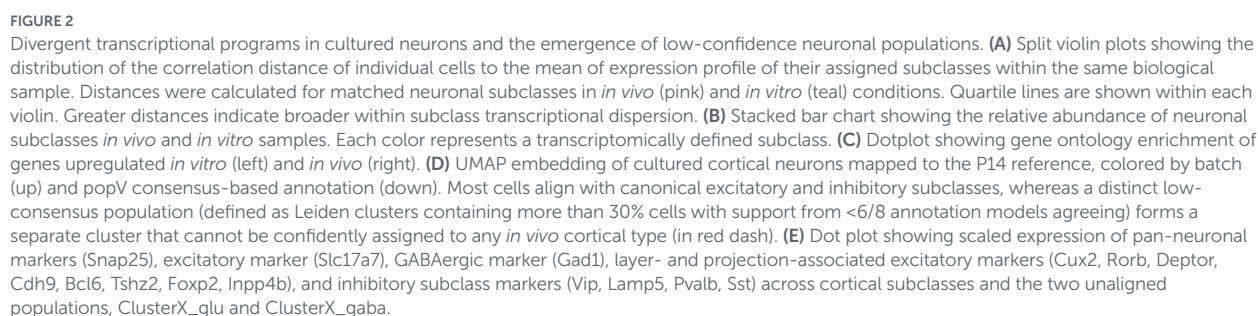


FIGURE 1
Global molecular signatures of cell identity are preserved in cultured cortical neurons. **(A)** Scheme of workflow to assign cell subclass annotations for a query dataset by using multiple reference sets using the PopV framework, which integrates predictions from eight automated annotation methods across multiple reference datasets. **(B)** Stacked bar plot showing the proportion of cells assigned to high-consensus (grey) or low-consensus (red) Leiden clusters when mapping *in vitro*-derived neurons to postnatal mouse cortical reference atlases (P8, P14, P17, and P21) at subclass-level. Cell-level agreement was defined by the number of annotation models supporting the majority-vote label. Clusters with fewer than 30% cells supported by five or fewer of eight models were classified as high-consensus. **(C)** Heatmap showing the Jaccard index of overlap between the top marker genes ($n = 30$ per subclass, identified by differential expression) for *in vitro*-derived neuronal subclasses and *in vivo* cortical references at postnatal stages P8, P14, P17, and P21. Higher Jaccard indices indicate greater similarity in subclass-specific transcriptional programs. Color scale ranges from 0 to 0.5. **(D)** UMAP visualization of the P14 cortical reference (left), and of the *in vitro* dataset alone (right). Only high-consensus cells were included. Cells are colored by annotated neuronal subclass. **(E)** UMAP embeddings of *in vitro* cortical neurons using a curated 172-gene ion channel panel (GO:0005244). Cells are colored by PopV subclass annotations derived from 4,000 highly variable genes (left) or by Leiden clusters (right). Agreement between ion channel-based clusters and PopV subclass labels was quantified using the ARI. At the whole-population level, concordance was modest (Global ARI = 0.24), with very low agreement among glutamatergic subclasses (Glut = 0.12) but high concordance within GABAergic subclasses (GABA = 0.77).



transcriptional signatures, their transcriptomes also show enrichment of context-specific gene programs.

Despite the overall transcriptional resemblance of cultured neurons to mid-postnatal cortical stages, our analysis revealed the emergence of a distinct cluster that did not align confidently with any *in vivo* reference population. This cluster consistently formed a distinct group in the low dimensional embedding and showed low consensus scores in popV-based annotations, indicating a lack of clear correspondence to known cortical *in vivo* subclasses (Figure 2D). We assigned two unaligned populations as ClusterX_glu and ClusterX_gaba, which occupy distinct spaces in between the canonical excitatory and inhibitory clusters (Figure 2D). To further investigate the nature of the ambiguous low-consensus cluster, we examined the expression of canonical excitatory and inhibitory marker genes across all subclasses, including ClusterX_glu and ClusterX_gaba (Figure 2E). These clusters showed reduced expression of canonical excitatory and inhibitory marker genes compared to their *in vivo*-aligned counterparts. One possible explanation is that these populations are dominated by state-associated transcriptional programs that reduce the relative prominence of canonical type-defining marker genes used by reference-based annotation methods. Consistent with this interpretation, *in vitro* neurons showed broader transcriptional dispersion within assigned subclasses than their *in vivo* counterparts, indicating an expansion of transcriptional state space under culture conditions. Under such conditions, neurons may retain broad excitatory or inhibitory identity while exhibiting transcriptional profiles that map poorly to existing *in vivo* reference subclasses.

