Title: Single-cell genomics reveals region-specific developmental trajectories underlying neuronal diversity in the human hypothalamus

Authors: Brian R. Herb† (1), Hannah J. Glover† (2), Aparna Bhaduri (3), Carlo Colantuoni (1,5), Tracy L. Bale (6), Kimberly Siletti (7), Sten Linnarsson (7), Rebecca Hodge (8), Ed Lein (8), Arnold R. Kriegstein (9,10), Claudia A. Doege* (11), Seth A. Ament* (1,12)

Affiliations:

[1]Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD
[2]Naomi Berrie Diabetes Center, Columbia Stem Cell Initiative, Department of Pediatrics, Columbia University Irving Medical Center, New York, NY
[3]Department of Biological Chemistry, University of California—Los Angeles, Los Angeles, CA
[4]Medical Scientist Training Program, University of Maryland School of Medicine, Baltimore, MD
[5]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA
[6]Department of Pharmacology, University of Maryland School of Medicine, Baltimore, MD
[7]Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden
[8]Allen Institute for Brain Science, Seattle, WA 98109
[9]Department of Neurology, University of California—San Francisco, San Francisco, CA
[10]The Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, University of California, San Francisco, CA
[11]Naomi Berrie Diabetes Center, Columbia Stem Cell Initiative, Department of Pathology and Cell Biology, Columbia University Irving Medical Center, New York, NY
[12]Department of Psychiatry, University of Maryland School of Medicine, Baltimore, MD

*Corresponding author. Email: sament@som.umaryland.edu, cad2114@cumc.columbia.edu

†These authors contributed equally to this work

Abstract: The development and diversity of neuronal subtypes in the human hypothalamus has been insufficiently characterized. We sequenced the transcriptomes of 40,927 cells from the prenatal human hypothalamus spanning from 6 to 25 gestational weeks and 25,424 mature neurons in regions of the adult human hypothalamus, revealing a temporal trajectory from proliferative stem cell populations to mature neurons and glia. Developing hypothalamic neurons followed branching trajectories leading to 170 transcriptionally distinct neuronal subtypes in ten hypothalamic nuclei in the adult. The uniqueness of hypothalamic neuronal lineages was examined developmentally by comparing excitatory lineages present in cortex and inhibitory lineages in ganglionic eminence from the same individuals, revealing both distinct and shared drivers of neuronal maturation across the human forebrain. Cross-species comparisons to the mouse hypothalamus identified human-specific *POMC* populations expressing unique combinations of transcription factors and neuropeptides. These results provide the first comprehensive transcriptomic view of human hypothalamus development at cellular resolution.

One-Sentence Summary: Using single-cell genomics, we reconstructed the developmental lineages by which precursor populations give rise to 170 distinct neuronal subtypes in the human hypothalamus.

Main Text:

The hypothalamus is a small but anatomically complex brain region that controls a large variety of evolutionally requisite physiological and homeostatic functions, including body temperature, circadian rhythms, sleep, stress responses, satiety, and hunger, and aspects of mood, social behavior, and memory(*1–3*). These functions are subdivided amongst specialized neuronal subtypes, which are organized into distinct anatomical nuclei(*4–11*). Environmental and genetic perturbations to hypothalamic development result in long-lasting changes in physiology and behavior(*12–20*) and are thought to contribute to risk for human diseases including obesity, anxiety, and depression(*21*, *22*). These clinical consequences motivate deeper investigation into the timing and regulation of hypothalamic development. However, much of what we know about hypothalamic development is derived from animal models(*2–12*, *23*, *24*). While many of these functions are thought to be evolutionarily conserved, the molecular identities of human hypothalamic cells and the timing and regulation of their development remain inadequately characterized(*25*, *26*). Here, we sought to address this deficiency through single-cell transcriptomics of the prenatal and adult human hypothalamus to define its transcriptional cell types and developmental trajectories.

Results

An atlas of neuronal and non-neuronal lineages in the developing human hypothalamus

We performed 10x Genomics single-cell RNA sequencing (scRNA-seq) of prenatal hypothalamus from 11 human fetuses (4 female, 7 male) at ~6 to 25 gestational weeks (GW); Carnegie Stages (CS) 13-15, CS22, GW16, GW18-20, GW22, GW25), yielding 40,927 high-quality single-cell transcriptomes (**Fig. 1A and tables S1**). These samples were collected in parallel with samples from several additional brain regions in the same fetuses, and scRNA-seq of cortical samples from these fetuses has been reported previously(*27–30*). In addition, we performed single-nucleus RNA sequencing (snRNA-seq) of neurons from the post-mortem hypothalamus of a neurotypical donor (male, 50 years of age) (**Fig. 1A and fig. S1**), yielding 22,365 high-quality single-nucleus transcriptomes.

We integrated the 40,927 prenatal hypothalamic cells (**Fig. 1B and fig. S2**) and annotated the resulting cell clusters using a curated, comprehensive list of literature-based marker genes (**tables S2 and S3**). We identified 24 broad cell classes, including radial glia, dividing cells, neuronal, oligodendrocyte, and astrocyte lineages, as well as other supporting cell types (**Fig. 1, C to H and table S4**). Cell type distributions shifted over developmental time (**Fig. 1, B to D**). The earliest timepoints – CS13, 14 and 15 – were composed largely of dividing progenitors

(*MKI67*$^{+}$) and vascular leptomeningeal cells (*LUM*$^{+}$). Radial glia (*HOPX*$^{+}$) cells were identified at GW16 and onward. Neuronal and oligodendrocyte progenitor populations were present from the earliest time point. Post-mitotic neurons (*STMN2*$^{+}$) emerged as early as CS22 and were abundant starting at GW16. Oligodendrocyte populations (*PDGFRA*$^{+}$/*OLIG2*$^{+}$/*MKI67*$^{-}$) were detected by GW16 and began to reach maturation (*MAG*$^{+}$/*MOG*$^{+}$) by GW22. Astrocytes (*GFAP*$^{+}$/*AQP4*$^{+}$/*HOPX*$^{-}$), ependymocytes (*CCDC153*$^{+}$), and tanycytes (*CRYM*$^{+}$) were established by CS22.

Most neurons in the prenatal samples appeared immature, with weak expression of many canonical markers. To better annotate these populations and define their developmental trajectories, we integrated the 11,446 prenatal hypothalamic neurons (CS22 to GW25) with 22,365 adult hypothalamic neurons (**Fig. 2, A and B, and table S5**). Using Monocle3, we reconstructed a branching pseudotime trajectory, rooted at the earliest timepoint (CS22) (**Fig. 2, B and C**). We identified branch points where the lineages diverge, and ‘leaves’ where the lineages terminate and followed the trajectories of these branches and leaves to generate a lineage tree (**Fig. 2D**). The highest-order branch points distinguish GABAergic (*SLC32A1*$^{+}$) and glutaminergic (*SLC17A6*+) neurons (**Fig. 2E**). At the next level, branches could primarily be assigned to specific hypothalamic nuclei (**Fig. 2, F and G and tables S6 and S7**). Ten hypothalamic nuclei were identified, including the Arcuate (ARC; *TBX3*+), Tuberomammillary Terminal (TM; *HDC*+), Paraventricular Nucleus of the Hypothalamus (PVH; *SIM1*+), Ventromedial Nucleus of the Hypothalamus (VMH; *FEZF1*+), Lateral Hypothalamus (LH; *LHX9*+), Suprachiasmatic Nucleus (SCN; *LHX8*+/*SIX6*+), Supramammillary Nucleus (SMN; *LMX1A*+), Mammillary Nucleus (MN; *FOXB1*+), Zona Incerta (ZI; *MEIS2*+) and Intrahypothalamic Diagonal (ID; *LHX6*+). The transcription factor signature for these nuclei could be detected by CS22 (first trimester), except for the SMN where robust expression could be detected in the second trimester (**Fig. 2, G to J and fig. S3**).

The expression of canonical hypothalamic neuropeptides demarcated the maturation of each nucleus, revealing variation across cell types. In the Arcuate nucleus, *POMC* and *GHRH* were expressed at low levels in the first semester and increased to robust levels by the second trimester. Other Arcuate neuropeptides such as *AGRP*, *NPY* (co-expressed in AGRP neurons), and *CARTPT* (co-expressed in POMC neurons) were expressed only starting in the second trimester. *AVP*, a canonical neuropeptide of the PVH could be detected by CS22, while other PVH neuropeptides such as *OXT* and *CRH* were not expressed until the second trimester (**Figure 2G**).

We applied a hierarchical iterative clustering approach (scrattch.hicat package) to obtain a more detailed atlas of neuronal subtypes. This revealed 170 transcriptionally distinct cell clusters in adult neurons. 147 of which could be mapped to a specific nucleus (**Fig. 2, K and L)**, including all ten nuclei represented in our lineage tree and three additional nuclei that were more sparsely represented in the data **(fig. S4 and tables S3 and S8**). 74 clusters were glutaminergic, 67 were GABAergic and 6 were histaminergic (**table S9**). Most clusters could be assigned to established neuronal subtypes by the expression of neuropeptides and other canonical markers. For instance, we identified discrete neuronal populations expressing POMC, AGRP, OXT, GHRH, and AVP, each of which was restricted to just 1-2 clusters. Our data also support diverse subpopulations expressing certain neuropeptides, including CRH, and TRH, and SST, each of which was expressed in ten or more clusters. In summary, our data define the development and diversity of neuronal subtypes in the human hypothalamus in unprecedented detail.

Cross-species comparison of neuronal subtypes in humans and mice

Major classes of hypothalamic neurons are shared in humans and mice, but species differences in more refined neuronal subtypes are poorly understood. We integrated our sc/snRNA-seq data from human prenatal and adult hypothalamic neurons with existing scRNA-seq from developing(*11*) and adult mouse hypothalamus(*4–6*, *8*, *9*, *11*) to create a comprehensive atlas of hypothalamic neurons (*n* = 175,754, **Fig. 3A and table S10**). As expected, human and mouse neurons co-clustered within major groups corresponding to immature (prenatal) neurons, GABAergic neurons, glutamatergic neurons, and histaminergic neurons (**Fig. 3, B and C**). Louvain clustering revealed 28 neuronal sub-classes, most of which could be annotated to established subtypes and nuclei (**Fig. 3, D and E and fig. S5** and **table S11**). Most sub-classes contained both human and mouse cells, consistent with the evolutionary conservation of most hypothalamic neuronal sub-classes. While certain clusters of adult neurons appeared more abundant in one species than the other, we believe this most likely results from sampling differences in the human vs. mouse datasets rather than true species differences at this sub-class level. Thus, our analysis supports dozens of evolutionary conserved neuronal sub-classes and describes their marker genes.

To explore species differences in more refined sub-populations, we focused on POMC neurons, which were well represented in both the human and mouse datasets (**Fig. 3F**). Sub-clustering of POMC neurons revealed 11 sub-clusters (**Fig. 3G and tables S12 and S13**). Overall, cells from mouse and human samples were well mixed (**Fig. 3I**). However, cluster 11 (*POMC*/*PCSK1*/*LEPR*/*NPY1R*/*TBX3*+) was over-represented for human cells (79% human), while clusters 1, 4, 7, 8 and 10 were over-represented for mouse cells (95%, 98%, 82%, 100% and 100% respectively) (**Fig. 3, H to J**). We used co-expression patterns of transcription factors and neuropeptides within POMC neurons to gain insight into both conserved and species-specific regulatory networks (**tables S14 to S18**). This analysis predicted a core set of TFs as master regulators shared by mouse and human POMC neurons, including *ISL1*, *KLF6*, *SIX3*, *PRDM12*, *NHLH2* and *PROX1*. However, human POMC neurons also utilize *TBX3*, *JUND*, *CXXC4*, *YBX1*, *ZFHX4*, *MEIS3*, and *POU2F2*, whereas mouse POMC neurons utilize *PEG3*, *SIX3*, *PLAGL1* and *NFE2L1* (**Fig. 3K**). Thus, our data provide preliminary evidence for human-specific POMC neuron subtypes and their regulatory networks, which may merit further investigation.

Unique and shared features of hypothalamic vs. cortical germinal zones in the human forebrain

Little is known about the genes and trajectories distinguishing neurogenesis in the hypothalamus versus other forebrain regions. To address this, we obtained matched samples from the cortex, ganglionic eminence (the source of telencephalic inhibitory neurons), and hypothalamus from the same individuals at GW18, GW19 and GW20. Co-embedding these samples in a shared low-dimensional space enabled direct comparisons among the progenitor populations and excitatory and inhibitory neuron lineages (n=95,107, **Fig. 4 A and B and fig. S6 A and B**). Both neuronal and non-neuronal lineages emanate from a large cluster of radial glia that were abundantly represented in all three stem cell niches. Radial glia give rise to multiple neuronal and numerous non-neuronal progenitor populations. Of these, among the non-neuronal populations, we identified tanycytes and ependymal cells that were specific to the hypothalamus, as well as astrocytes and oligodendrocytes that were abundant in both the hypothalamus and cortex. A large population of dividing progenitor cells formed a distinct barrel shape in 3D UMAP space indicative of actively cycling cells (cell cycle stage, **fig. S6C**, seurat clusters 6, 9, 10, **fig. S6 D**). These dividing progenitors, along with the radial glia, feed into a neuronal lineage that split between excitatory and inhibitory lineages and further subdivided into major neuron sub-classes. While cortex and GE samples split mid-way along the developmental lineage between progenitors and mature neuron populations in a largely expected way, we took a keen interest in similarities and differences among these lineages and how progenitors from the hypothalamus fit into this framework.

The co-embedded neurogenic map retains well-known features of cortical neurogenesis. Excitatory neurons of the cortex branched off from inhibitory neurons about halfway along the bridge from early neuron progenitors into mature neuron populations. This lineage led directly into intratelencephalic (IT) excitatory neuron populations and reflected the temporal ordering with which deep layer vs. superficial layer IT neurons are born: newborn neurons at the beginning of the lineage proximal to progenitor cells primarily expressed markers for later-born superficial layer excitatory neurons such as *CUX2*, while more mature neurons at the lineage terminus primarily expressed markers of earlier-born deep layer 5/6 neurons such as *RORB*. A distinct branch in the trajectory expressed markers of extratelencephalic (ET) excitatory neurons. Separately, inhibitory neuron lineages led into three distinct inhibitory neuron clusters, which expressed markers of medial ganglionic eminence-derived (*LHX6*+), caudal ganglionic eminence-derived (*CALB2*$^{+}$), and lateral ganglionic eminence-derived (*FOXP1*$^{+}$) sub-populations. The hypothalamic germinal zone is distinct from germinal zones of the cortex and GE in that it gives rise to both excitatory and inhibitory neurons, and these locally born neurons represent most of the mature neurons within the hypothalamus. Most of the hypothalamic cells were distributed among excitatory and inhibitory lineages that also included cortical and GE-derived cells. This suggests that at a high level of organization the development of excitatory vs. inhibitory neurons in the hypothalamus involves shared regulatory programs with the development of excitatory vs. inhibitory lineages in the cortex and GE.

The high-level similarity of the developmental trajectories in the cortex, GE, and hypothalamus enabled direct comparisons of the lineages to reveal finer-scale distinctions. Region-specific gene expression differences were already apparent at early stages of development. In radial glia and actively dividing progenitor populations (clusters 4, 6, 9, 10), we found region-specific expression for developmental transcription factors such as *PAX6* and *POU3F2* in the cortex, *DLX1* and *DLX5* in the GE, and *FOXB1* and *SIX6* in hypothalamus (**Fig. 4D and tables S19 and S20**). We also identified shared regulators, including *HES1* and *EGR1* (**Fig. 4E**). At a later stage of development, in neural progenitors (cluster 11), we found hypothalamus-specific expression for the transcription factor *TSHZ2*, which remained elevated in mature neuron populations. GE-specific transcripts in neural progenitors included *COL1A2*, a collagen, along with elevated expression of the NOTCH signaling ligand *DLL1* and the cell-cell signaling mediator *LGALS1*. We confirmed cortex-specific expression for canonical progenitor marker genes such as *EOMES* and *NEUROD2*. Many other genes showed complex spatiotemporal patterns in neural progenitors. For instance, neural progenitors of the hypothalamus and cortex had high expression (compared to GE) for cytoskeletal proteins such as *MYO10*, *CALD1*, and *VIM*, whereas neural progenitors in the hypothalamus and GE (but not cortex) had shared expression of *RBP1*, involved in retinoic acid signaling(*31*). Other genes had different temporal patterns or different cell type specificity. For example, *PPP1R17*, a phosphatase regulatory subunit, was identified as a neuroprogenitor marker within the cortex(*30*). However, in the hypothalamus, *PPP1R17* is instead expressed most highly in radial glia. These results provide insight into the unique regulatory programs that give rise to specialized neuronal populations.

We applied a network-biology approach to gain further insight into the shared vs. unique drivers of neuronal development across lineages. Gene co-expression network modeling identified 28 gene co-expression modules, many of which were expressed in region- and stage-specific patterns across the development of excitatory and inhibitory neuronal lineages (**fig. S6E**). In addition, we reconstructed a gene regulatory network (GRN) model(*32*) to predict the target genes for 1317 TFs in each germinal zone and the key regulator TFs for each gene co-expression module. We found both region-specific and shared TF-gene associations, where the extent of overlap across models provides a measure for the re-wiring of these networks across brain regions. We illustrate this concept for the 20 most highly-connected 'hub' transcription factors (**Fig. 4F**). For these TFs, we detect many TF-gene edges that are reproducible within different samples from the same brain region but not across brain regions, suggesting substantial re-wiring. Interestingly, network edges detected in the hypothalamus appear to be split between cortical and GE-derived groups, suggesting an early lineage difference not distinguished in the UMAP in **Fig. 4A**. Gene network modeling also provided insights into the regulation of specific neuronal lineages. The gene co-expression module M73 (**fig. S6E**) was expressed across the entire excitatory neuron lineage from dividing progenitor cells to mature neurons, enabling us to compare the regulators of excitatory neuron development across regions. Several TFs were identified as key regulators of excitatory neuron development in both hypothalamus and cortex, including *NEUROD6*, *ANK2*, *MYT1L*, and *CSRNP3*. Strikingly, we also identified region-specific regulators of excitatory neurons, including hypothalamus-specific (e.g., *CSRNP3*, *ARID4A*, *PEG3*, *BASP1*) and cortex-specific TFs (*MEF2C*, *ZBTB18*, *NEUROD2*, *SATB2*). Similarly, module M70 was expressed across the entire inhibitory lineage, and we used this module to compare the regulators of inhibitory lineages in hypothalamus and GE (**fig. S6E**). We identified *SP9*, *DLX5*, *DLX2*, *SOX11*, and *SOX4* as shared regulators of inhibitory neuron development across regions, in addition to hypothalamus-specific (*SCAND1*, *ZNF428*, *YBX1*, and *PTMA*) and GE-specific TFs (*CITED2*, *TCF4*, *LHX6*, and *ARX*). This same pattern of shared and region-specific regulators could be identified in several additional modules, including M42, which is activated at later stages of excitatory neuron development (**fig. S6E**). These results suggest that there are shared "core" regulatory programs that govern the development of excitatory and inhibitory neurons in multiple brain regions, which may be refined by region-specific regulatory programs that give rise to the unique properties of each region's excitatory and inhibitory populations.

