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

Brian R. Herb (1), Hannah J. Glover (2), Aparna Bhaduri (3), Alex M. Casella (1,4), Tracy L. Bale (5,6), Arnold R. Kriegstein (7,8), Claudia A. Doege (9), Seth A. Ament (1,10)*

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 Pharmacology, University of Maryland School of Medicine, Baltimore, MD
6. Center for Epigenetic Research in Child Health and Brain Development, University of Maryland School of Medicine, Baltimore, MD
7. Department of Neurology, University of California—San Francisco, San Francisco, CA
8. The Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, University of California, San Francisco, CA
9. Naomi Berrie Diabetes Center, Columbia Stem Cell Initiative, Department of Pathology and Cell Biology, Columbia University Irving Medical Center, New York, NY
10. Department of Psychiatry, University of Maryland School of Medicine, Baltimore, MD

* Corresponding author:

Seth A. Ament
670 West Baltimore Street, HSF3-3182
Baltimore, MD 21201
sament@som.umaryland.edu

ORCID ID's :
Brian R. Herb - 0000-0002-5910-9647
Hannah J. Glover - 0000-0002-7319-645X
Aparna Bhaduri- 0000-0003-4625-6899
Alex M. Casella- 0000-0001-6873-7867
Tracy L. Bale- 0000-0002-3577-2428
Arnold R. Kriegstein - 0000-0001-5742-2990
Claudia A. Doege - 0000-0003-3871-4765
Seth A. Ament - 0000-0001-6443-7509

Abstract

The hypothalamus is critically important for regulating most autonomic, metabolic, and behavioral functions essential for life and species propagation, yet a comprehensive understanding of neuronal subtypes and their development in the human brain is lacking. Here, we characterized the prenatal human hypothalamus by sequencing the transcriptomes of 45,574 single-cells from 12 embryos, spanning gestational weeks 4 through 25. These cells describe a temporal trajectory from proliferative stem cell populations to maturing neurons and glia, including 38 distinct excitatory and inhibitory neuronal subtypes. Merging these data with paired samples from the cortex and ganglionic eminences (GE) revealed two distinct neurogenesis pathways, one shared between GE and hypothalamus and a second unique to cortex. Gene regulatory network modeling predicted that these distinct maturation trajectories involve the activation of region- and cell type-specific transcription factor networks. These results provide the first comprehensive transcriptomic view of human hypothalamus development at cellular resolution.

Introduction

Though small in size, the hypothalamus is one of the most functionally diverse and neuroanatomically complex regions of the human brain[1]. It plays critical roles in an astonishing variety of essential functions, including the regulation of body temperature, sexual dimorphisms, circadian rhythms, sleep, stress responses, satiety, and hunger, and aspects of mood, social behavior, and memory[2–4]. These functions are subdivided amongst specialized nuclei, each of which contains a unique set of transcriptionally and functionally distinct neuronal subtypes[5–12].

Hypothalamic neuronal and glial diversity arises during prenatal brain development. In mice, precursor populations first attain molecular identities corresponding one of the distinct nuclei, then further differentiate toward distinct nucleus-specific subtypes[12]. Hypothalamic development is sensitive to both environmental and genetic perturbations. Adverse prenatal environments such as maternal stress or malnutrition result in long-lasting changes in physiology and behavior, mediated by neurons within specific hypothalamic nuclei[13–15]. Hypothalamic neurons specifically express a number of genes known to be sensitive to the neuropeptide leptin, and/or are located near single nucleotide polymorphisms (SNPs) associated with obesity[16]. These clinical consequences motivate deeper investigation into the timing and regulation of hypothalamic development.

Much of what we know about the biological functions, anatomical and physiological characteristics, and development of the hypothalamus is derived from experimental studies in animal models. 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[17,18]. Here, we sought to address this deficiency through single cell RNA sequencing (scRNA-seq) of the prenatal human hypothalamus to define its transcriptional cell types and developmental trajectories.