Discussion

Since Robert Hooke observed the thin slice of cork under the microscope and discovered cells, biologists have sought to characterize and classify them. These efforts have been crucial, not only for organizing the complexity of biological systems but also for understanding the developmental and functional trajectories of specific cell types. This dates back to the early 1900s, when growing disputes and discussions about neuronal structure and function pushed the field to develop tools that could isolate and characterize single cells. This culminated in the prominent advancement in the field by Ross Harrison, who developed the first *in vitro* neuronal culture (Harrison, 1906). He noted that this approach offered the key advantage of allowing the study of individual cells independent of environmental and spatial influences. Now, in a time defined by vast datasets and high-resolution cellular profiling, our study took advantage of this reductionist approach to examine how much of neuronal transcriptional identity is preserved when cells are removed from their native environment.

Our findings show that cortical neurons cultured for 14 days from P0-P2 brains show transcriptomic profiles most closely aligning with P14 reference data. Additionally, the neurons maintain core cellular heterogeneity, including excitatory glutamatergic and inhibitory GABAergic subclasses. This suggests that key aspects of maturation can unfold *in vitro*, even without the *in vivo* spatial and complex network context, indicating that several

major transcriptional features of neuronal subclasses remain identifiable *in vitro*. The inhibitory neurons particularly display a stronger correspondence to *in vivo* transcriptional signatures, whereas transcriptional separation between excitatory subclasses weakens *in vitro*, with reduced marker gene coherence and increased cluster overlap. This suggests that excitatory identities are more sensitive to extrinsic environmental cue or network context that are absent in culture. Additionally, it may be that transcriptional differentiation between excitatory subtypes is more subtle, and more easily masked by *in vitro*-specific changes in gene expression. For instance, a recent multi-modal cell characterization of excitatory neurons from human cortex found that position in cortical layer plays a strong role in distinguishing subtype classes (Dalley et al., 2026). The differential stability across classes implies that use of transcriptomic clustering to define identity may be especially fragile for excitatory neurons whose phenotypes depend on network context, while inhibitory neuron types exhibit greater transcriptional autonomy.

The emergence of an extra cluster *in vitro*, disjointed from any established *in vivo* cortical identities, highlights the difficulty of mapping some *in vitro* transcriptional profiles onto existing *in vivo* reference identities. Using gene-set enrichment analyses, we demonstrated that the transcriptome mirrors the physiological properties dictated by the environment- metabolic and biosynthetic transcriptional programs enriched in *in vitro*, whereas the *in vivo* environment drives the expression of genes essential for synaptic cellular communication. Consistent with this, Baumann et al. (2025) recently reported clusters of transcriptionally undefined cells *in vitro* and demonstrated that this molecular divergence is driven by extrinsic cues. Notably, they also found that excitatory neurons were disproportionately affected by these context-dependent transcriptional changes. Together, these observations suggest that single-cell transcriptomes capture both cell-type-associated programs and context-dependent state-associated gene expression. Hence, transcriptomic marker expression should be interpreted in the context of cellular state and environment. Notably, reliable identification of cell-type-specific markers has been shown to require validation across multiple datasets, at least five, according to Fischer and Gillis (2021).

Waddington painted a landscape, an analogy for cell fate determination, of a ball rolling down a branching slope onto valleys, which represent stable differentiated cell states (Waddington, 2014). However, given the current understanding of cell identity, it is essential to look at these stable cell fates as high-dimensional attractor states in a gene regulatory landscape (Huang et al., 2005). This perspective implies that plastic and context-dependent attractor states may shift under altered environmental conditions. Although previously studied in developmental settings, our findings reveal that differentiated postnatal neurons *in vitro* can also display plastic transcriptional profiles. This underscores the need to evaluate whether observed transcriptomic differences correspond to altered cellular function or represent adaptive, context-dependent states.

A recent study elegantly demonstrated that a small subset of genes, particularly those encoding ion channels, is sufficient to capture the identity and variability of neuronal subtypes (Reva et al., 2025). Similarly, our results show that, while the *in vitro* system does not fully recapitulate global transcriptomic signatures and exhibits some culture-specific expression biases, the

key functional transcriptional features of neurons appear to be preserved. This concept aligns with studies integrating multi-modal datasets where morphoelectric characteristics of cortical GABAergic cells were integrated with their transcriptomic profile (Gouwens et al., 2020). The study found 28 cell types with the three integrated features in contrast to the previously described 60 ‘transcriptomic-types’. The study by the NIH’s BRAIN Initiative Cell Census Network (BICCN) also demonstrated how integrating morphology, electrophysiology, and spatial context can refine our understanding of neuronal diversity beyond transcriptomic labels alone (BICCN, 2021).