Discussion

Here, we have described the development and diversity of neuronal and non-neuronal cell types of the human hypothalamus across prenatal stages to adulthood at single-cell resolution. The window into prenatal development starting at ~6 GW and including 25 GW is the largest to date, and our data represent the first single-nucleus transcriptomic atlas for the adult human hypothalamus. Our comparisons of hypothalamic cells in humans vs. mice expands knowledge about both shared and human-specific aspects of hypothalamic development. Comparisons among matched prenatal brain samples provided a unique opportunity to evaluate specializations of neurodevelopmental trajectories in the hypothalamus vs. other forebrain regions.

Our analysis revealed a sequential specification of hypothalamic neurons. First, developing neurons attain a unique transcriptional identity corresponding to the nucleus in which they reside. Later, specific neuronal subtypes differentiate. This process was well underway by Carnegie Stage 22, just ~10 weeks of gestational age, at which time we were able to discern most nuclei by the expression of established markers. We confidently identified ten hypothalamic nuclei in our human prenatal data, and human cells from three more nuclei could be identified after including complementary data from human adult. Notably, certain nuclei appeared to mature at different rates. Distinct developmental dynamics among hypothalamic nuclei are consistent with previous reports using immunohistochemistry(*33*) as well as a recently published single-cell transcriptomics study of the human prenatal hypothalamus covering GW7 - 20(*34*).

Differentiation of neuronal subtypes occurred primarily in the second trimester samples, with evidence that the maturation of many hypothalamic neurons continues into the third trimester and beyond. Our results confirm previous observations that *OXT* is detectable by GW14(*35–37*) and *CRH* and *AVP* as early as GW12(*38*), although we observe *AVP* as early as GW10. However, even late in the second trimester, many neuronal subtypes were identifiable primarily by the expression of key transcription factors, while the expression of some neuropeptides remained sparse, such as *KISS1*. Such sparse expression of neuropeptides in the second trimester might be interpreted as still ongoing differentiation, diversification, and maturation of the various peptidergic subtypes.

We detected 170 transcriptionally distinct neuronal populations in the adult hypothalamus, 147 of which could be mapped to specific nuclei. These numbers are within the range of detected neuronal cell types by scRNA-seq at postnatal timepoints in mice(*4–11*), and more than three times the number of distinct populations described previously for the human hypothalamus(*34*). The increased resolution of our dataset can be attributed to adult timepoints with more mature neurons. However, future large-scale studies will be needed to better define some of the rare subtypes.

As noted above, variation in the prenatal environment can have lasting consequences on many hypothalamic functions(*12*, *13*, *22*), but it is not well understood why adverse prenatal environments lead to worse or different outcomes in some individuals and not others. Precisely delineating the timing at which hypothalamic nuclei mature could provide insight into these exposures and outcomes, including the possibility that differences in the developmental timing among hypothalamic nuclei could produce distinct sensitive periods within the developmental trajectory and different disease risks.

Mouse models have been used extensively to study the relationships between the prenatal environment, the development of the hypothalamus, and the emergence of behavioral and physiological variation in hypothalamic functions. Thus, a critical question is whether these developmental processes are strongly conserved in mice vs. humans. Overall, we found that patterns of cell type-specific gene expression were quite similar in mouse vs. human. Known markers from mice enabled us to assign nucleus and cell type identities for all nuclei and many neuronal subtypes(*11*, *39–46*). However, numerous human-specific gene expression patterns were identified, including human-specific utilization of certain transcription factors, which could provide a substrate for subtle changes in regulation and function. Therefore, as has been demonstrated in other brain regions, many cell type-specific gene expression programs for neuronal subtypes are conserved from mice to humans. For example, among the conserved genes are the transcription factors *ISL1* and *NHLH2* which have been implicated in the regulation of body weight via hypothalamic circuitry in mice and humans(*41*, *47–49*). Nonetheless, results in mice should be interpreted cautiously when specific genes show divergent patterns across species.

Our comparison of forebrain neurogenic niches revealed shared lineages leading to excitatory and inhibitory neurons in the hypothalamus, cortex, and GE. However, gene expression differences were detectable between regions starting at early stages of development in radial glia. Gene regulatory network models suggested that the differentiation of excitatory vs. inhibitory neurons occurs early in hypothalamic development and involves ‘core’ regulatory programs that are shared, respectively, by excitatory and inhibitory neuronal lineages in the cortex and GE. We present evidence that these core programs are refined by region-specific TFs that may give rise to the unique properties of the excitatory and inhibitory programs in each brain region. While some of these predicted regulators have been described previously, our models make many novel predictions for the roles of key regulator TFs, and these should be tested experimentally. As some of these networks are predicted to be human-specific, testing their functions will likely require the further development of human cell culture systems that enable direct comparisons across forebrain stem cell niches.

Several limitations should be noted, setting the stage for future work. As with most human studies, information on the third trimester is lacking. Additional scRNA-seq data from late gestation and early postnatal time points could improve the resolution to characterize the maturation of nuclei and cell types. In addition, there is a need for more spatially-resolved human transcriptomic data to more fully annotate neuronal subtypes to the nuclei in which they reside. Finally, there is limited information about the prenatal environment experienced by the fetuses that we studied. An exciting future direction will be to expand our analysis to human and model organism samples with known variation in the prenatal environment, enabling a more direct evaluation of sensitive periods and their developmental consequences. It is our expectation that continuing to characterize human-specific patterns of gene expression in the hypothalamus will enable researchers to more precisely evaluate genes and hypothalamic cell types in the pathophysiology of human diseases.

References and Notes

1. Y. Xie, R. I. Dorsky, Development of the hypothalamus: Conservation, modification and innovation. *Development (Cambridge)*. **144** (2017), pp. 1588–1599.
2. J. L. Bedont, E. A. Newman, S. Blackshaw, Patterning, specification, and differentiation in the developing hypothalamus. *Wiley Interdiscip Rev Dev Biol*. **4** (2015), pp. 445–468.
3. C. B. Saper, B. B. Lowell, The hypothalamus. *Current Biology*. **24** (2014), pp. R1111–R1116.
4. J. N. Campbell, E. Z. Macosko, H. Fenselau, T. H. Pers, A. Lyubetskaya, D. Tenen, M. Goldman, A. M. J. Verstegen, J. M. Resch, S. A. McCarroll, E. D. Rosen, B. B. Lowell, L. T. Tsai, A molecular census of arcuate hypothalamus and median eminence cell types. *Nat Neurosci*. **20**, 484–496 (2017).
5. R. Chen, X. Wu, L. Jiang, Y. Zhang, Single-Cell RNA-Seq Reveals Hypothalamic Cell Diversity. *Cell Rep*. **18**, 3227–3241 (2017).
6. J. R. Moffitt, D. Bambah-Mukku, S. W. Eichhorn, E. Vaughn, K. Shekhar, J. D. Perez, N. D. Rubinstein, J. Hao, A. Regev, C. Dulac, X. Zhuang, Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science (1979)*. **362** (2018), doi:10.1126/science.aau5324.
7. D. W. Kim, Z. Yao, L. T. Graybuck, T. K. Kim, T. N. Nguyen, K. A. Smith, O. Fong, L. Yi, N. Koulena, N. Pierson, S. Shah, L. Lo, A. H. Pool, Y. Oka, L. Pachter, L. Cai, B. Tasic, H. Zeng, D. J. Anderson, Multimodal Analysis of Cell Types in a Hypothalamic Node Controlling Social Behavior. *Cell*. **179**, 713-728.e17 (2019).
8. L. E. Mickelsen, M. Bolisetty, B. R. Chimileski, A. Fujita, E. J. Beltrami, J. T. Costanzo, J. R. Naparstek, P. Robson, A. C. Jackson, Single-cell transcriptomic analysis of the lateral hypothalamic area reveals molecularly distinct populations of inhibitory and excitatory neurons. *Nat Neurosci*. **22**, 642–656 (2019).
9. S. Wen, D. Ma, M. Zhao, L. Xie, Q. Wu, L. Gou, C. Zhu, Y. Fan, H. Wang, J. Yan, Spatiotemporal single-cell analysis of gene expression in the mouse suprachiasmatic nucleus. *Nat Neurosci*. **23**, 456–467 (2020).
10. R. A. Romanov, E. O. Tretiakov, M. E. Kastriti, M. Zupancic, M. Häring, S. Korchynska, K. Popadin, M. Benevento, P. Rebernik, F. Lallemend, K. Nishimori, F. Clotman, W. D. Andrews, J. G. Parnavelas, M. Farlik, C. Bock, I. Adameyko, T. Hökfelt, E. Keimpema, T. Harkany, Molecular design of hypothalamus development. *Nature*. **582**, 246–252 (2020).
11. D. W. Kim, P. W. Washington, Z. Q. Wang, S. H. Lin, C. Sun, B. T. Ismail, H. Wang, L. Jiang, S. Blackshaw, The cellular and molecular landscape of hypothalamic patterning and differentiation from embryonic to late postnatal development. *Nat Commun*. **11** (2020), doi:10.1038/s41467-020-18231-z.
12. G. Q. Chang, V. Gaysinskaya, O. Karatayev, S. F. Leibowitz, Maternal high-fat diet and fetal programming: Increased proliferation of hypothalamic peptide-producing neurons that increase risk for overeating and obesity. *Journal of Neuroscience*. **28**, 12107–12119 (2008).
13. K. Franke, T. Harder, L. Aerts, K. Melchior, S. Fahrenkrog, E. Rodekamp, T. Ziska, F. A. van Assche, J. W. Dudenhausen, A. Plagemann, "Programming" of orexigenic and anorexigenic hypothalamic neurons in offspring of treated and untreated diabetic mother rats. *Brain Res*. **1031**, 276–283 (2005).
14. M. Nishi, Effects of Early-Life Stress on the Brain and Behaviors: Implications of Early Maternal Separation in Rodents. *Int J Mol Sci*. **21** (2020), , doi:10.3390/ijms21197212.
15. A. Caqueret, C. Yang, S. Duplan, F. Boucher, J. L. Michaud, Looking for trouble: a search for developmental defects of the hypothalamus. *Horm Res*. **64**, 222–230 (2005).
16. S. L. Fyffe, J. L. Neul, R. C. Samaco, H. T. Chao, S. Ben-Shachar, P. Moretti, B. E. McGill, E. H. Goulding, E. Sullivan, L. H. Tecott, H. Y. Zoghbi, Deletion of Mecp2 in Sim1-expressing neurons reveals a critical role for MeCP2 in feeding behavior, aggression, and the response to stress. *Neuron*. **59**, 947–958 (2008).
17. B. E. McGill, S. F. Bundle, M. B. Yaylaoglu, J. P. Carson, C. Thaller, H. Y. Zoghbi, Enhanced anxiety and stress-induced corticosterone release are associated with increased Crh expression in a mouse model of Rett syndrome. *Proc Natl Acad Sci U S A*. **103**, 18267–18272 (2006).
18. A. F. M. van Abeelen, S. G. Elias, T. J. Roseboom, P. M. M. Bossuyt, Y. T. van der Schouw, D. E. Grobbee, C. S. P. M. Uiterwaal, Postnatal acute famine and risk of overweight: the dutch hungerwinter study. *Int J Pediatr*. **2012**, 1–9 (2012).
19. J. S. Huang, T. A. Lee, M. C. Lu, Prenatal programming of childhood overweight and obesity. *Matern Child Health J*. **11**, 461–473 (2007).

20. S. D. Parlee, O. A. MacDougald, Maternal nutrition and risk of obesity in offspring: the Trojan horse of developmental plasticity. *Biochim Biophys Acta*. **1842**, 495–506 (2014).
21. M. C. de Rosa, H. J. Glover, G. Stratigopoulos, C. A. LeDuc, Q. Su, Y. Shen, M. W. Sleeman, W. K. Chung, R. L. Leibel, J. Y. Altarejos, C. A. Doege, Gene expression atlas of energy balance brain regions. *JCI Insight*. **6** (2021), doi:10.1172/JCI.INSIGHT.149137.
22. F. Inoue, W. L. Eckalbar, Y. Wang, K. K. Murphy, N. Matharu, C. Vaisse, N. Ahituv, Genomic and epigenomic mapping of leptin-responsive neuronal populations involved in body weight regulation. *Nat Metab*. **1**, 475–484 (2019).
23. T. Shimogori, D. A. Lee, A. Miranda-Angulo, Y. Yang, H. Wang, L. Jiang, A. C. Yoshida, A. Kataoka, H. Mashiko, M. Avetisyan, L. Qi, J. Qian, S. Blackshaw, A genomic atlas of mouse hypothalamic development. *Nat Neurosci*. **13**, 767–775 (2010).
24. Y. H. Zhang, M. Xu, X. Shi, X. L. Sun, W. Mu, H. Wu, J. Wang, S. Li, P. Su, L. Gong, M. He, M. Yao, Q. F. Wu, Cascade diversification directs generation of neuronal diversity in the hypothalamus. *Cell Stem Cell*. **28**, 1483-1499.e8 (2021).
25. K. Tessmar-Raible, F. Raible, F. Christodoulou, K. Guy, M. Rembold, H. Hausen, D. Arendt, Conserved Sensory-Neurosecretory Cell Types in Annelid and Fish Forebrain: Insights into Hypothalamus Evolution. *Cell*. **129**, 1389–1400 (2007).
26. X. Zhou, S. Zhong, H. Peng, J. Liu, W. Ding, L. Sun, Q. Ma, Z. Liu, R. Chen, Q. Wu, X. Wang, Cellular and molecular properties of neural progenitors in the developing mammalian hypothalamus. *Nat Commun*. **11** (2020), doi:10.1038/s41467-020-17890-2.
27. T. J. Nowakowski, A. Bhaduri, A. A. Pollen, B. Alvarado, M. A. Mostajo-Radji, E. di Lullo, M. Haeussler, C. Sandoval-Espinosa, S. J. Liu, D. Velmeshev, J. R. Ounadjela, J. Shuga, X. Wang, D. A. Lim, J. A. West, A. A. Leyrat, W. J. Kent, A. R. Kriegstein, Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science (1979)*. **358** (2017), doi:10.1126/science.aap8809.
28. A. A. Pollen, A. Bhaduri, M. G. Andrews, T. J. Nowakowski, O. S. Meyerson, M. A. Mostajo-Radji, E. di Lullo, B. Alvarado, M. Bedolli, M. L. Dougherty, I. T. Fiddes, Z. N. Kronenberg, J. Shuga, A. A. Leyrat, J. A. West, M. Bershteyn, C. B. Lowe, B. J. Pavlovic, S. R. Salama, D. Haussler, E. E. Eichler, A. R. Kriegstein, Establishing Cerebral Organoids as Models of Human-Specific Brain Evolution. *Cell*. **176**, 743-756.e17 (2019).
29. A. Bhaduri, M. G. Andrews, W. Mancia Leon, D. Jung, D. Shin, D. Allen, D. Jung, G. Schmunk, M. Haeussler, J. Salma, A. A. Pollen, T. J. Nowakowski, A. R. Kriegstein, Cell stress in cortical organoids impairs molecular subtype specification. *Nature*. **578** (2020), doi:10.1038/s41586-020-1962-0.
30. U. C. Eze, A. Bhaduri, M. Haeussler, T. J. Nowakowski, A. R. Kriegstein, Single-cell atlas of early human brain development highlights heterogeneity of human neuroepithelial cells and early radial glia. *Nat Neurosci*. **24**, 584–594 (2021).
31. H. Flach, T. Basten, C. Schreiner, P. Dietmann, S. Greco, L. Nies, N. Roßmanith, S. Walter, M. Kühl, S. J. Kühl, Retinol binding protein 1 affects Xenopus anterior neural development via all-trans retinoic acid signaling. *Dev Dyn*. **250**, 1096–1112 (2021).
32. V. A. Huynh-Thu, A. Irrthum, L. Wehenkel, P. Geurts, Inferring Regulatory Networks from Expression Data Using Tree-Based Methods. *PLoS One*. **5** (2010), doi:10.1371/journal.pone.0012776.
33. Y. Koutcherov, J. K. Mai, K. W. S. Ashwell, G. Paxinos, Organization of human hypothalamus in fetal development. *Journal of Comparative Neurology*. **446**, 301–324 (2002).
34. X. Zhou, Y. Lu, F. Zhao, J. Dong, W. Ma, S. Zhong, M. Wang, B. Wang, Y. Zhao, Y. Shi, Q. Ma, T. Lu, J. Zhang, X. Wang, Q. Wu, Deciphering the spatial-temporal transcriptional landscape of human hypothalamus development. *Cell Stem Cell*. **29**, 328-343.e5 (2022).
35. D. F. Swaab, Development of the human hypothalamus. *Neurochem Res*. **20** (1995), doi:10.1007/BF01694533.
36. F. Schubert, J. M. George, M. Bhaskar Rao, Vasopressin and oxytocin content of human fetal brain at different stages of gestation. *Brain Res*. **213** (1981), doi:10.1016/0006-8993(81)91251-8.
37. G. Mengod, E. Goudsmit, A. Probst, J. M. Palacios, "Chapter 4 In situ hybridization histochemistry in the human hypothalamus" in (1992).
38. G. MASTORAKOS, I. ILIAS, Maternal and Fetal Hypothalamic-Pituitary-Adrenal Axes During Pregnancy and Postpartum. *Ann N Y Acad Sci*. **997** (2003), doi:10.1196/annals.1290.016.
39. D. W. Kim, K. Liu, Z. Q. Wang, Y. S. Zhang, A. Bathini, M. P. Brown, S. H. Lin, P. W. Washington, C. Sun, S. Lindtner, B. Lee, H. Wang, T. Shimogori, J. L. R. Rubenstein, S. Blackshaw, Gene regulatory networks controlling differentiation, survival, and diversification of hypothalamic Lhx6-expressing GABAergic neurons. *Commun Biol*. **4** (2021), doi:10.1038/s42003-020-01616-7.