Results

A map of neuronal and non-neuronal lineages in the prenatal human hypothalamus

We generated 10x Genomics scRNA-seq of prenatal human hypothalamus from 12 human fetuses at ~4 to 25 gestational weeks, yielding 45,574 high-quality single-cell transcriptomes with 752 – 7,690 cells per sample (Supplementary Table 1). 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[19–22]. Integration of all 45,574 hypothalamic cells revealed clusters and trajectories corresponding to each of the major cell types. Progenitor cells emanated from a common pool of dividing cells to form three lineages leading to neurons, oligodendrocytes, or astrocytes and other glia (Figure 1a; Figure S1a,b). Each cell type was defined by a unique set of cell type-specific transcription factors (Figure 1b) and other marker genes (Supplementary Table 2). Gene Ontology (GO) analysis confirmed enrichments for known markers for neurons, oligodendrocytes, and other glial populations (Supplementary Tables 3-5). Cell type distributions shifted over developmental time (Figure 1c), with major neuronal populations emerging as early as CS20 (~8 gestational weeks). As expected, oligodendrocytes, astrocytes, and specialized glial populations such as tanycytes and ependymal cells developed later, with progenitor populations diversifying during the first trimester, and populations of mature glial cells expanded during the second trimester.

Specification of hypothalamic nuclei during the first trimester

We re-clustered 20,574 cells from the first trimester -- gestational weeks (GW) 4.5-7.5; Carnegie Stages (CS) 13, 14, 15, 20, and 22 -- to gain insight into early patterning of the hypothalamus. These samples encompass developmental stages spanning from the early formation of the prosencephalon (CS12-CS13) to the development of an anatomically defined hypothalamus with distinct nuclei (CS20-CS22). Cell type composition across these developmental stages supported a progression from progenitor populations to maturing neuronal subtypes that cluster by hypothalamic nucleus. Multiple populations of radial glia and related progenitor cells are already apparent at these time points, with post-mitotic neurons beginning to appear, emanating from radial glia (Figure S2a-c).

In humans, hypothalamic nuclei begin to form late in the first trimester and can be identified histologically at this time by the expression of early marker genes and varying cell densities[23]. To gain insight into the gene expression programs underlying these patterning events, we again re-clustered the cells focusing specifically on two samples at CS22 (GW ~7.5). We used known markers from spatial gene expression profiling in mouse to assign neurons in these samples to specific nuclei[5–12,24]. Nucleus identity was reflected in discrete clusters in UMAP space (Figure 2a). These clusters were distinguished by hundreds of differentially expressed genes, including transcription factors and neurotransmitters that marked the unique functionality of each nucleus (Figure 2b, Figure S2d-h, Supplementary Table 6). Distinct functional categories of developmental genes were active in each nucleus, revealing nucleus-specific developmental programs (Supplementary tables 7-9*).* Among the transcription factors detected were well-established drivers of mammalian hypothalamic development including *NKX2-1, OTP, SIM1, BRN2, LHX6, IRX5, LHX8, LHX1, DLX2, ARX, LHX3*[24]. Specific neuronal subtypes within and across nuclei matured at different times. For example, among canonical neurons of the arcuate,

POMC was already expressed by CS22, whereas expression of *NPY* was not yet apparent. These dynamics resemble the time course of expression of *POMC* and *NPY* as seen during mouse embryonic development. Specifically, *POMC* expression was first observed in the hypothalamic ventricular zone at E10.5–E11.5 and *NPY* in laterally-situated cells in the rostral most presumptive ARC at E13.5[25]. Other major hypothalamic neuropeptides expressed at low levels in first trimester include *GHRH, TRH, AVP* and *GAL.* all of which continue expression in the second trimester samples. Additionally, *PNOC* expression was restricted to the PMN nucleus at CS22, then spread to neuronal subtypes in other nuclei later in maturation. Neuropeptides expressed in the second trimester but not first trimester included *AGRP, KISS1, OXT,* and *CRH*.

Maturation of hypothalamic neural and glial subtypes during the second trimester

Hypothalamic nuclei and post-mitotic neurons mature during the second trimester. At this stage, dividing progenitors constitute a minority of cells, giving way to well defined clusters of maturing neurons and glia marked by unique sets of TFs and vast gene expression differences (Figure S2i-j). Maturation of cell populations at this timepoint reflects the overall refinement of hypothalamic structure established in the second trimester that persists through development[23]. We re-clustered the cells from five samples at GW18-GW25 to characterize signatures of maturing neuronal, oligodendrocyte, and glial populations along with their respective progenitor populations.