Transcriptomics provides a powerful lens to define a cell through its fundamental molecular signatures; yet, these signatures are inherently complex and transient. Our findings suggest that key molecular signatures are robust and can emerge even under simplified spatial and network contexts. Nevertheless, these observations underscore the importance of reframing neuronal identity, not solely as a fixed transcriptional entity defined by a molecular signature but as a dynamic, context-dependent state shaped by both intrinsic programs and extrinsic environments (Morris, 2019). As data grows denser and algorithms become more opaque, there is a growing need for unified and concerted analytical frameworks, particularly for cell type annotation. A multimodal approach not only reflects the full complexity of biological identity but also helps disentangle genuine biological transitions from artifacts introduced by analytic frameworks. Defining a cell type is not a matter of assigning it to a fixed transcriptional category but of understanding the regulatory, spatial, and functional conditions under which cell states emerge and stabilize.

Methods in brief

All animal experiments were performed in accordance with institutional and local authorities, as permitted by Landesamt für Gesundheit und Soziales (LaGeSo) Berlin under license numbers T0220/09.

Neurons were dissociated from P0-P2 male C56/BL6N mouse cortex and maintained for 14 days *in vitro* (DIV14), allowing us to observe transcriptional changes that emerge outside the native brain environment. Frozen cultures were processed to undergo single-nucleus encapsulation, barcoding, and sequencing library preparation using the Chromium GEM 3’ Kit v3.1 from 10X Genomics, and the sequencing was performed on a NovaSeq X Plus sequencer (Illumina).

Raw reads were processed using Cell Ranger v9.0.1 and aligned to the GRChm29 (2024-A) mouse reference to generate the gene expression count matrix. Ambient RNA was removed using CellBender (Fleming et al., 2023). The subsequent analysis was performed using SCANPY (Wolf et al., 2018). After removing doublets (Gayoso and Shor, 2022) and low-quality cells based on standard QC metrics (n_genes, n_counts, and mitochondrial content; Supplementary Figure S2), a total of 14,723 cells and 26,253 genes were retained for analysis.

Detailed Methods can be found in the [Supplementary Materials](#) section.

Data availability statement

Publicly available datasets were analyzed in this study. Raw and processed snRNA-seq data reported in this study are publicly available on the Zenodo repository at Doi: [10.5281/zenodo.19554693](https://doi.org/10.5281/zenodo.19554693).

Ethics statement

The animal study was approved by Landesamt für Gesundheit und Soziales (LaGeSo) Berlin. The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

SY: Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Writing – review & editing. ST: Conceptualization, Visualization, Writing – original draft, Writing – review & editing. CR: Resources, Supervision, Writing – review & editing. MH: Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy—EXC-2049-390688087 as a Research Fellowship to MH and Research Grant Project Number 543999006 to CR. ST was funded by a stipend from the Einstein Foundation Berlin.

Acknowledgments

snRNA-seq was performed by the Genomics Technology Platform of the Max Delbrück Center and the Berlin Institute of Health/Charité.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that Generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncel.2026.1819116/full#supplementary-material>