40. D. P. Orquera, M. B. Tavella, F. S. J. de Souza, S. Nasif, M. J. Low, M. Rubinstein, The homeodomain transcription factor NKX2.1 is essential for the early specification of melanocortin neuron identity and activates Pomc expression in the developing hypothalamus. *Journal of Neuroscience*. **39**, 4023–4035 (2019).
41. S. Nasif, F. S. J. de Souza, L. E. González, M. Yamashita, D. P. Orquera, M. J. Low, M. Rubinstein, Islet 1 specifies the identity of hypothalamic melanocortin neurons and is critical for normal food intake and adiposity in adulthood. *Proc Natl Acad Sci U S A*. **112**, E1861–E1870 (2015).
42. F. S. J. de Souza, A. M. Santangelo, V. Bumaschny, M. E. Avale, J. L. Smart, M. J. Low, M. Rubinstein, Identification of Neuronal Enhancers of the Proopiomelanocortin Gene by Transgenic Mouse Analysis and Phylogenetic Footprinting. *Mol Cell Biol*. **25**, 3076–3086 (2005).
43. E. Sanz, A. Quintana, J. D. Deem, R. A. Steiner, R. D. Palmiter, G. S. McKnight, Fertility-regulating kiss1 neurons arise from hypothalamic pomc-expressing progenitors. *Journal of Neuroscience*. **35**, 5549–5556 (2015).
44. M. M. Li, J. C. Madara, J. S. Steger, M. J. Krashes, N. Balthasar, J. N. Campbell, J. M. Resch, N. J. Conley, A. S. Garfield, B. B. Lowell, The Paraventricular Hypothalamus Regulates Satiety and Prevents Obesity via Two Genetically Distinct Circuits. *Neuron*. **102**, 653-667.e6 (2019).
45. G. Alvarez-Bolado, F. A. Paul, S. Blaess, Sonic hedgehog lineage in the mouse hypothalamus: From progenitor domains to hypothalamic regions. *Neural Dev*. **7** (2012), doi:10.1186/1749-8104-7-4.
46. N. Morales-Delgado, B. Castro-Robles, J. L. Ferrán, M. Martinez-De-La-Torre, L. Puelles, C. Díaz, Regionalized differentiation of CRH, TRH, and GHRH peptidergic neurons in the mouse hypothalamus. *Brain Struct Funct*. **219**, 1083–1111 (2014).
47. P. Loid, T. Mustila, R. E. Mäkitie, H. Viljakainen, A. Kämpe, P. Tossavainen, M. Lipsanen-Nyman, M. Pekkinen, O. Mäkitie, Rare Variants in Genes Linked to Appetite Control and Hypothalamic Development in Early-Onset Severe Obesity. *Front Endocrinol (Lausanne)*. **11** (2020), doi:10.3389/FENDO.2020.00081.
48. A. K. Topaloglu, E. Simsek, M. A. Kocher, J. Mammadova, E. Bober, L. D. Kotan, I. Turan, C. Celiloglu, F. Gurbuz, B. Yuksel, D. J. Good, Inactivating NHLH2 variants cause idiopathic hypogonadotropic hypogonadism and obesity in humans. *Hum Genet*. **141**, 295–304 (2022).
49. N. AL-Rayyan, J. Zhang, A. S. Burnside, D. J. Good, Leptin signaling regulates hypothalamic expression of nescient helix-loop-helix 2 (Nhlh2) through signal transducer and activator 3 (Stat3). *Mol Cell Endocrinol*. **384**, 134–142 (2014).
50. R. D. Hodge, T. E. Bakken, J. A. Miller, K. A. Smith, E. R. Barkan, L. T. Graybuck, J. L. Close, B. Long, N. Johansen, O. Penn, Z. Yao, J. Eggermont, T. Höllt, B. P. Levi, S. I. Shehata, B. Aevermann, A. Beller, D. Bertagnolli, K. Brouner, T. Casper, C. Cobbs, R. Dalley, N. Dee, S. L. Ding, R. G. Ellenbogen, O. Fong, E. Garren, J. Goldy, R. P. Gwinn, D. Hirschstein, C. D. Keene, M. Keshk, A. L. Ko, K. Lathia, A. Mahfouz, Z. Maltzer, M. McGraw, T. N. Nguyen, J. Nyhus, J. G. Ojemann, A. Oldre, S. Parry, S. Reynolds, C. Rimorin, N. v. Shapovalova, S. Somasundaram, A. Szafer, E. R. Thomsen, M. Tieu, G. Quon, R. H. Scheuermann, R. Yuste, S. M. Sunkin, B. Lelieveldt, D. Feng, L. Ng, A. Bernard, M. Hawrylycz, J. W. Phillips, B. Tasic, H. Zeng, A. R. Jones, C. Koch, E. S. Lein, Conserved cell types with divergent features in human versus mouse cortex. *Nature*. **573**, 61–68 (2019).
51. T. E. Bakken, N. L. Jorstad, Q. Hu, B. B. Lake, W. Tian, B. E. Kalmbach, M. Crow, R. D. Hodge, F. M. Krienen, S. A. Sorensen, J. Eggermont, Z. Yao, B. D. Aevermann, A. I. Aldridge, A. Bartlett, D. Bertagnolli, T. Casper, R. G. Castanon, K. Crichton, T. L. Daigle, R. Dalley, N. Dee, N. Dembrow, D. Diep, S. L. Ding, W. Dong, R. Fang, S. Fischer, M. Goldman, J. Goldy, L. T. Graybuck, B. R. Herb, X. Hou, J. Kancherla, M. Kroll, K. Lathia, B. van Lew, Y. E. Li, C. S. Liu, H. Liu, J. D. Lucero, A. Mahurkar, D. McMillen, J. A. Miller, M. Moussa, J. R. Nery, P. R. Nicovich, S. Y. Niu, J. Orvis, J. K. Osteen, S. Owen, C. R. Palmer, T. Pham, N. Plongthongkum, O. Poirion, N. M. Reed, C. Rimorin, A. Rivkin, W. J. Romanow, A. E. Sedeño-Cortés, K. Siletti, S. Somasundaram, J. Sulc, M. Tieu, A. Torkelson, H. Tung, X. Wang, F. Xie, A. M. Yanny, R. Zhang, S. A. Ament, M. M. Behrens, H. C. Bravo, J. Chun, A. Dobin, J. Gillis, R. Hertzano, P. R. Hof, T. Höllt, G. D. Horwitz, C. D. Keene, P. v. Kharchenko, A. L. Ko, B. P. Lelieveldt, C. Luo, E. A. Mukamel, A. Pinto-Duarte, S. Preissl, A. Regev, B. Ren, R. H. Scheuermann, K. Smith, W. J. Spain, O. R. White, C. Koch, M. Hawrylycz, B. Tasic, E. Z. Macosko, S. A. McCarroll, J. T. Ting, H. Zeng, K. Zhang, G. Feng, J. R. Ecker, S. Linnarsson, E. S. Lein, Comparative cellular analysis of motor cortex in human, marmoset and mouse. *Nature*. **598**, 111–119 (2021).
52. Y. Hao, S. Hao, E. Andersen-Nissen, W. M. Mauck, S. Zheng, A. Butler, M. J. Lee, A. J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E. P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L. M. Fleming, B. Yeung, A. J. Rogers, J. M. McElrath, C. A. Blish, R. Gottardo, P. Smibert, R. Satija, Integrated analysis of multimodal single-cell data. *Cell*. **184** (2021), doi:10.1016/j.cell.2021.04.048.

53. W. Huang, A. Bhaduri, D. Velmeshev, S. Wang, L. Wang, C. A. Rottkamp, A. Alvarez-Buylla, D. H. Rowitch, A. R. Kriegstein, Origins and Proliferative States of Human Oligodendrocyte Precursor Cells. *Cell*. **182**, 594-608.e11 (2020).
54. R. C. Bandler, I. Vitali, R. N. Delgado, M. C. Ho, E. Dvoretskova, J. S. Ibarra Molinas, P. W. Frazel, M. Mohammadkhani, R. Machold, S. Maedler, S. A. Liddelow, T. J. Nowakowski, G. Fishell, C. Mayer, Single-cell delineation of lineage and genetic identity in the mouse brain. *Nature*. **601**, 404–409 (2022).
55. E. A. Newman, D. Wu, M. M. Taketo, J. Zhang, S. Blackshaw, Canonical Wnt signaling regulates patterning, differentiation and nucleogenesis in mouse hypothalamus and prethalamus. *Dev Biol*. **442**, 236–248 (2018).
56. A. Zeisel, H. Hochgerner, P. Lönnerberg, A. Johnsson, F. Memic, J. van der Zwan, M. Häring, E. Braun, L. E. Borm, G. la Manno, S. Codeluppi, A. Furlan, K. Lee, N. Skene, K. D. Harris, J. Hjerling-Leffler, E. Arenas, P. Ernfors, U. Marklund, S. Linnarsson, Molecular Architecture of the Mouse Nervous System. *Cell*. **174**, 999-1014.e22 (2018).
57. A. Bhaduri, C. Sandoval-Espinosa, M. Otero-Garcia, I. Oh, R. Yin, U. C. Eze, T. J. Nowakowski, A. R. Kriegstein, An atlas of cortical arealization identifies dynamic molecular signatures. *Nature*. **598**, 200–204 (2021).
58. J. Cao, M. Spielmann, X. Qiu, X. Huang, D. M. Ibrahim, A. J. Hill, F. Zhang, S. Mundlos, L. Christiansen, F. J. Steemers, C. Trapnell, J. Shendure, The single-cell transcriptional landscape of mammalian organogenesis. *Nature 2019 566:7745*. **566**, 496–502 (2019).
59. K. Benoit, K. Watanabe, H. Wang, P. Nulty, A. Obeng, S. Müller, A. Matsuo, quanteda: An R package for the quantitative analysis of textual data. *J Open Source Softw*. **3**, 774 (2018).
60. B. J. Molyneaux, L. A. Goff, A. C. Brettler, H. H. Chen, J. R. Brown, S. Hrvatin, J. L. Rinn, P. Arlotta, DeCoN: Genome-wide analysis of invivo transcriptional dynamics during pyramidal neuron fate selection in neocortex. *Neuron*. **85**, 275–288 (2015).
61. M. J. Lewis, M. R. Barnes, K. Blighe, K. Goldmann, S. Rana, J. A. Hackney, N. Ramamoorthi, C. R. John, D. S. Watson, S. K. Kummerfeld, R. Hands, S. Riahi, V. Rocher-Ros, F. Rivellese, F. Humby, S. Kelly, M. Bombardieri, N. Ng, M. DiCicco, D. van der Heijde, R. Landewé, A. van der Helm-van Mil, A. Cauli, I. B. McInnes, C. D. Buckley, E. Choy, P. C. Taylor, M. J. Townsend, C. Pitzalis, Molecular Portraits of Early Rheumatoid Arthritis Identify Clinical and Treatment Response Phenotypes. *Cell Rep*. **28**, 2455-2470.e5 (2019).
62. P. Langfelder, S. Horvath, WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*. **9**, 1–13 (2008).

Acknowledgments: We thank Owen White and the Neuroscience Multi-Omic Archive team and Ronna Hertzano and the gEAR / NeMO Analytics team, respectively, for their assistance in building and hosting the data resources and web browsers that accompany this study. Some primary human tissue was obtained from the Human Developmental Biology Resource (HDBR), with special thanks to S. Lisgo and M. Crosier.

Funding:

National Institutes of Health grant U01MH114825 (ARK)

National Institutes of Health grant U01MH114812 (EL)

National Institutes of Health grant R24MH114788 (Owen White, PI)

National Institutes of Health grant R24MH114815 (Owen White and Ronna Hertzano, PIs)

National Institutes of Health grant R01DC019370 (Ronna Hertzano, PI)

Author contributions:

Conceptualization: BRH, HJG, TLB, CAD, SAA

Methodology: BRH, SAA, CC

Investigation: BRH, HJG

Resources: AB, KS, SL, RH, EL, ARK

Supervision: CAD, SAA

Writing – original draft: BRH, HJG, CAD, SAA

Writing – review & editing: all authors

Competing interests: A.R.K. is a cofounder and board member of Neurona Therapeutics. The remaining authors declare no competing interests.

Data and materials availability:

The data analyzed in this study were produced through the Brain Initiative Cell Census Network (BICCN:RRID:SCR_015820) and deposited in the NEMO Archive (RRID:SCR_002001) under identifier nemo:dat-917e9vs accessible at https://assets.nemoarchive.org/dat-917e9vs.

Data can be explored on NeMO Analytics: https://nemoanalytics.org/p?l=a856c14e&g=gad2

Code availability:

https//github.com/brianherb/HumanHypothalamusDev

Supplementary Materials

Materials and Methods

Figs. S1 to S6

Tables S1 to S24

Figure 1. Neuronal and non-neuronal lineages across the developing human hypothalamus. A. Overview of sample collection, including single-cell RNAseq analysis of Trimester 1 and 2 prenatal hypothalamus, plus single-nucleus RNAseq from adult hypothalamic neurons dissected into four regions. B. 3D UMAP showing integrated samples of the prenatal hypothalamic samples including Trimester 1 (CS13, CS14, CS15 and CS22) and Trimester 2 (GW15, GW18, GW19, GW20, GW22 and GW25). C. 3D UMAP showing samples after clustering to show cell subpopulations. D. Stacked plot showing distribution of cell subpopulation across samples. E. 3D visualization of the neuronal lineage. F. 3D visualization of the astro-ependymal lineage. G. 3D visualization of the oligodendrocyte lineage. H. Dot Plot showing expression of marker genes.

Figure 2. Neuronal trajectories across development and the adult human hypothalamus. A. scRNAseq data from embryonic postmitotic neurons (CS22-GW25) and neurons from the 50YO adult hypothalamus. B. 3D UMAP showing overlay of samples by timepoint. C. Monocle3 was used to run pseudotime analysis of both prenatal and adult neurons starting with the node

containing the highest percentage of cells from CS22. D. The pseudotime trajectory highlighting branch points (green) and the terminal leaves (orange) E. The pseudotime trajectory can be used to generate a 2D lineage tree. The lineage tree can be overlaid with gene expression (red). GABAergic (*GAD1*), glutaminergic (*SLC17A6*) and histaminergic (*HDC*) expression demarcates cell populations. F. Nuclei were annotated using the marker genes in fig. S4 and table S3. G. Dot Plot showing expression of marker genes and neuropeptides split between Trimester 1 and 2 and adult. H. The lineage tree was subdivided into 24 nodes to check sample distribution across tree. I. The 24 nodes were plotted to show the temporal distribution of nodes by time point. J. Temporal expression was also plotted to show distribution across trimester 1, 2 and adult samples. K. Hicat was used to generate 170 clusters on the adult neurons. 147/170 clusters generated by Hicat could be assigned to nuclei. The number of individual nuclei are shown in parentheses, with nuclei visualized in shades within the Monocle3 lineage. L. Heatmap showing mean $\log_2$ fold change in expression of key marker genes for each individual subpopulation.

Figure 3. Cross-species comparison of mouse and human hypothalamic neurons. A. UMAP of single-cell RNAseq data from the human and mouse hypothalamus neurons, including prenatal and adult samples. Mammill, mammillary dissection; PO, preoptic dissection; SO, Supraoptic dissection; Tub, Tuberal dissection, ARC, Arcuate; SCN, Suprachiasmatic Nucleus; VMH, Ventromedial Nucleus of the Hypothalamus; LH, Lateral Hypothalamus. B. Split UMAP showing cell distribution across human/mouse and prenatal/adult samples. C. FeaturePlots showing localization of GABAergic (*SLC32A1* and *GAD1*), Glutaminergic (*SLC17A6*) and histaminergic (*HDC*) neurons. D. Louvain clustering generates 28 cell clusters. Anatomical nuclei annotations are indicated for clusters based on sample dissection and gene expression

profile. E. Stacked plot showing the percentage of samples contributing to each cluster. Percentage is normalized to the number of cells in the input, with legend shown on right. F. FeaturePlot showing localization of *POMC* neurons. G. Subclustering of POMC neurons reveals 11 unique POMC$^+$ clusters. H. Stacked plot showing the percentage of input samples contributing to each POMC cluster. Percentage is normalized to the number of cells in the input, with legend shown Fig. 3E above. I. DotPlot showing expression of genes localized within POMC subclusters. J. Transcription factor co-expression analysis across human/mouse and fetal/adult samples. Transcription factors coexpressing with POMC cells with a co-expression value three standard deviations above the mean are highlighted for each category. Dot size indicates relative the expression level for the gene of interest.

Figure 4. Developmental principles across distinct human brain regions. A. scRNAseq data from hypothalamus, cortex and ganglionic eminence for three matched human embryonic samples at GW18, GW19 and GW20. Cell types presented in 3D UMAP, inset panels are rotations of UMAP showing additional clusters. B. 3D UMAP showing overlay of samples by brain region. C. Distribution of cells within cell-type cluster by brain region. D. Polar plot of TF expression across three brain regions for dividing progenitor cells. Points off-axis reflect shared expression between regions. Grey points indicate no significant difference between regions. E. Heatmap depicting genes with similar or divergent expression patterns across neuronal lineages. Colors represent scaled expression of cells only within a given cell-type and brain region, as described in x-axis legend. F. Subsets of gene regulatory networks for hypothalamus, cortex and ganglionic eminence were created based on genes expressed in neural progenitor cells, radial glia and dividing progenitors. Resulting overlapping networks show edges from hypothalamus

network (blue) are roughly evenly split and adjacent to the ganglionic eminence network (green) and cortex network (red). Shared edges follow the color scheme introduced in Fig. 4D. Top 20 TFs with the greatest number of connections are highlighted in white.