References

- Baumann, N., Morassut, I., Roig-Puigros, S., Klingler, E., Bartolini, G., Fièvre, S., et al. (2025). Cell-extrinsic controls over neocortical neuron fate and diversity. *Sci. Adv.* 11:eaw0218. doi: 10.1126/sciadv.adw0218
- BICCN. (2021). A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature* 598, 86–102. doi: 10.1038/s41586-021-03950-0
- Cheng, S., Butrus, S., Tan, L., Xu, R., Sagireddy, S., Trachtenberg, J. T., et al. (2022). Vision-dependent specification of cell types and function in the developing cortex. *Cell* 185:e24, 311–327. doi: 10.1016/j.cell.2021.12.022
- Cole, K. S., and Curtis, H. J. (1939). Electric impedance of the squid giant axon during activity. *J. Gen. Physiol.* 22, 649–670. doi: 10.1085/jgp.22.5.649
- Contreras, D. (2004). Electrophysiological classes of neocortical neurons. *Neural Netw.* 17, 633–646. doi: 10.1016/j.neunet.2004.04.003
- Dalley, R., Walling-Bell, S., Miller, J. A., Waleboer, F., Thijssen, J., McCutcheon, A., et al. (2026). Human neocortical glutamatergic neurons revealed through multimodal profiling. *bioRxiv*, 2026–2001. doi: 10.64898/2026.01.08.698257
- Ergen, C., Xing, G., Xu, C., Kim, M., Jayasuriya, M., McGeever, E., et al. (2024). Consensus prediction of cell type labels in single-cell data with popV. *Nat. Genet.* 56, 2731–2738. doi: 10.1038/s41588-024-01993-3
- Fischer, S., and Gillis, J. (2021). How many markers are needed to robustly determine a cell's type? *IScience* 24:103292. doi: 10.1016/j.isci.2021.103292
- Fleming, S. J., Chaffin, M. D., Arduini, A., Akkad, A. D., Banks, E., Marioni, J. C., et al. (2023). Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nat. Methods* 20, 1323–1335. doi: 10.1038/s41592-023-01943-7
- Garcia-Lopez, P., Garcia-Marin, V., and Freire, M. (2010). The histological slides and drawings of Cajal. *Front. Neuroanat.* 4:1156. doi: 10.3389/fnana.2010.009.2010
- Gayoso, A., and Shor, J. (2022). *JonathanShor/doubletdetection: doubletdetection v4.2*. Geneva: Zenodo. doi: 10.5281/zenodo.6349517
- Gouwens, N. W., Sorensen, S. A., Baftizadeh, F., Budzillo, A., Lee, B. R., Jarsky, T., et al. (2020). Integrated morphoelectric and transcriptomic classification of cortical GABAergic cells. *Cell* 183, 935–953.e19. doi: 10.1016/j.cell.2020.09.057
- Greig, L. C., Woodworth, M. B., Galazo, M. J., Padmanabhan, H., and Macklis, J. D. (2013). Molecular logic of neocortical projection neuron specification, development and diversity. *Nat. Rev. Neurosci.* 14, 755–769. doi: 10.1038/nrn3586
- Harrison, R. G. (1906). Observations on the living developing nerve fiber. *Exp. Biol. Med.* 4, 140–143. doi: 10.3181/00379727-4-98
- Hodgkin, A. L., and Huxley, A. F. (1939). Action potentials recorded from inside a nerve fibre. *Nature* 144, 710–711. doi: 10.1038/144710a0
- Holguera, I., and Desplan, C. (2018). Neuronal specification in space and time. *Science* 362, 176–180. doi: 10.1126/science.aas9435
- Huang, S., Eichler, G., Bar-Yam, Y., and Ingber, D. E. (2005). Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Phys. Rev. Lett.* 94:128701. doi: 10.1103/PhysRevLett.94.128701
- Imoto, Y., Nakamura, T., Escobar, E. G., Yoshiwaki, M., Kojima, Y., Yabuta, Y., et al. (2022). Resolution of the curse of dimensionality in single-cell RNA sequencing data analysis. *Life Sci. Allian.* 5:e202201591. doi: 10.26508/lsa.202201591
- Kharchenko, P. V. (2021). The triumphs and limitations of computational methods for scRNA-seq. *Nat. Methods* 18, 723–732. doi: 10.1038/s41592-021-01171-x
- Kim, E. J., Zhang, Z., Huang, L., Ito-Cole, T., Jacobs, M. W., Juvinett, A. L., et al. (2020). Extraction of distinct neuronal cell types from within a genetically continuous population. *Neuron* 107, 274–282.e6. doi: 10.1016/j.neuron.2020.04.018
- Kornfeld, J., and Denk, W. (2018). Progress and remaining challenges in high-throughput volume electron microscopy. *Curr. Opin. Neurobiol.* 50, 261–267. doi: 10.1016/j.conb.2018.04.030
- Kretzschmar, K., and Watt, F. M. (2012). Lineage tracing. *Cell* 148, 33–45. doi: 10.1016/j.cell.2012.01.002