Figure 1

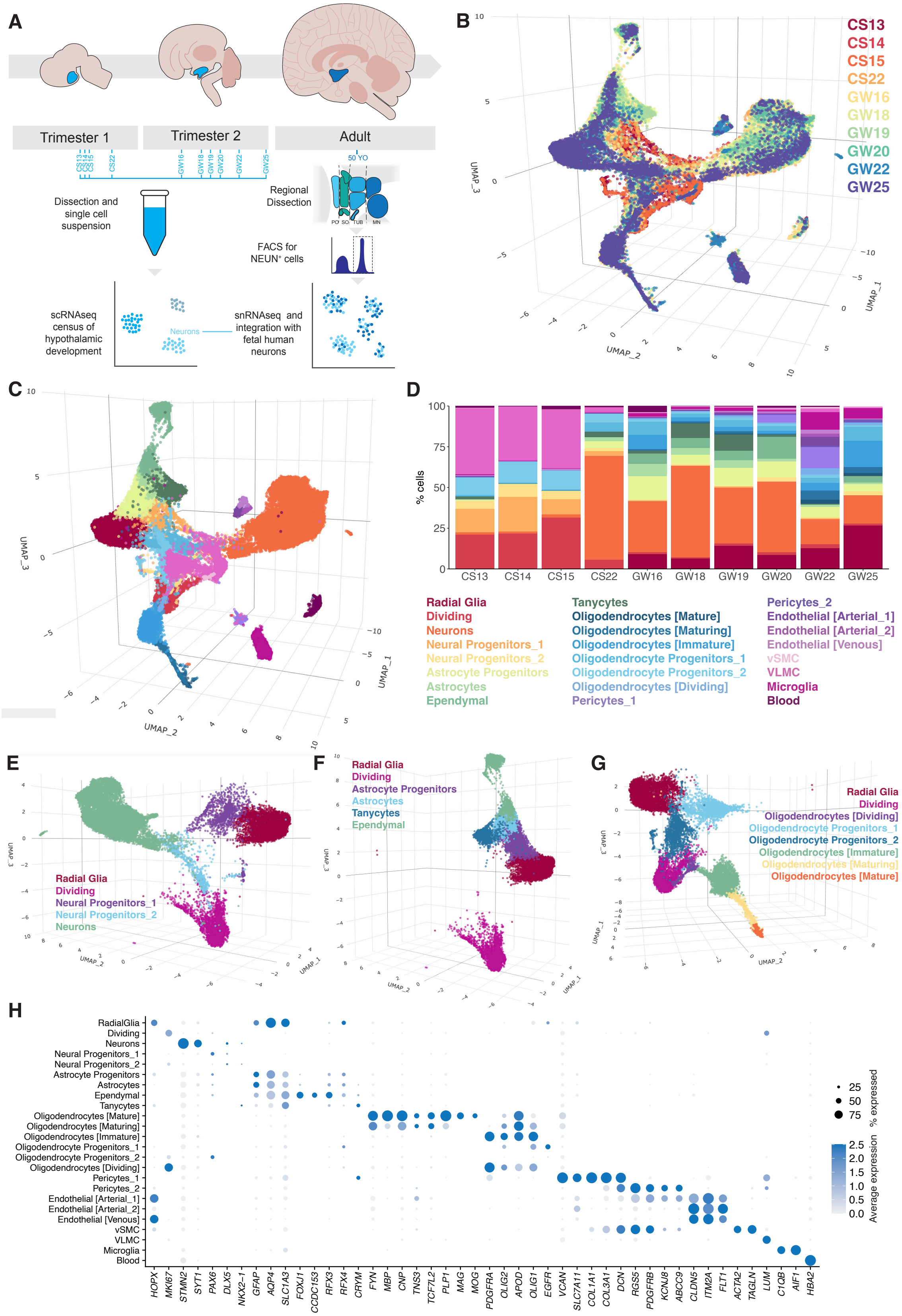

A
Trimester 1
Trimester 2
Adult
CS13
CS14
CS15
CS22
GW16
GW18
GW19
GW20
GW22
GW25
50 YO
Dissection and single cell suspension
Regional Dissection
PO SO TUB MN
FACS for NEUN⁺ cells
scRNAseq census of hypothalamic development
Neurons
snRNAseq and integration with fetal human neurons
B
CS13
CS14
CS15
CS22
GW16
GW18
GW19
GW20
GW22
GW25
UMAP_3
UMAP_2
UMAP_1
C
UMAP_3
UMAP_2
UMAP_1
D
% cells
100
75
50
25
0
CS13
CS14
CS15
CS22
GW16
GW18
GW19
GW20
GW22
GW25
Radial Glia
Dividing
Neurons
Neural Progenitors_1
Neural Progenitors_2
Astrocyte Progenitors
Astrocytes
Ependymal
Tanycytes
Oligodendrocytes [Mature]
Oligodendrocytes [Maturing]
Oligodendrocytes [Immature]
Oligodendrocyte Progenitors_1
Oligodendrocyte Progenitors_2
Oligodendrocytes [Dividing]
Pericytes_1
Pericytes_2
Endothelial [Arterial_1]
Endothelial [Arterial_2]
Endothelial [Venous]
vSMC
VLMC
Microglia
Blood
E
Radial Glia
Dividing
Neural Progenitors_1
Neural Progenitors_2
Neurons
UMAP_3
UMAP_2
UMAP_1
F
Radial Glia
Dividing
Astrocyte Progenitors
Astrocytes
Tanycytes
Ependymal
UMAP_3
UMAP_2
UMAP_1
G
Radial Glia
Dividing
Oligodendrocytes [Dividing]
Oligodendrocyte Progenitors_1
Oligodendrocyte Progenitors_2
Oligodendrocytes [Immature]
Oligodendrocytes [Maturing]
Oligodendrocytes [Mature]
UMAP_3
UMAP_2
UMAP_1
H
RadialGlia
Dividing
Neurons
Neural Progenitors_1
Neural Progenitors_2
Astrocyte Progenitors
Astrocytes
Ependymal
Tanycytes
Oligodendrocytes [Mature]
Oligodendrocytes [Maturing]
Oligodendrocytes [Immature]
Oligodendrocyte Progenitors_1
Oligodendrocyte Progenitors_2
Oligodendrocytes [Dividing]
Pericytes_1
Pericytes_2
Endothelial [Arterial_1]
Endothelial [Arterial_2]
Endothelial [Venous]
vSMC
VLMC
Microglia
Blood
HOPX
MKI67
STMN2
SYT1
PAX6
DLX5
NKX2−1
GFAP
AQP4
SLC1A3
FOXJ1
CCDC153
RFX3
RFX4
CRYM
FYN
MBP
CNP
TNS3
TCF7L2
PLP1
MAG
MOG
PDGFRA
OLIG2
APOD
OLIG1
EGFR
VCAN
SLC7A11
COL1A1
COL3A1
DCN
RGS5
PDGFRB
KCNJ8
ABCC9
CLDN5
ITM2A
FLT1
ACTA2
TAGLN
LUM
C1QB
AIF1
HBA2
25
50
75
% expressed
2.5
2.0
1.5
1.0
0.5
0.0
Average expression

Figure 2

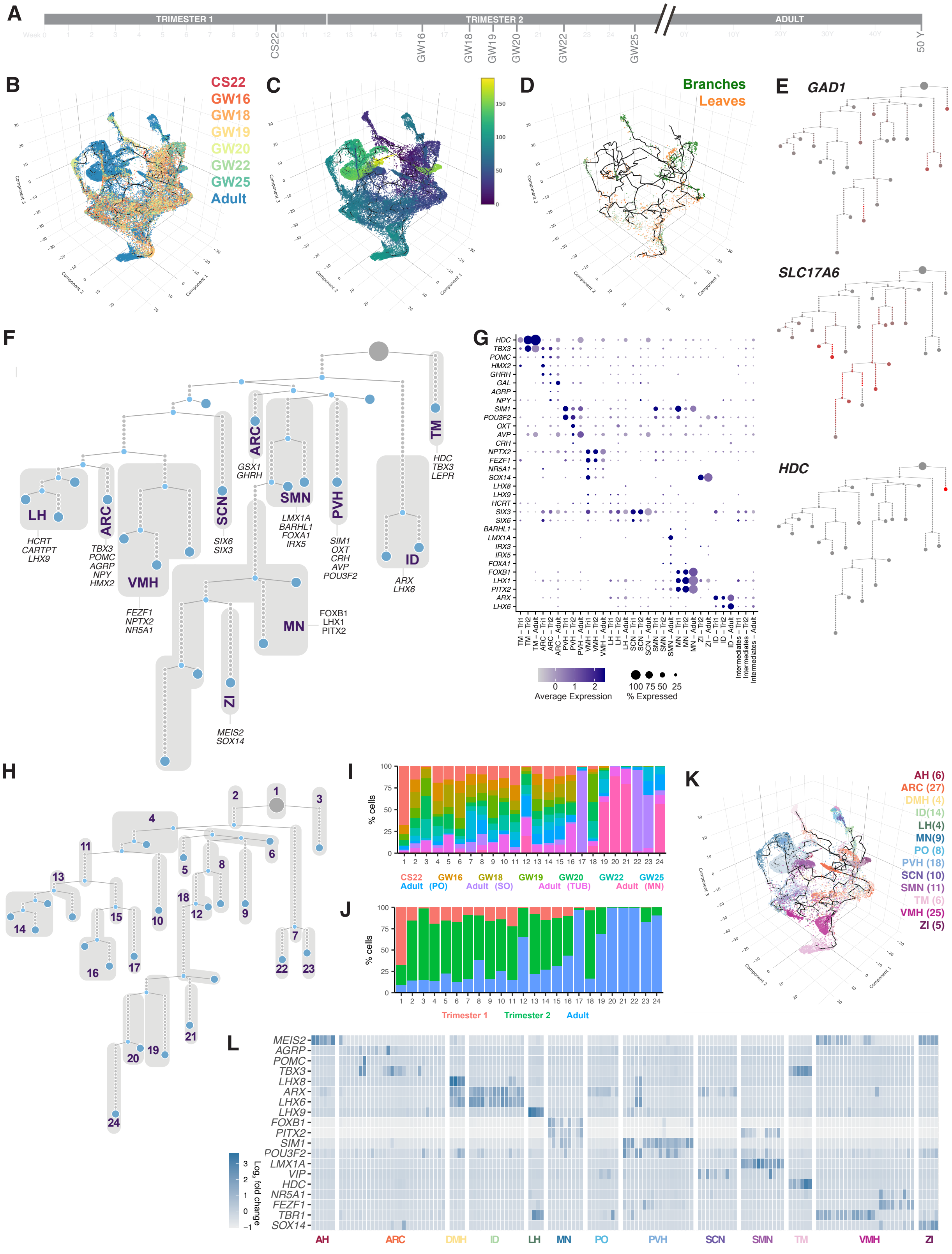

A
TRIMESTER 1
TRIMESTER 2
ADULT
CS22
GW16
GW18
GW19
GW20
GW22
GW25
50 Y
B
CS22
GW16
GW18
GW19
GW20
GW22
GW25
Adult
C
D
Branches
Leaves
E
GAD1
SLC17A6
HDC
F
LH
HCRT
CARTPT
LHX9
ARC
TBX3
POMC
AGRP
NPY
HMX2
VMH
FEZF1
NPTX2
NR5A1
SCN
SIX6
SIX3
ARC
GSX1
GHRH
SMN
LMX1A
BARHL1
FOXA1
IRX5
PVH
SIM1
OXT
CRH
AVP
POU3F2
ID
ARX
LHX6
TM
HDC
TBX3
LEPR
MN
FOXB1
LHX1
PITX2
ZI
MEIS2
SOX14
G
Average Expression
% Expressed
H
I
% cells
CS22
GW16
GW18
GW19
GW20
GW22
GW25
Adult (PO)
Adult (SO)
Adult (TUB)
Adult (MN)
J
% cells
Trimester 1
Trimester 2
Adult
K
AH (6)
ARC (27)
DMH (4)
ID(14)
LH(4)
MN(9)
PO (8)
PVH (18)
SCN (10)
SMN (11)
TM (6)
VMH (25)
ZI (5)
L
MEIS2
AGRP
POMC
TBX3
LHX8
ARX
LHX6
LHX9
FOXB1
PITX2
SIM1
POU3F2
LMX1A
VIP
HDC
NR5A1
FEZF1
TBR1
SOX14
Log2 fold change
AH
ARC
DMH
ID
LH
MN
PO
PVH
SCN
SMN
TM
VMH
ZI

Figure 3

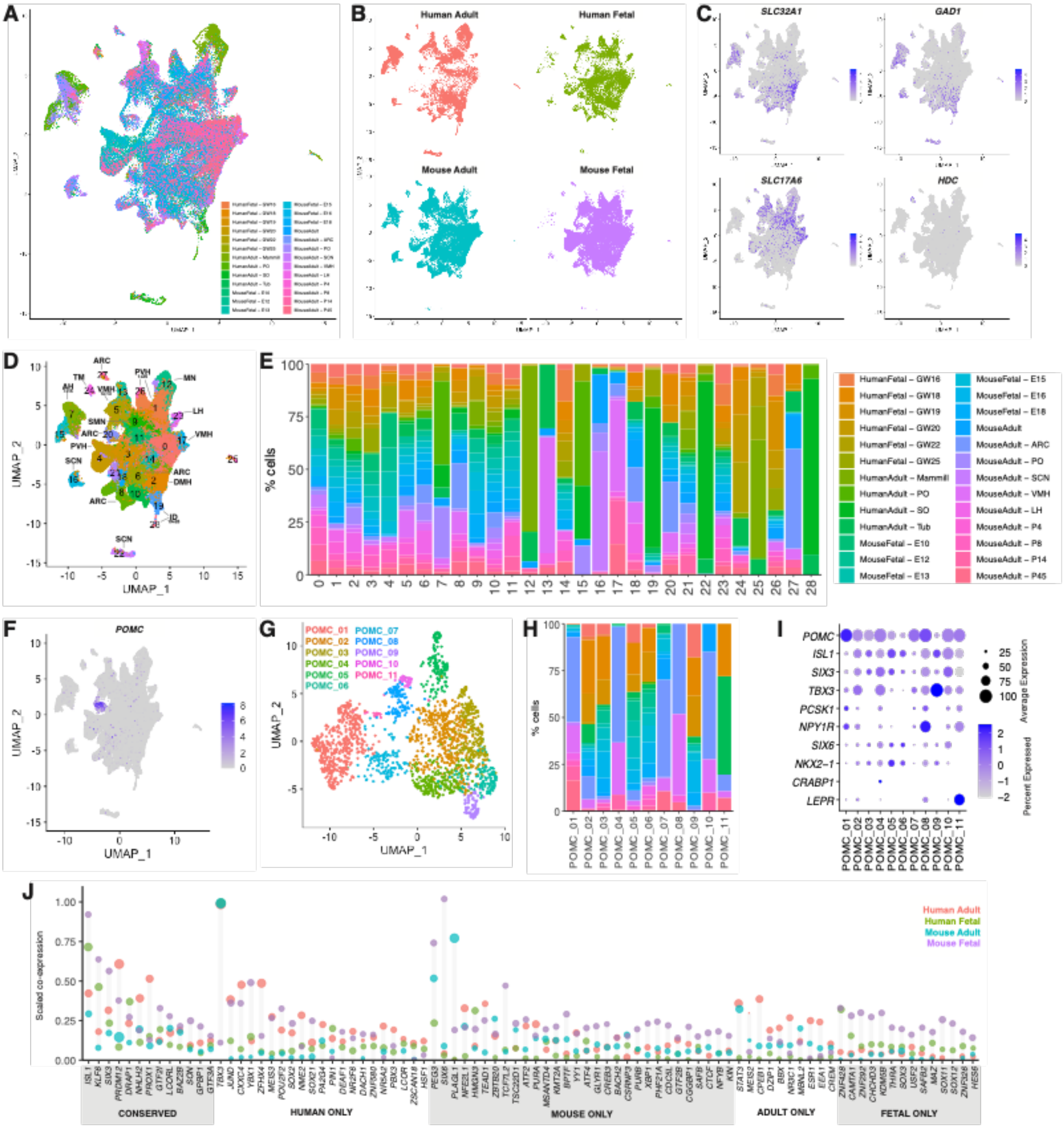
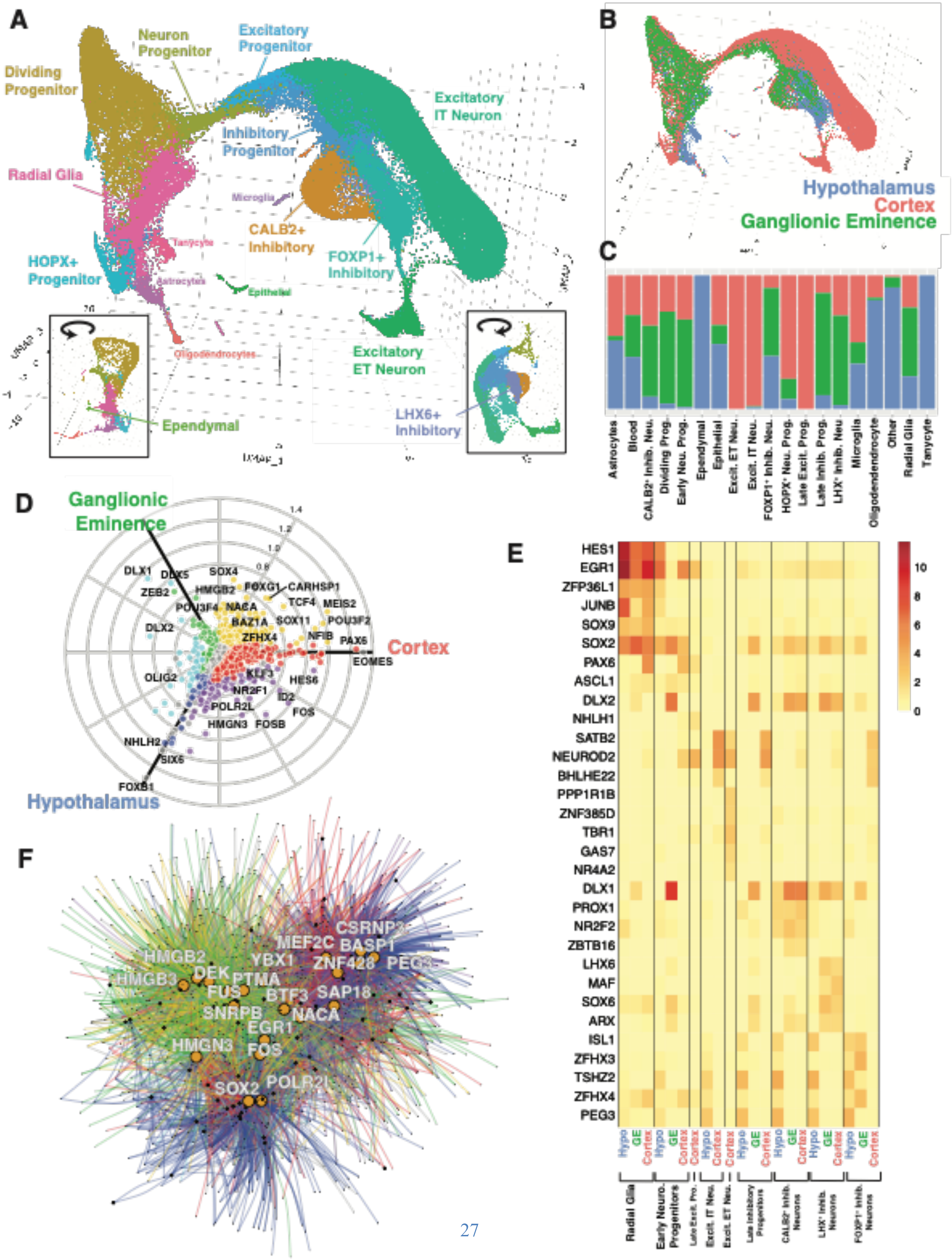

A
B
Human Adult
Human Fetal
Mouse Adult
Mouse Fetal
C
SLC32A1
GAD1
SLC17A6
HDC
D
UMAP_1
UMAP_2
ARC
TM
AH
PVH
VMH
MN
LH
SMN
SCN
DMH
ID
E
% cells
HumanFetal – GW16
HumanFetal – GW18
HumanFetal – GW19
HumanFetal – GW20
HumanFetal – GW22
HumanFetal – GW25
HumanAdult – Mammill
HumanAdult – PO
HumanAdult – SO
HumanAdult – Tub
MouseFetal – E10
MouseFetal – E12
MouseFetal – E13
MouseFetal – E15
MouseFetal – E16
MouseFetal – E18
MouseAdult
MouseAdult – ARC
MouseAdult – PO
MouseAdult – SCN
MouseAdult – VMH
MouseAdult – LH
MouseAdult – P4
MouseAdult – P8
MouseAdult – P14
MouseAdult – P45
F
POMC
G
POMC_01
POMC_02
POMC_03
POMC_04
POMC_05
POMC_06
POMC_07
POMC_08
POMC_09
POMC_10
POMC_11
H
% cells
I
POMC
ISL1
SIX3
TBX3
PCSK1
NPY1R
SIX6
NKX2−1
CRABP1
LEPR
Average Expression
Percent Expressed
J
Scaled co-expression
Human Adult
Human Fetal
Mouse Adult
Mouse Fetal
ISL1
KLF8
SIX3
PROM12
DRAP1
NHLH2
PROX1
GTF2I
LCORL
BAZ2B
SON
GPBP1
GTF3A
TBX3
JUND
CXXC4
YBX1
ZFHX4
MEIS3
POU2F2
SOX2
NME2
SOX21
PA2G4
PIN1
DEAF1
NR2F6
DACH1
ZNF680
NR4A2
PBX3
LCOR
ZSCAN18
HSF1
PEG3
SIX6
PLAGL1
NFE2L1
HMGN3
TEAD1
ZBTB20
TCF7L2
TSC22D1
ATF2
PURA
MSANTD4
KMT2A
BPTF
YY1
ATF4
GLYR1
CREB3
BACH2
CSRNP3
PURB
XBP1
PHF21A
CDC5L
GTF2B
CGGBP1
SAFB
CTCF
NFYB
KIN
STAT3
MEIS2
CPEB1
DZIP1
BBX
NR3C1
MBNL2
ESR1
EEA1
CREM
ZNF428
CAMTA1
ZNF292
CHCHD3
KDM5B
THRA
SOX3
USF2
SAFB2
MAZ
SOX11
SOX12
ZNF528
HES6
CONSERVED
HUMAN ONLY
MOUSE ONLY
ADULT ONLY
FETAL ONLY

Figure 4


A
Neuron Progenitor
Excitatory Progenitor
Dividing Progenitor
Excitatory IT Neuron
Inhibitory Progenitor
Radial Glia
Microglia
CALB2+ Inhibitory
Tanycyte
FOXP1+ Inhibitory
HOPX+ Progenitor
Astrocytes
Epithelial
Oligodendrocytes
Excitatory ET Neuron
Ependymal
LHX6+ Inhibitory
UMAP_1
UMAP_2
UMAP_3
B
Hypothalamus
Cortex
Ganglionic Eminence
C
Astrocytes
Blood
CALB2+ Inhib. Neu.
Dividing Prog.
Early Neu. Prog.
Ependymal
Epithelial
Excit. ET Neu.
Excit. IT Neu.
FOXP1+ Inhib. Neu.
HOPX+ Neu. Prog.
Late Excit. Prog.
Late Inhib. Prog.
LHX+ Inhib. Neu
Microglia
Oligodendendrocyte
Other
Radial Glia
Tanycyte
D
Ganglionic Eminence
Cortex
Hypothalamus
1.4
1.2
1.0
0.8
DLX1
DLX5
SOX4
ZEB2
HMGB2
FOXG1
CARHSP1
TCF4
MEIS2
POU3F4
NACA
BAZ1A
SOX11
POU3F2
DLX2
ZFHX4
NFIB
PAX6
EOMES
OLIG2
KLF3
HES6
NR2F1
ID2
POLR2L
FOS
HMGN3
FOSB
NHLH2
SIX6
FOXB1
E
HES1
EGR1
ZFP36L1
JUNB
SOX9
SOX2
PAX6
ASCL1
DLX2
NHLH1
SATB2
NEUROD2
BHLHE22
PPP1R1B
ZNF385D
TBR1
GAS7
NR4A2
DLX1
PROX1
NR2F2
ZBTB16
LHX6
MAF
SOX6
ARX
ISL1
ZFHX3
TSHZ2
ZFHX4
PEG3
10
8
6
4
2
0
Hypo
GE
Cortex
Radial Glia
Early Neuro. Progenitors
Late Excit. Pro.
Excit. IT Neu.
Excit. ET Neu.
Late Inhibitory Progenitors
CALB2+ Inhib. Neurons
LHX+ Inhib. Neurons
FOXP1+ Inhib. Neurons
F
CSRNP3
MEF2C
BASP1
HMGB2
YBX1
ZNF428
PEG3
HMGB3
DEK
PTMA
FUS
BTF3
SAP18
SNRPB
NACA
EGR1
HMGN3
FOS
SOX2
POLR2L

Supplementary Materials for

Title: Single-cell genomics reveals region-specific developmental trajectories underlying neuronal diversity in the human hypothalamus

Authors: Brian R. Herb† (1), Hannah J. Glover† (2), Aparna Bhaduri (3), Carlo Colantuoni (1,5), Tracy L. Bale (6), Kimberly Siletti (7), Sten Linnarsson (7), Rebecca Hodge (8), Ed Lein (8), Arnold R. Kriegstein (9,10), Claudia A. Doege* (11), Seth A. Ament* (1,12)

Affiliations:

[1]Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD
[2]Naomi Berrie Diabetes Center, Columbia Stem Cell Initiative, Department of Pediatrics, Columbia University Irving Medical Center, New York, NY
[3]Department of Biological Chemistry, University of California—Los Angeles, Los Angeles, CA
[4]Medical Scientist Training Program, University of Maryland School of Medicine, Baltimore, MD
[5]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA
[6]Department of Pharmacology, University of Maryland School of Medicine, Baltimore, MD
[7]Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden
[8]Allen Institute for Brain Science, Seattle, WA 98109
[9]Department of Neurology, University of California—San Francisco, San Francisco, CA
[10]The Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, University of California, San Francisco, CA
[11]Naomi Berrie Diabetes Center, Columbia Stem Cell Initiative, Department of Pathology and Cell Biology, Columbia University Irving Medical Center, New York, NY
[12]Department of Psychiatry, University of Maryland School of Medicine, Baltimore, MD

*Corresponding author. Email: sament@som.umaryland.edu, cad2114@cumc.columbia.edu

†These authors contributed equally to this work

This PDF file includes:

Materials and Methods

Figs. S1 to S6

Materials and Methods

Prenatal sample collection and processing. Acquisition of all primary human tissue samples was approved by the UCSF Human Gamete, Embryo and Stem Cell Research Committee (approval nos. 10-03379 and 10-05113). All experiments were performed in accordance with protocol guidelines. Informed consent was obtained before sample collection and use for this study. First and second trimester human hypothalamus tissue was collected from elective pregnancy termination specimens from San Francisco General Hospital and the Human Developmental Biology Resource (HDBR). Cortical and ganglionic eminence tissue was collected in parallel from the same specimens, as previously described(*29*, *30*). Hypothalamic tissue samples were dissociated using Papain (Worthington) containing DNase. Samples were grossly chopped and then placed in 1mL of Papain and incubated at 37C for 15 min. Samples were inverted three times and continued incubating for another 15 min. Next, samples were triturated by manually pipetting with a glass pasteur pipette approximately ten times. Dissociated cells were spun down at 300g for 5min and Papain removed.

Adult sample collection and processing. Adult hypothalamus tissue was obtained by the Allen Institute from a 50 year old Male donor (subject ID H18.30.002, Right hemisphere, Cause of death = Cardiovascular, PMI = 10hr, Tissue RIN score = 8.2 +/- 0.4) with no history of neuropsychiatric or neurological disorders and negative for infectious disease. Tissue collection was performed in accordance with the provisions of the United States Uniform Anatomical Gift Act of 2006 described in the California Health and Safety Code section 7150 (effective 1/1/2008) and other applicable state and federal laws and regulations. The Western Institutional Review Board reviewed tissue collection procedures and determined that they did not constitute human subjects research requiring institutional review board (IRB) review. The de-identified postmortem brain sample was obtained after receiving permission from the decedent's legal next-of-kin and prepared as described previously (full methods described in Hodge 2019(*50*) (dx.doi.org/10.17504/protocols.io.bf4ajqse) and Bakken 2021(*51*) describes subject ID

H18.30.002 sample). Briefly, coronal brain slabs of 1cm thickness were frozen in dry-ice cooled isopentane, and transferred to vacuum-sealed bags for storage at -80°C until the time of further use. To isolate the brain regions of interest, tissue slabs were briefly transferred to -20°C and the region of interest was removed and subdivided into smaller blocks on a custom temperature controlled cold table. Tissue blocks were stored at -80°C in vacuum-sealed bags until later use. **fig. S1** details the four broad regions – preoptic, supraoptic, tuberal and mammillary, identified by gross anatomical landmarks for the hypothalamus dissections. Nucleus isolation for 10x Chromium Single Cell 3' RNA sequencing V3 was conducted as described (dx.doi.org/10.17504/protocols.io.y6rfzd6). Gating on DAPI and NeuN fluorescence intensity was carried out as described previously(*50*). NeuN+ and NeuN- nuclei were sorted into separate tubes and were pooled at a defined ratio after sorting. Sorted samples were centrifuged, frozen in a solution of 1X PBS, 1% BSA, 10% DMSO, and 0.5% RNAsin Plus RNase inhibitor (Promega, N2611), and stored at -80°C until the time of shipment on dry ice from the Allen Institute to the Karolinska Institute for 10x chip loading.

Sequencing. For prenatal samples, single-cell capture was performed following the 10X v2 Chromium manufacturer's instructions. Each sample was its own batch. For each batch, 10,000 cells were targeted for capture and 12 cycles of amplification for each of the complementary DNA and library amplifications were performed. Libraries were sequenced according to the manufacturer's instructions on the Illumnia NovaSeq 6000 S2 flow cell (RRID:SCR_016387). For adult samples, immediately before loading on the 10x Chromium instrument, frozen nuclei were thawed at 37°C, washed, and quantified for loading as described (dx.doi.org/10.17504/protocols.io.nx3dfqn). In brief, suspensions were thawed in a 37˚C water bath, spun down briefly, and pipetted several times to mix. Nuclei were then processed according to the 10X Genomics protocol, targeting 5000 cells. We aimed to sequence two replicates per sample. Nearly all samples were processed with the 10X Genomics V3 kit. Samples were first sequenced to a shallow depth (approximately 1000 reads/cell) on the Illumina NextSeq platform to validate sample concentrations. Samples were then sequenced to approximately 100,000 reads/cell on the Illumina NovaSeq platform. After sequencing, saturation was calculated for each sample using the Preseq package (https://github.com/smithlabcode/preseq). Any samples that were not saturated to 60% were sequenced more deeply using Preseq predictions.

Quality filtering and integration. All fetal samples were quality filtered to include cells with the number of detected genes between 200 & 4000, a total UMI between 1,000 and 15,000 and a maximum 10 percent of reads mapping to mitochondrial genes. Adult neurons were filtered to include cells with at least 500 reads, and total UMI above 1000. For published datasets, cells passing quality filters defined in publication were used in this study. Doublets were detected using scDblFinder (default parameters, clust.method='overcluster') and discarded from the analysis. Unless otherwise noted, the *Seurat* package v4.0.5(*52*) was used for normalization, sample integration, clustering, differential gene expression analysis and plotting. Prior to integration, samples were normalized using SCT transform, regressing on percent mitochondrial content.

Cell type assignments. Integrated prenatal samples (n cells = 40,927) were subject to dimension reduction with PCA and were visualized in the 3D UMAP space. Cells were clustered using Louvain clustering and identified using established markers from published studies(*2*, *4–11*, *23*, *53–57*). Aberrant cells were manually reassigned based on UMAP coordinates to best reflect gene expression in feature plots. Extrahypothalamic cells such as *FOXG1+* cells (Telencephalon) and *NEUROD6+* cells (forebrain) were removed prior downstream analysis. For the adult sample etrahypothalamic cells such as *PPP1R1B*/*DRD1*/*PDYN*/*DRD2*/*ADORA2A+* cells (Nucleus accumbens) were removed from the anterior samples (preoptic and supraoptic) prior downstream analysis.

Neuronal lineage analysis. Prenatal and adult hypothalamic neurons (CS22-Adult; n= 33,811) were integrated. *Seurat* was then used for dimension reduction using PCA and generation of a 3D UMAP. The scrattch.hicat package was used to generate 170 neuronal clusters from the adult hypothalamic neurons. Nuclei were identified based on expression of marker genes for each nuclei based on previously published studies (**table S3**) and/or colocalization in the Allen Brain Atlas (**fig. S4**).

Monocle3(*58*) was used to generate pseudotime lineages encompassing both prenatal and adult neurons. The Monocle3 trajectory generated 438 vertexes. Vertex 8 was chosen as the starting node as it contained the highest proportion of the earliest time point - the CS22 cells (66%). Using this pseudotime lineage, we identified branch points where the lineages diverge, and ‘leaves’ where the lineages terminate. Using the all_simple_paths function in Monocle3 and the *igraph* package, we were able to generate 2D lineage trees to represent the lineage. Gene expression was overlaid onto these lineage trees which showed the cells organizing into distinct lineages representing 10 anatomically distinct regions.

Comparison across human/mouse and adult/prenatal neurons. Human adult hypothalamic neurons (n=34,662) and human fetal neurons (n= 16,451) were merged with mouse fetal samples (n=71,149) from Kim *et al.* 2020(*11*), and human adult data either from the whole mouse adult hypothalamus (n = 17,416) from Kim *et al.* 2020(*11*) and Chen *et al.* 2017(*5*), or specific nuclei from the adult mouse including Arcuate (n = 9,368) from Campbell *et al.* 2017(*4*), Preoptic nucleus (n=10,314) from Moffitt *et al.* 2018(*6*), Suprachiasmatic Nucleus (n=2,015) from Wen *et al.* 2020(*9*), Ventromedial Hypothalamus (n=12,062) from Kim *et al.* 2019(*7*), and the Lateral Hypothalamus (n=2,317) from Mickelsen *et al.* 2019(*8*). Louvain clustering was used to assign cells to 28 clusters. The distribution of cell types across these clusters was determined by first normalizing to the number of cells in each sample of the input.

The *POMC*$^+$ cluster was cleaned and then subclustered using louvain clustering. 11 clusters were determined based on distribution of arcuate marker genes (**table S7**). Transcription factors with high correlations to neuropeptides were determined for each sample type by co-occurance analysis using the *quanteda* package(*59*) in R and calculating the log likelihood significance for each pair of transcription factors and neuropeptides. From each of these matrices, the top transcription factors were determined as genes with co-expression above one standard deviation above the average of all correlations. If a gene appeared as a top transcription factor in all four matrices, they were considered conserved. Genes conserved in human were top transcription factors in both human datasets, but not in at least one of the mouse datasets. The transcription factors expressed in mouse, adult and fetal data were calculated accordingly.

Comparison across brain regions during human development. Three fetuses from gestational week 18, 19 and 20 were chosen for their broad range of dissections covering the hypothalamus, three regions of the ganglionic eminence and five regions of the cortex. Adult neurons were filtered to include cells with at least 1000 reads (maximum 15000), number of genes 500 - 4000, a maximum 10 percent of reads mapping to mitochondrial genes and total UMI above 1000. The 95,107 cells passing these filters were integrated using the *Seurat* package in R with the same parameters as used for other datasets in this study. Cell type assignments were determined by a mixture of our work with the hypothalamus and cortex assignments previously reported(*57*). Three dimensional UMAP projections assisted in identification of multiple excitatory and inhibitory lineages as well as distinguishing ET and IT excitatory neurons and the three inhibitory neuron populations. Identification of IT and ET neuron types were further confirmed using *projectR* package and the *DeCoN* dataset(*60*). Seurat clustering at a resolution of 0.5 largely delineated major cell types. The Seurat function FindAllMarkers was used to both find top marker genes for each Seurat cluster as well as identifying brain region differences among Seurat clusters. Polar coordinate plots and Z-scores were calculated using the *volcano3D* (v2.0.1) package(*61*) and were used to identify top transcription factor differences among dividing progenitor cells. Genes selected for the heatmap in Figure 4E were a combination of top marker genes per Seurat cluster that were both similar across brain regions and as well different. Gene regulatory networks were reconstructed using the *Genie3* package(*32*). Networks were first built for each individual sample then edge weights were averaged across samples to identify region-specific networks. For summary analysis of networks in dividing cells, networks were combined by finding common edges and the resulting network was plotted using the *iGraph* package. Transcription factors with the greatest number of connections were displayed in white in figure 4F. Gene modules were created using K-means clustering starting with a K=100 and resulting in 28 groups, with gene eigenvalues calculated with the *WGCNA* package(*62*). Counts tables for both Genie3 and K-means clustering analysis were imputed using K nearest neighbors smoothing, where K = 5. We then compared gene regulatory networks across brain regions to identify region specific transcription factor drivers for each gene module. Transcription factors with poor correlation between the transcription factor expression and the gene module eigenvalue were discarded. Lists of transcription factor drivers per gene module word compared across brain regions and similarities and differences were reported.