- Li, T., Wang, Z., Liu, Y., He, S., Zou, Q., and Zhang, Y. (2025). An overview of computational methods in single-cell transcriptomic cell type annotation. *Brief. Bioinform.* 26:bba207. doi: 10.1093/bib/bba207
- Liu, L., Yun, Z., Manubens-Gil, L., Chen, H., Xiong, F., Dong, H., et al. (2025). Connectivity of single neurons classifies cell subtypes in mouse brains. *Nat. Methods* 22, 861–873. doi: 10.1038/s41592-025-02621-6
- Ma, J., Shen, Z., Yu, Y. C., and Shi, S. H. (2018). Neural lineage tracing in the mammalian brain. *Curr. Opin. Neurobiol.* 50, 7–16. doi: 10.1016/j.conb.2017.10.013
- Mao, X., and Staiger, J. F. (2024). Multimodal cortical neuronal cell type classification. *Pflugers Arch. Eur. J. Physiol.* 476, 721–733. doi: 10.1007/s00424-024-02923-2
- McInnes, L., Healy, J., and Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* 3:861. doi: 10.21105/joss.00861
- Morris, S. A. (2019). The evolving concept of cell identity in the single cell era. *Development* 146:dev169748. doi: 10.1242/dev.169748
- Nassi, J. J., Cepko, C. L., Born, R. T., and Beier, K. T. (2015). Neuroanatomy goes viral! *Front. Neuroanat.* 9:80. doi: 10.3389/fnana.2015.00080
- Oh, S. W., Harris, J. A., Ng, L., Winslow, B., Cain, N., Mihalas, S., et al. (2014). A mesoscale connectome of the mouse brain. *Nature* 508, 207–214. doi: 10.1038/nature13186
- Pasquini, G., Arias, J. E. R., Schäfer, P., and Busskamp, V. (2021). Automated methods for cell type annotation on scRNA-seq data. *Comput. Struct. Biotechnol. J.* 19, 961–969. doi: 10.1016/j.csbj.2021.01.015
- Poulin, J. F., Tasic, B., Hjerling-Lefler, J., Trimarchi, J. M., and Awatramani, R. (2016). Disentangling neural cell diversity using single-cell transcriptomics. *Nat. Neurosci.* 19, 1131–1141. doi: 10.1038/nn.4366
- Reva, M., Arnaudon, A., Zbili, M., Makkeh, A., Markram, H., Goaillard, J. M., et al. (2025). Higher-order interactions in neuronal function: from genes to ionic currents in biophysical models. *Proc. Natl. Acad. Sci.* 122:e2500048122. doi: 10.1073/pnas.2500048122
- Sporns, O. (2011). The human connectome: a complex network. *Ann. N. Y. Acad. Sci.* 1224, 109–125. doi: 10.1111/j.1749-6632.2010.05888.x
- Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* 6, 377–382. doi: 10.1038/nmeth.1315
- Tasic, B., Menon, V., Nguyen, T. N., Kim, T. K., Jarsky, T., Yao, Z., et al. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* 19, 335–346. doi: 10.1038/nn.4216
- Tasic, B., Yao, Z., Graybiel, L. T., Smith, K. A., Nguyen, T. N., Bertagnolli, D., et al. (2018). Shared and distinct transcriptomic cell types across neocortical areas. *Nature* 563, 72–78. doi: 10.1038/s41586-018-0654-5
- Traag, V. A., Waltman, L., and Van Eck, N. J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* 9:5233. doi: 10.1038/s41598-019-41695-z
- Trapnell, C. (2015). Defining cell types and states with single-cell genomics. *Genome Res.* 25, 1491–1498. doi: 10.1101/gr.190595.115
- Turner, N. L., Macrina, T., Bae, J. A., Yang, R., Wilson, A. M., Schneider-Mizell, C., et al. (2022). Reconstruction of neocortex: organelles, compartments, cells, circuits, and activity. *Cell* 185, 1082–1100.e24. doi: 10.1016/j.cell.2022.01.023
- Van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9, 2579–2605.
- Waddington, C. H. (2014). *The Strategy of the genes*. Abingdon: Routledge.

- Wolf, F. A., Angerer, P., and Theis, F. J. (2018). SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* 19:15. doi: 10.1186/s13059-017-1382-0
- Yuste, R., Hawrylycz, M., Aalling, N., Aguilar-Valles, A., Arendt, D., Armañanzas, R., et al. (2020). A community-based transcriptomics classification and nomenclature of neocortical cell types. *Nat. Neurosci.* 23, 1456–1468. doi: 10.1038/s41593-020-0685-8
- Zeng, H., and Sanes, J. R. (2017). Neuronal cell-type classification: challenges, opportunities and the path forward. *Nat. Rev. Neurosci.* 18, 530–546. doi: 10.1038/nrn.2017.85
- Zhang, S., Li, X., Lin, J., Lin, Q., and Wong, K. C. (2023). Review of single-cell RNA-seq data clustering for cell-type identification and characterization. *RNA* 29, 517–530. doi: 10.1261/rna.078965.121