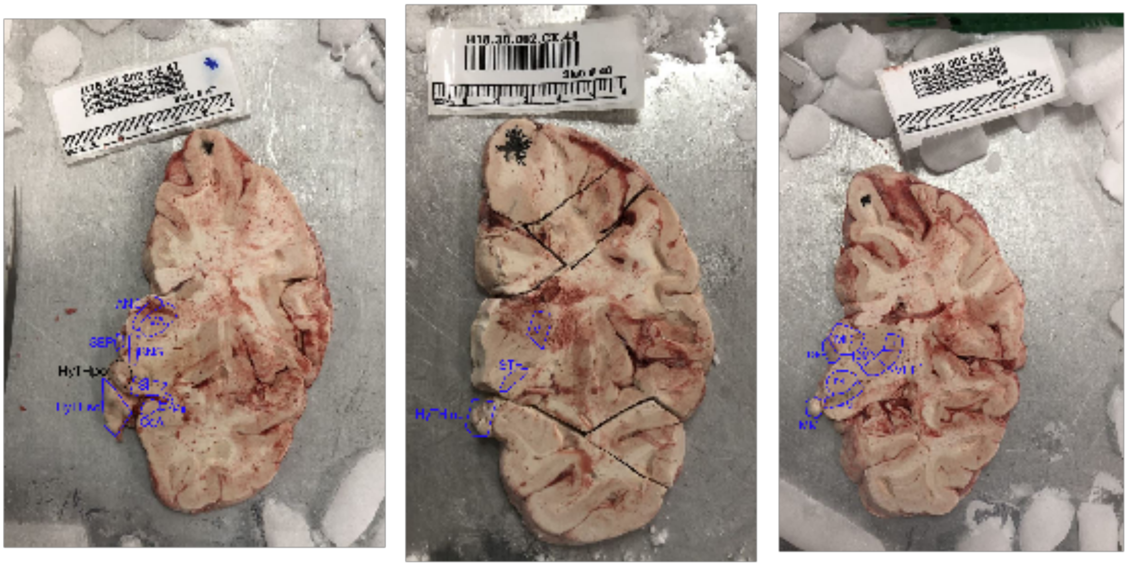


Fig. S1. Dissection of the human adult hypothalamus. Coronal sections were taken from a 50 year old healthy male. Anatomical landmarks were used to locate the preoptic hypothalamus (HyTHpo), supraoptic hypothalamus (HyTHso), tuberal hypothalamus (HyTHtub) and mammillary hypothalamus (MM).

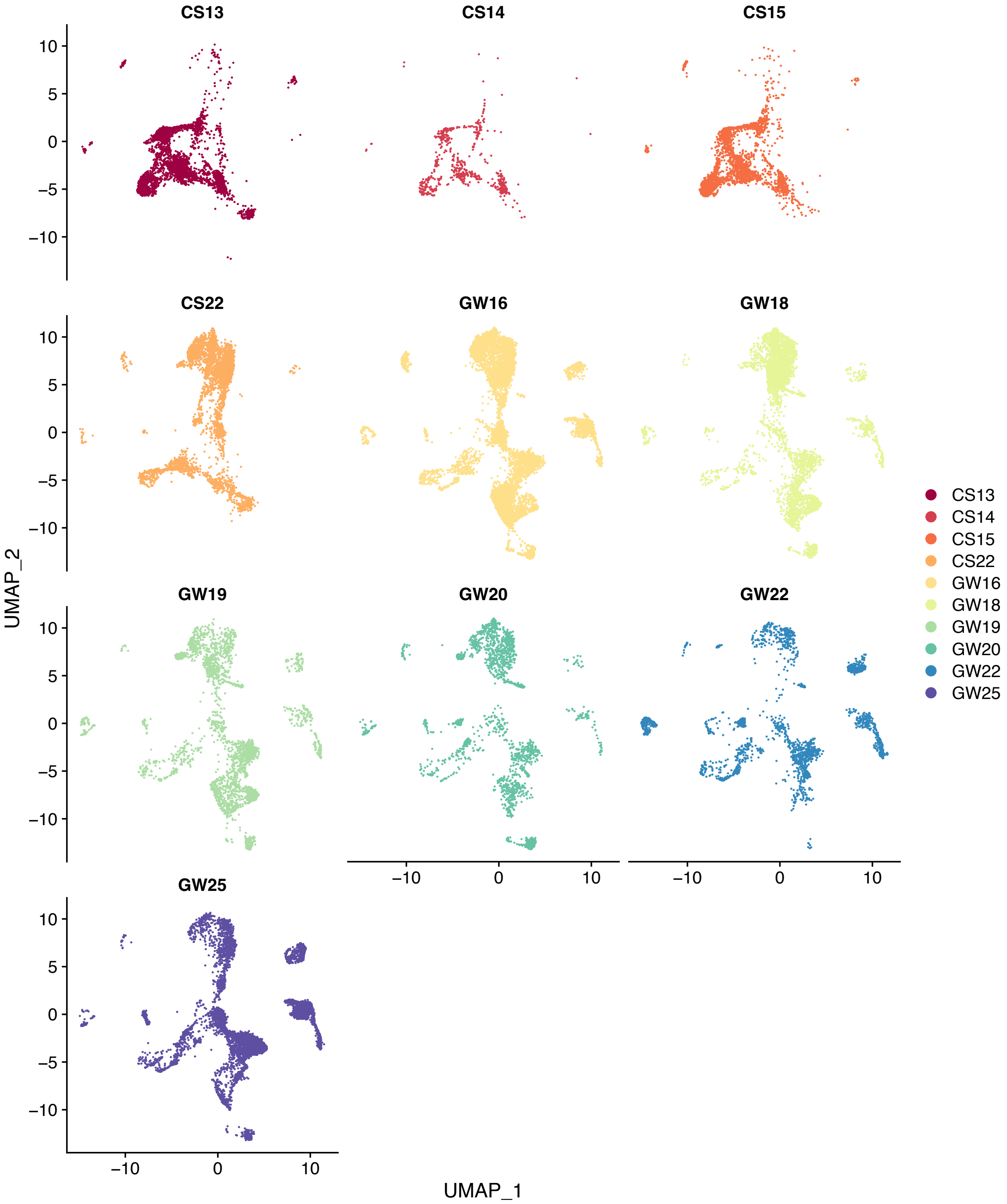
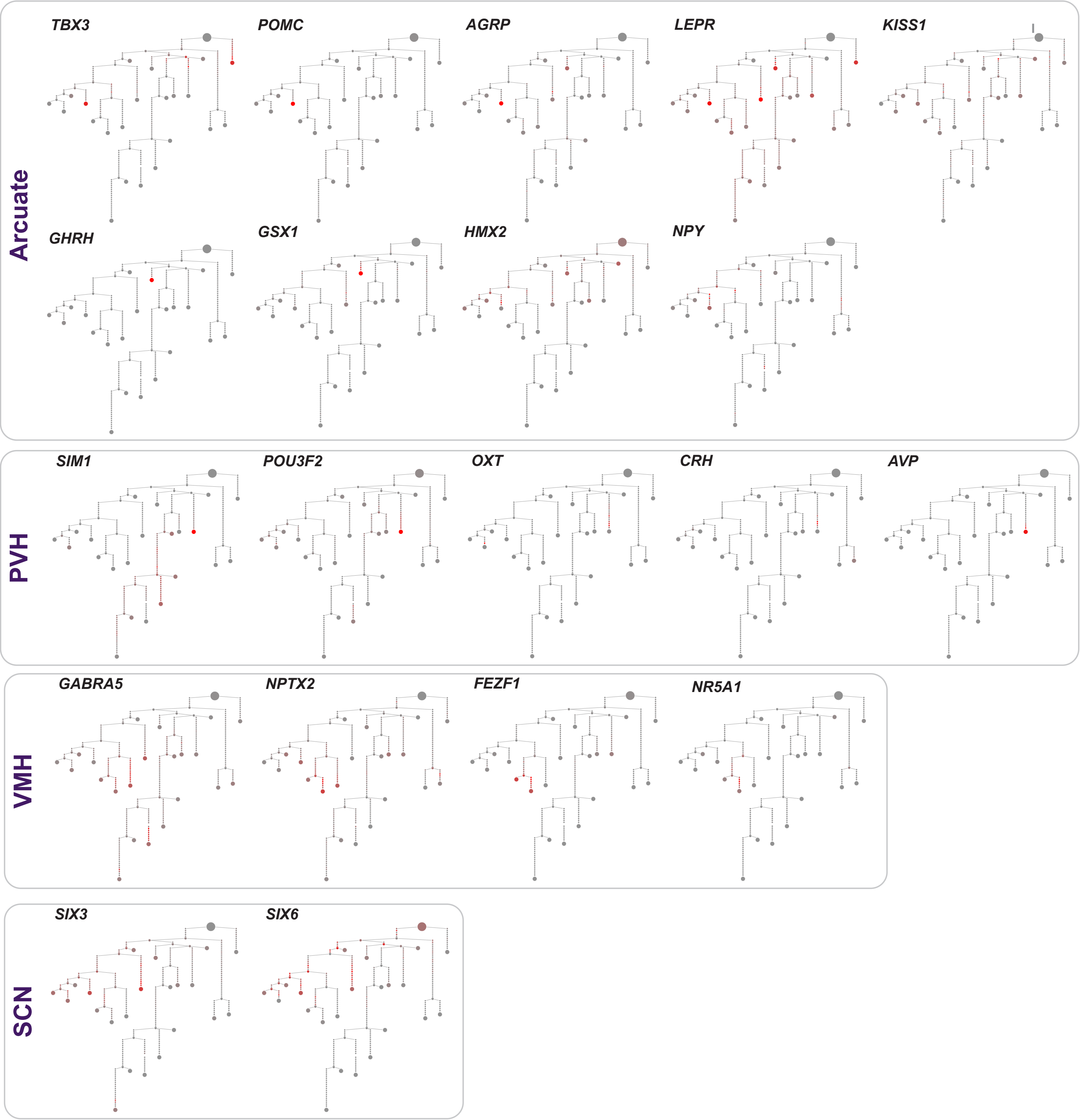


Fig. S2. Distribution of samples throughout UMAP space. 2D UMAP representation of integrated human embryonic samples, each plot depicts the coordinates of cells in each sample.


Arcuate
TBX3
POMC
AGRP
LEPR
KISS1
GHRH
GSX1
HMX2
NPY
PVH
SIM1
POU3F2
OXT
CRH
AVP
VMH
GABRA5
NPTX2
FEZF1
NR5A1
SCN
SIX3
SIX6

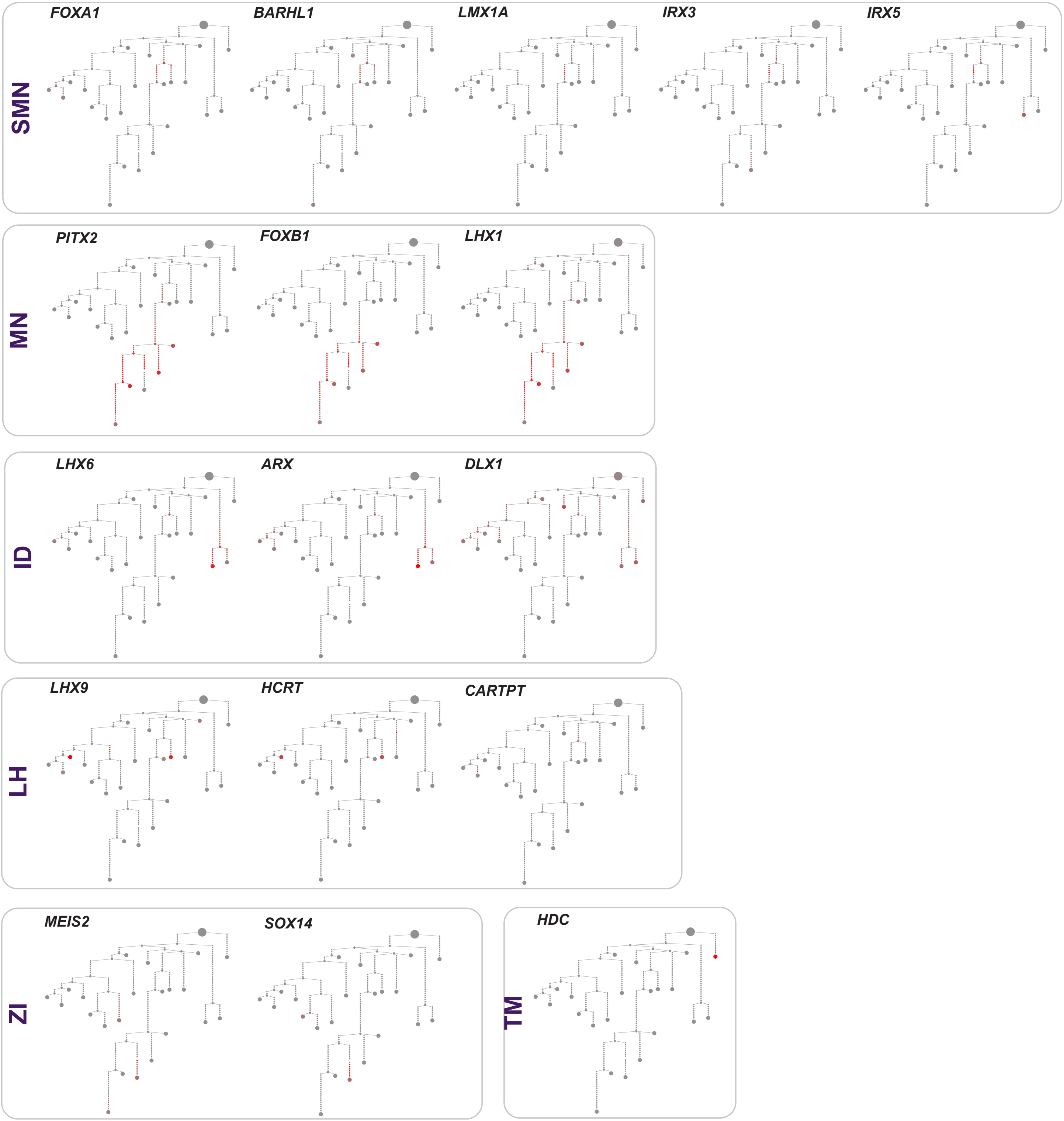
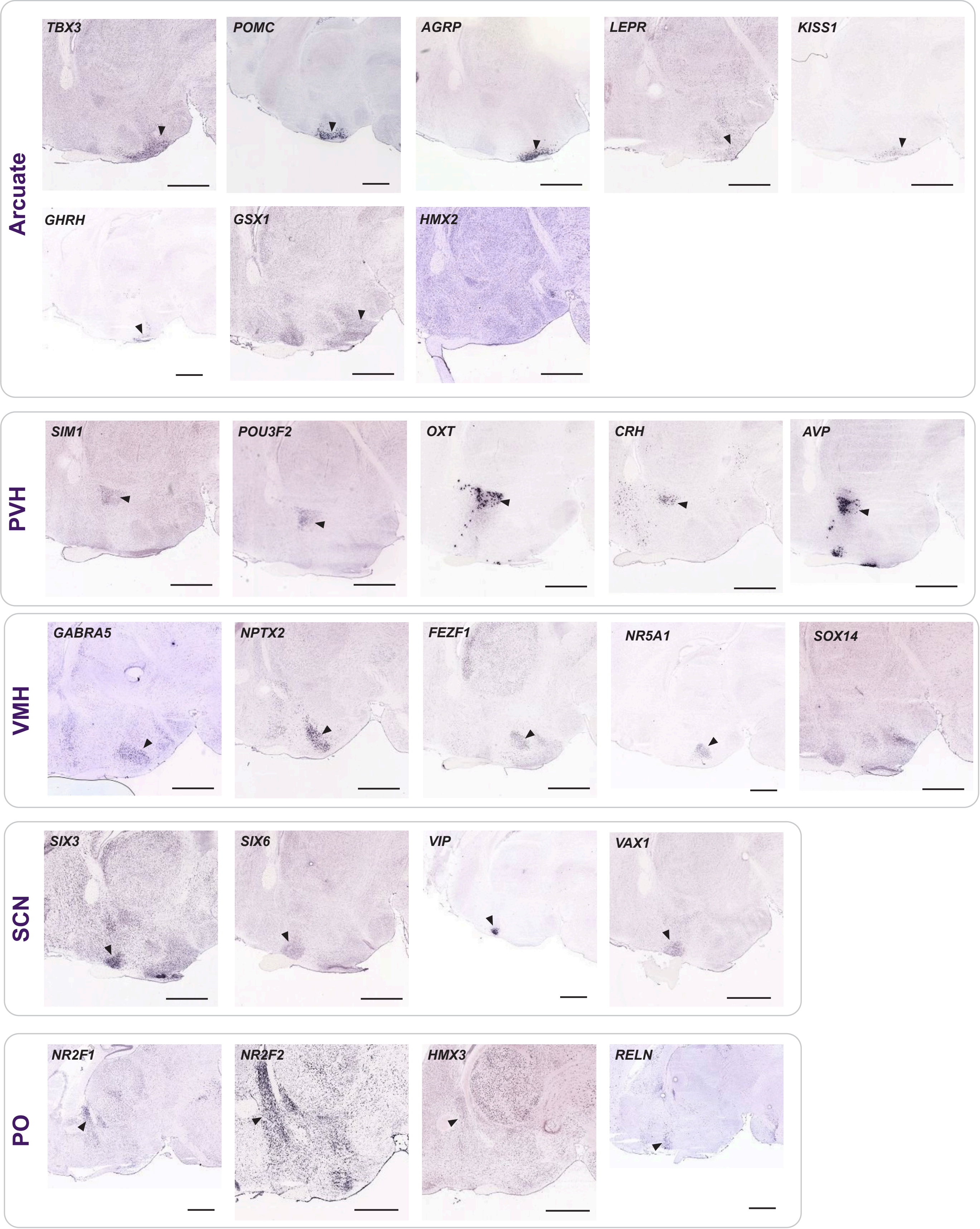

FOXA1
BARHL1
LMX1A
IRX3
IRX5
SMN
PITX2
FOXB1
LHX1
MN
LHX6
ARX
DLX1
ID
LHX9
HCRT
CARTPT
LH
MEIS2
SOX14
ZI
HDC
TM

Fig. S3. Expression of nuclei markers in embryonic and adult neuronal lineages. 2D lineage trees for all marker genes, with high expression indicated in red.


Arcuate
TBX3
POMC
AGRP
LEPR
KISS1
GHRH
GSX1
HMX2
PVH
SIM1
POU3F2
OXT
CRH
AVP
VMH
GABRA5
NPTX2
FEZF1
NR5A1
SOX14
SCN
SIX3
SIX6
VIP
VAX1
PO
NR2F1
NR2F2
HMX3
RELN

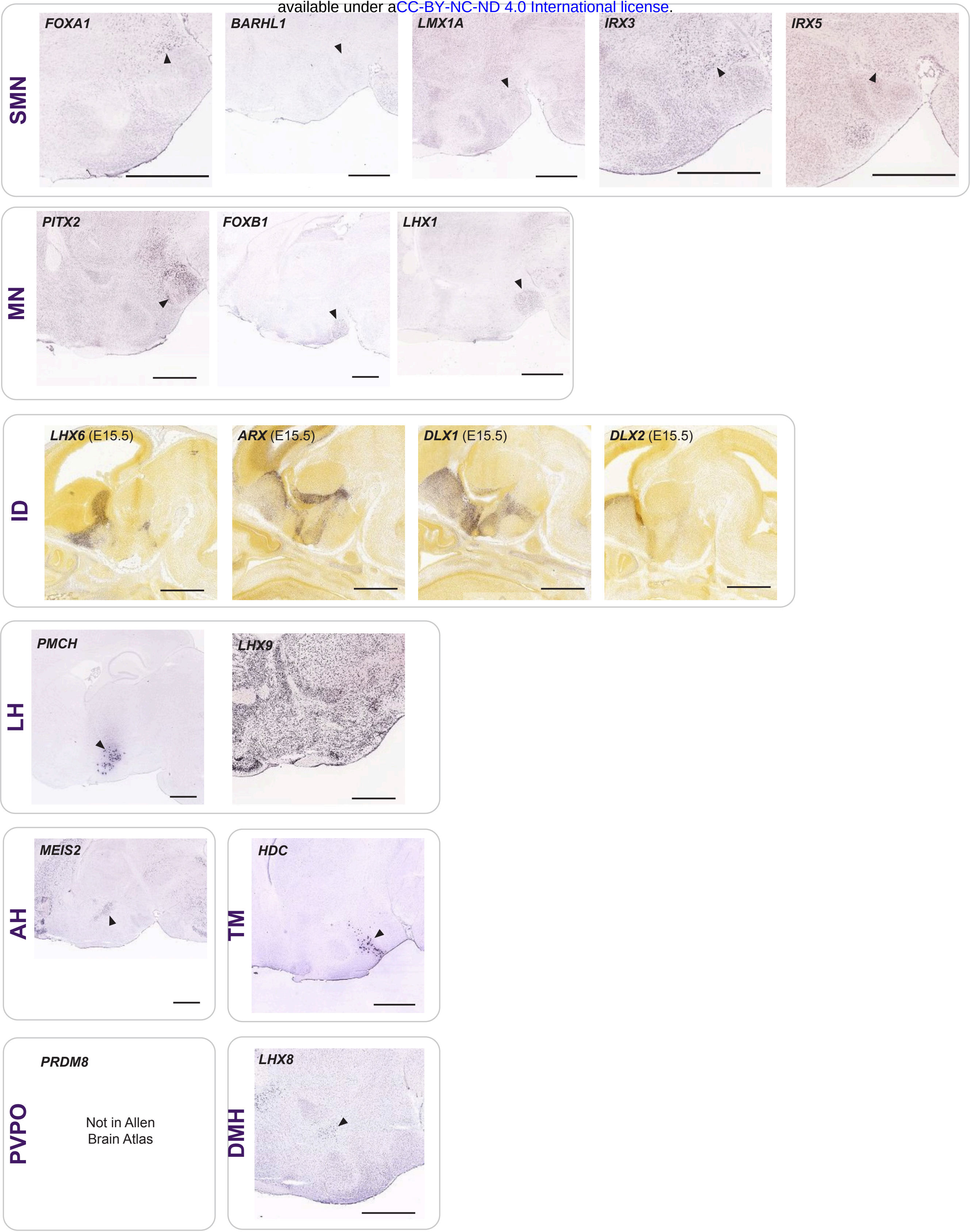

SMN
FOXA1
BARHL1
LMX1A
IRX3
IRX5
MN
PITX2
FOXB1
LHX1
ID
LHX6 (E15.5)
ARX (E15.5)
DLX1 (E15.5)
DLX2 (E15.5)
LH
PMCH
LHX9
AH
MEIS2
TM
HDC
PVPO
PRDM8
Not in Allen Brain Atlas
DMH
LHX8

Fig. S4. Localization of marker genes within the Allen Brain Atlas. In situ hybridization of marker genes used to discern neuronal lineages. Arrows show localization within the region of interest. Scale bar is 1000 microns.

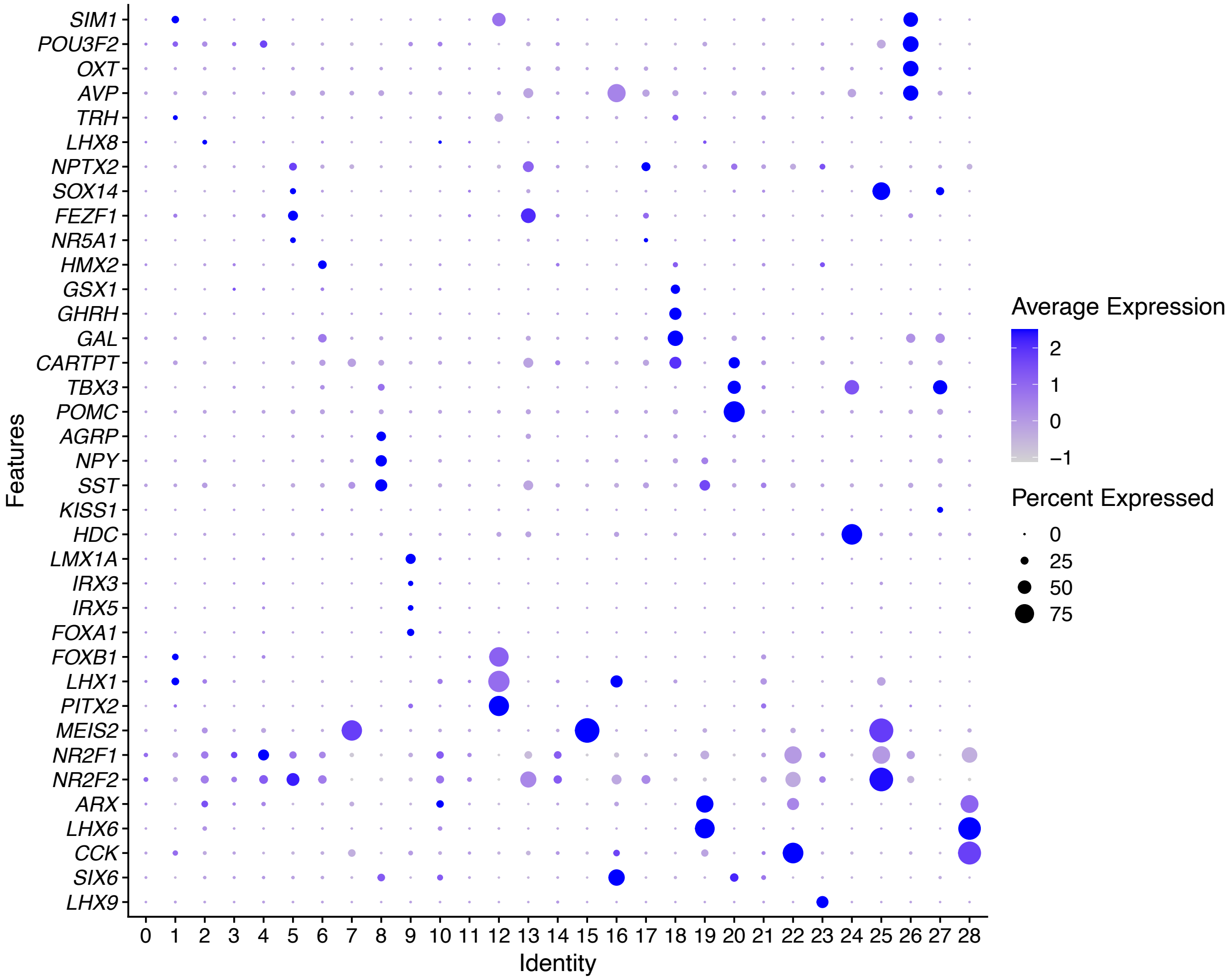
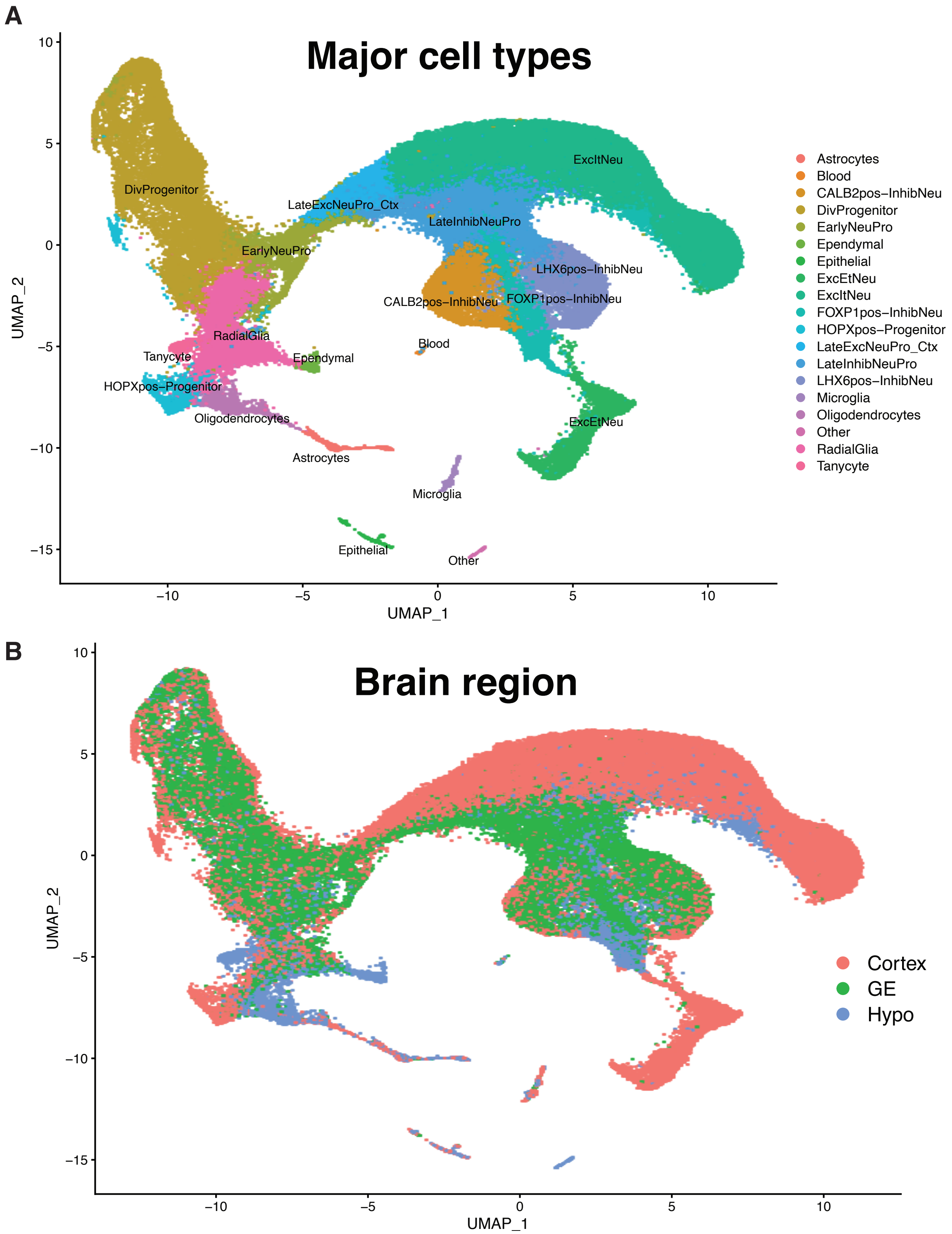
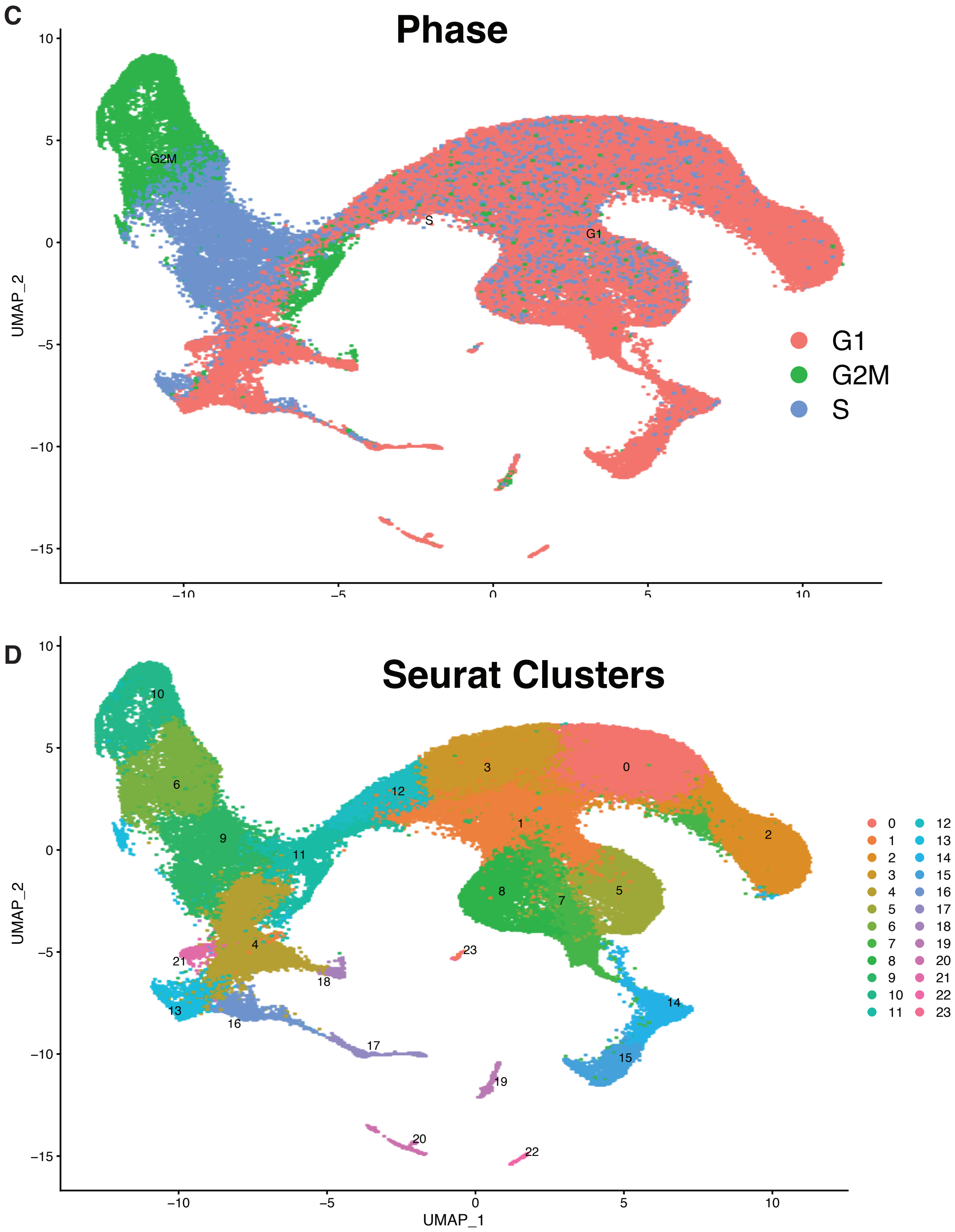
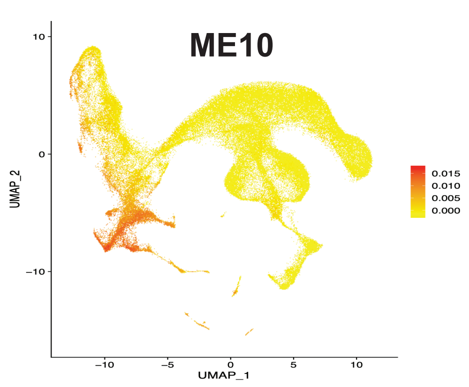
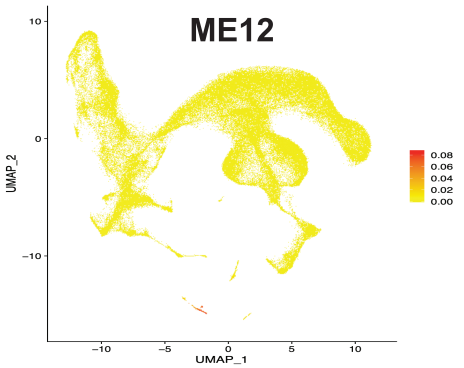


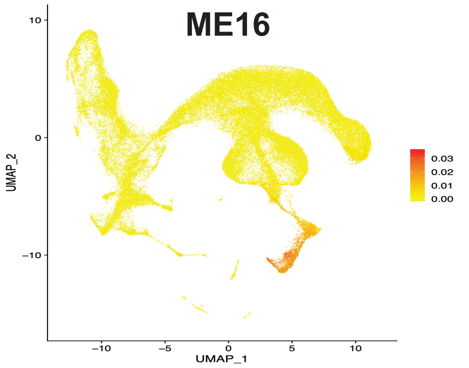
Fig. S5. Expression of nuclei markers across species and development. Louvain clustering was performed on the integrated data from mouse and human, prenatal and adult samples. Nuclei specific markers are localized with Seurat clusters.

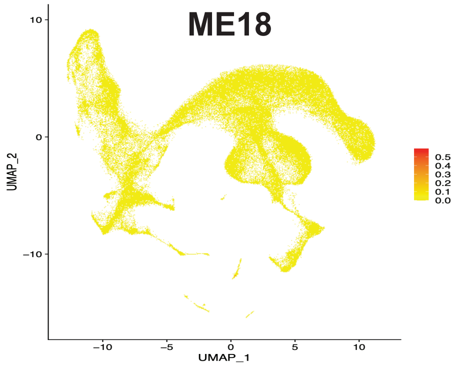

A
Major cell types
UMAP_2
UMAP_1
10
5
0
−5
−10
−15
−10
−5
0
5
10
DivProgenitor
LateExcNeuPro_Ctx
LateInhibNeuPro
ExcItNeu
EarlyNeuPro
LHX6pos−InhibNeu
CALB2pos−InhibNeu
FOXP1pos−InhibNeu
RadialGlia
Tanycyte
Ependymal
Blood
HOPXpos−Progenitor
Oligodendrocytes
ExcEtNeu
Astrocytes
Microglia
Epithelial
Other
● Astrocytes
● Blood
● CALB2pos−InhibNeu
● DivProgenitor
● EarlyNeuPro
● Ependymal
● Epithelial
● ExcEtNeu
● ExcItNeu
● FOXP1pos−InhibNeu
● HOPXpos−Progenitor
● LateExcNeuPro_Ctx
● LateInhibNeuPro
● LHX6pos−InhibNeu
● Microglia
● Oligodendrocytes
● Other
● RadialGlia
● Tanycyte
B
Brain region
UMAP_2
UMAP_1
10
5
0
−5
−10
−15
−10
−5
0
5
10
● Cortex
● GE
● Hypo

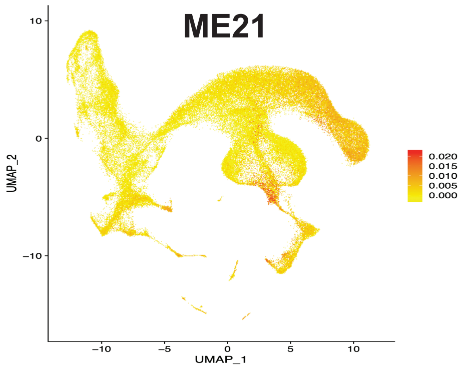

C
Phase
UMAP_2
10
5
0
−5
−10
−15
−10
−5
0
5
10
G2M
S
G1
G1
G2M
S
D
Seurat Clusters
UMAP_2
UMAP_1
10
6
9
11
12
3
0
1
2
8
7
5
4
21
18
13
16
17
23
14
15
19
20
22
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

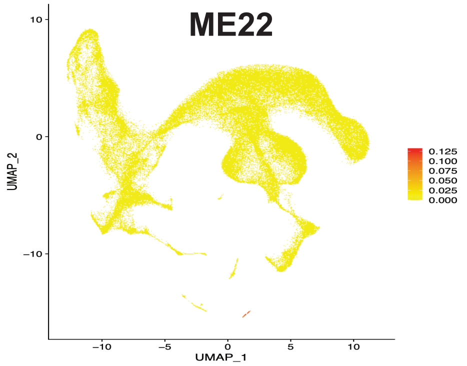

ME10
0.015
0.010
0.005
0.000
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10

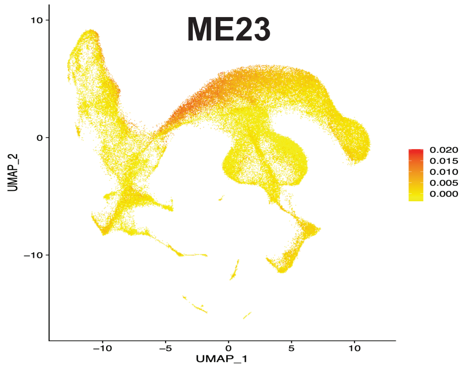

ME12
0.08
0.06
0.04
0.02
0.00
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10

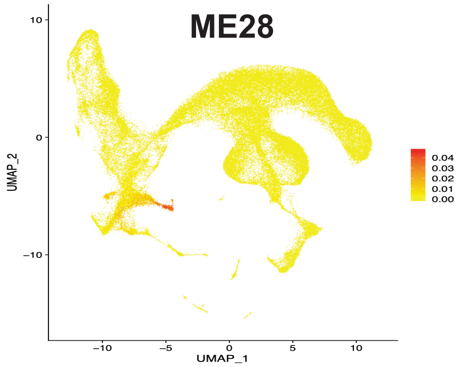

ME16
0.03
0.02
0.01
0.00
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10

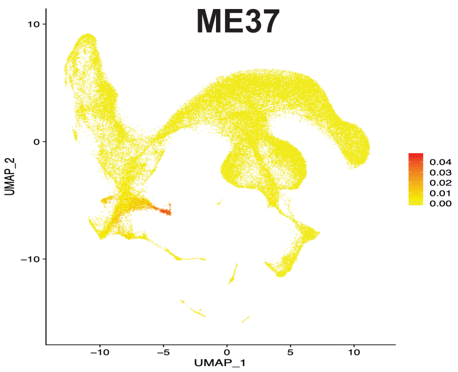

ME18
0.5
0.4
0.3
0.2
0.1
0.0
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10

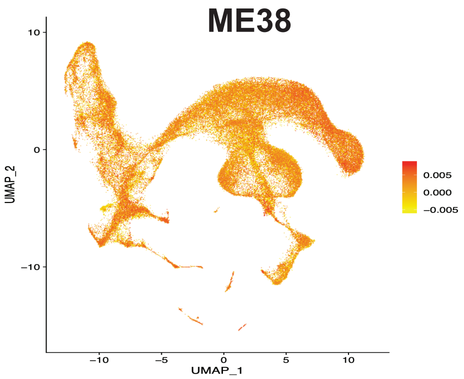

ME21
0.020
0.015
0.010
0.005
0.000
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10

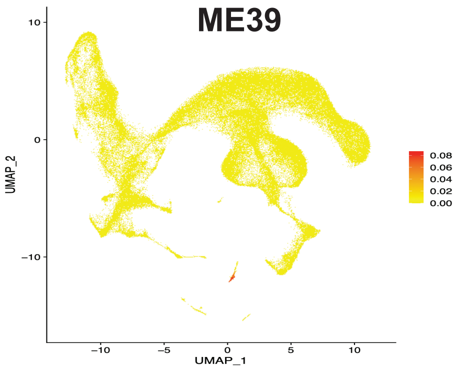

ME22
0.125
0.100
0.075
0.050
0.025
0.000
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10

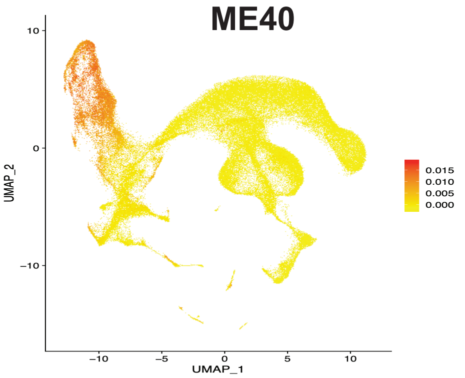

ME23
0.020
0.015
0.010
0.005
0.000
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10

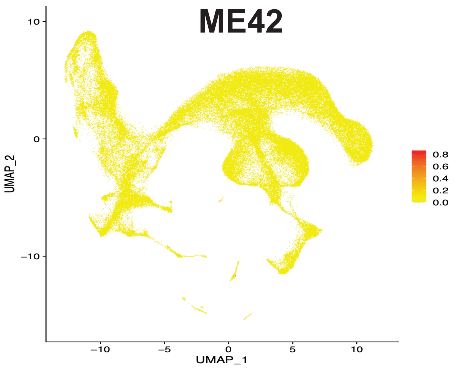

ME28
0.04
0.03
0.02
0.01
0.00
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10

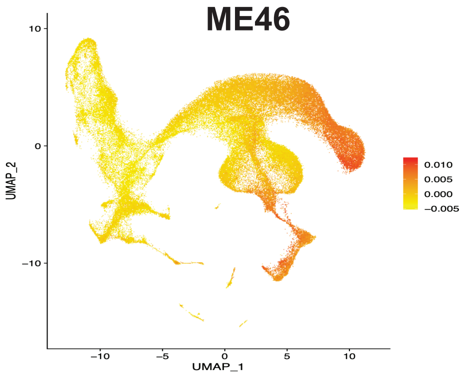

ME37
10
0
−10
UMAP_2
−10
−5
0
5
10
UMAP_1
0.04
0.03
0.02
0.01
0.00

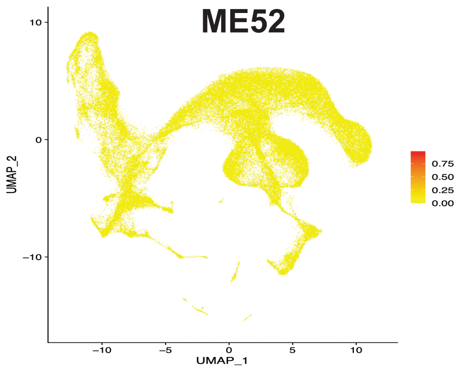

ME38
10
0
−10
UMAP_2
−10
−5
0
5
10
UMAP_1
0.005
0.000
−0.005

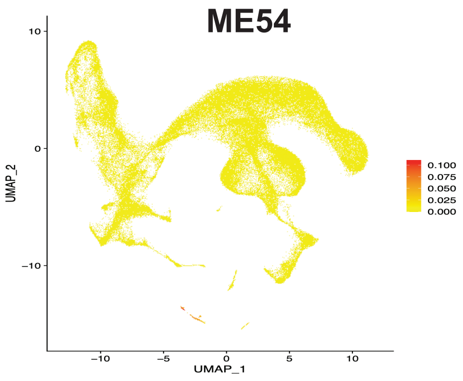

ME39
10
0
−10
UMAP_2
−10
−5
0
5
10
UMAP_1
0.08
0.06
0.04
0.02
0.00

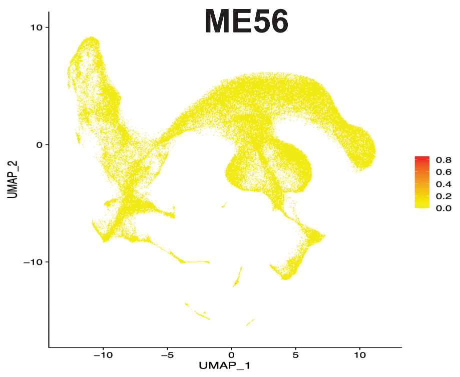

ME40
10
0
−10
UMAP_2
−10
−5
0
5
10
UMAP_1
0.015
0.010
0.005
0.000

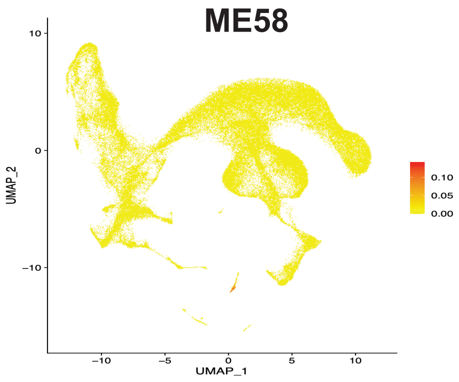

ME42
10
0
−10
UMAP_2
−10
−5
0
5
10
UMAP_1
0.8
0.6
0.4
0.2
0.0

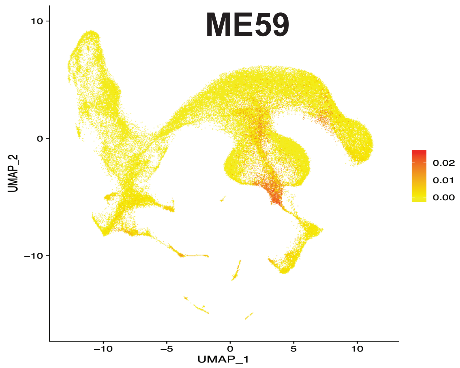

ME46
10
0
−10
UMAP_2
−10
−5
0
5
10
UMAP_1
0.010
0.005
0.000
−0.005

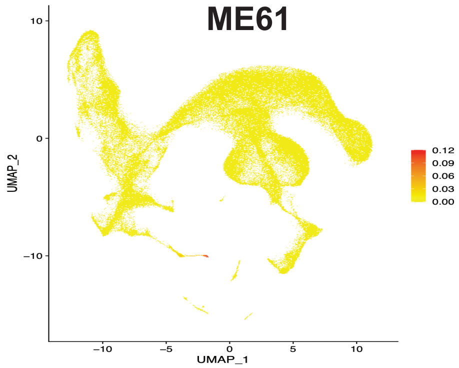

ME52
10
0
−10
UMAP_2
−10
−5
0
5
10
UMAP_1
0.75
0.50
0.25
0.00

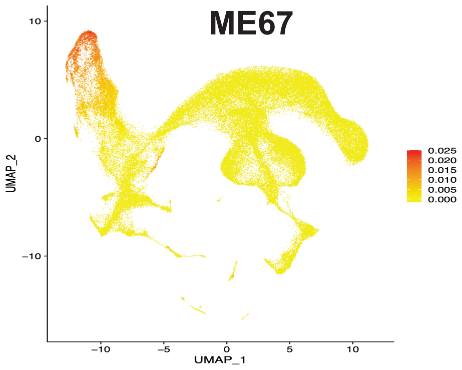

ME54
10
0
−10
UMAP_2
−10
−5
0
5
10
UMAP_1
0.100
0.075
0.050
0.025
0.000

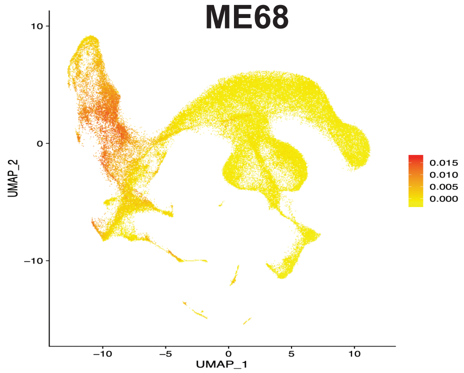

ME56
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10
0.8
0.6
0.4
0.2
0.0

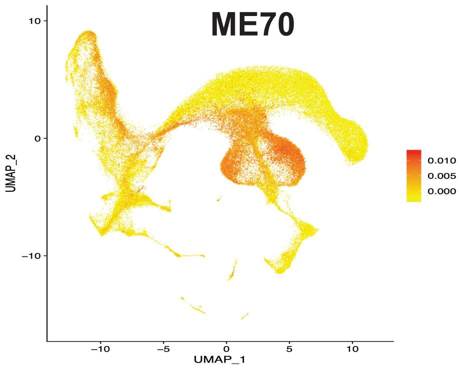

ME58
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10
0.10
0.05
0.00

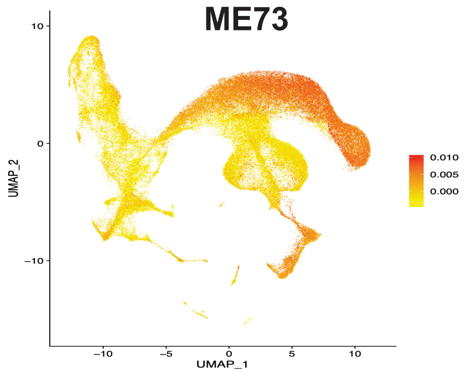

ME59
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10
0.02
0.01
0.00


ME61
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10
0.12
0.09
0.06
0.03
0.00


ME67
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10
0.025
0.020
0.015
0.010
0.005
0.000


ME68
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10
0.015
0.010
0.005
0.000


ME70
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10
0.010
0.005
0.000


ME73
UMAP_1
UMAP_2
10
0
−10
−10
−5
0
5
10
0.010
0.005
0.000

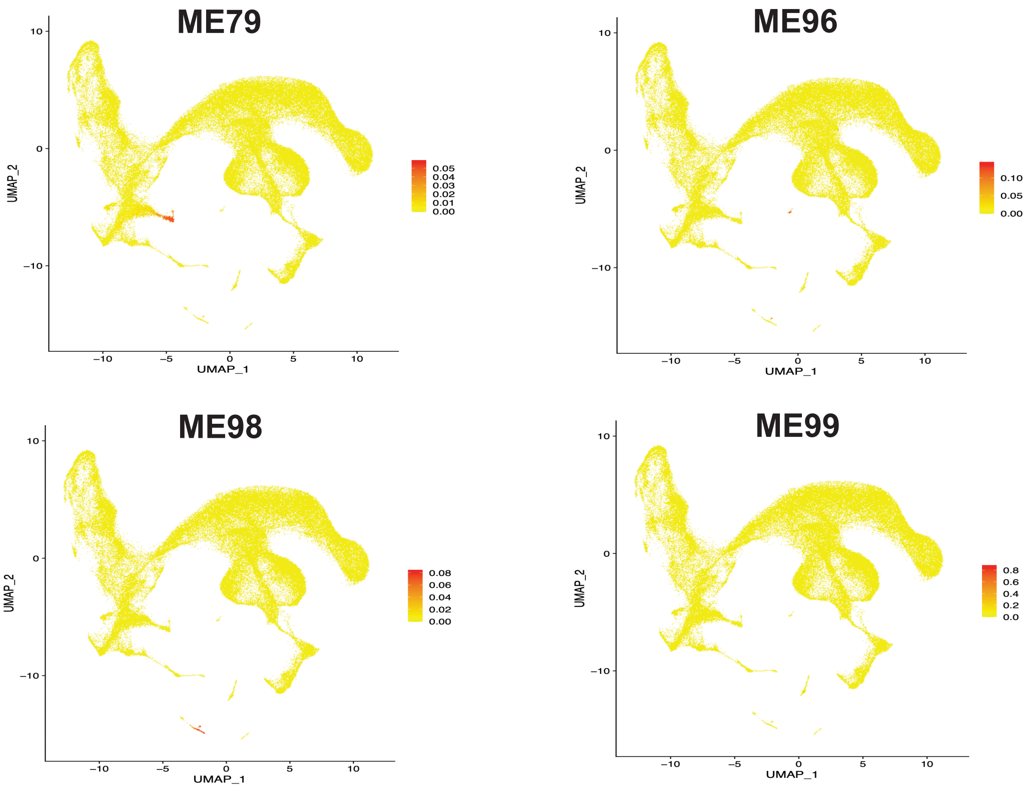


Fig. S6. Gene module activity across cell-types and neuronal lineages. **A.** Major cell types determined by marker genes and boundaries defined by Seurat clustering. (3D UMAP) **B.** Integration of cells from across the hypothalamus, ganglionic eminence and cortex are presented in a 2 dimensional UMAP. **C.** Cell cycle stage as calculated using the Seurat program. (3D UMAP) **D.** Louvain clustering of cells calculated using the FindClusters function in Seurat, default resolution set to 0.5. (3D UMAP) **E.** Gene modules were created using imputed counts and grouping genes by K-means clustering. Eigengene values were calculated using the WGCNA package and these values were plotted for each gene module.