Sub-clustering of 6,124 second trimester neurons and neural progenitors revealed 38 molecularly distinct clusters. Many clusters could be assigned to the nuclei in which they reside by the expression of known nucleus-specific markers (Figure 3a,b). Moreover, most clusters could be annotated to a known hypothalamic neuronal subtype by the expression of neurotransmitters and neuropeptides, their receptors, and TFs. Well-defined neuronal subtypes at this stage included components of the magnocellular system, including *OXT*-expressing neurons (PVH_1 cluster) and *AVP* neurons (PVH_2); components of the parvicellular system, including *GHRH* neurons (ARC_9), *POMC* neurons (ARC_3), and *TRH* neurons (PVH_3); subtypes involved in the regulation of hunger and satiety, including *SST* neurons (ARC_5), *AGRP/NPY* neurons (ARC_4), and *POMC* neurons (ARC_3); and subtypes involved in the regulation of wakefulness, including *HDC* neurons of the tuberomammillary nucleus (TM cluster) and glutamatergic neurons of the supramammillary nucleus (SMN_1, SMN_2) (Figure 3c)[4]. Hierarchical clustering of these neuron types revealed TFs enriched within each branch, predicting key regulators of neuronal subtype specification (Figure S3a).

Specialized hypothalamic non-neuronal cell types along the blood-brain barrier and ventricles enable chemosensing of the blood and cerebrospinal fluid. Tanycytes are a monociliated subtype of ependymal cell that are positioned along the ventral region of the third ventricle and send processes into the hypothalamus[26]. Multiciliated ependymal cells are positioned dorsally along the third ventricle. Tanycytes and multiciliated ependymal cells, as well as astrocytes, emerged as discrete cell types from a common pool of progenitors during the second trimester. Tanycytes were marked by genes involved in chemical sensing and hormonal signaling (GO terms ‘response to organic substance’ and ‘response to lipid’, Supplementary tables 10-12). Multiciliated ependymal cells were marked by genes involved in ‘cilium organization’ and ‘cilium movement’. Genes distinguishing astrocytes included those related to a central function of astrocytes in nutrient and ion transport[27,28].

Oligodendrocyte populations were established by the second trimester. A substantial pool of immature non-myelinating oligodendrocytes was marked by genes expressed in progenitor populations that later turned off in mature oligodendrocytes. Notably, only a few genes had isolated expression in the immature population. These include the canonical marker *PDGFRA*, a growth factor essential for oligodendrogenesis and myelin formation[29]. Additionally, a number of genes were expressed in progenitor populations and turn off in mature oligodendrocytes, including *FABP7*, *SCG3* and *PTPRZ1*, all of which are brain-specific and involved in regulating cell growth. Mature oligodendrocyte populations expressed many genes involved in myelination, including *MBP*, *CNP* and *NFASC*, as well as the epigenetic remodeling protein *SIRT2*.

Development lineages involved in the specification of hypothalamic neurons and glia.

Little is known about the transcriptional regulation of neuronal and glial development in the human hypothalamus, though key regulator transcription factors (TFs) have begun to be defined in the mouse[6,30]. We reconstructed gene co-expression networks and gene regulatory networks from our data using k-means clustering and GENIE3[31], respectively, to begin to unravel these human-specific aspects of hypothalamic development. These analyses revealed 54 gene co-expression modules, together with predictions for each module's key regulator TFs (Supplementary table 13: TF enrichment per module). These included modules expressed specifically in nearly all cell types (Extended Figure 1). In addition, many modules defined lineages from specific subsets of dividing progenitors to mature neurons and glia (Figure 4a,b, Supplementary table 14: Top TF per lineage). We analyzed these lineage-defining modules in detail to gain insight into developmental trajectories.

We identified two major neuronal lineages. A lineage marked by increasing expression of module M76 began in a subset of dividing progenitors, continued in radial glia, and finished in mature neurons. Excitatory vs. inhibitory branches of the M76 lineage were marked by M83/M16 and by M8, respectively. A second neuronal lineage, marked by M11, also began in dividing progenitors, yet this lineage bypassed radial glia populations, making a more direct path to neuronal precursors, potentially marking a separate developmental pathway. Key regulators for the M76 lineage included *CSRNP3*, *MYT1L*, and *DEAF1*, whereas key regulators for the M11 lineage included *NEUROD1*, *NEUROD6*, *BCL11B*, and *SATB2* (Figure 4a).

Oligodendrocyte development followed a single clear lineage from progenitor populations to mature oligodendrocytes. This lineage split from other dividing progenitor populations very early in maturation, with an unbroken trajectory from progenitors to mature oligodendrocytes most apparent when plotted using a 3 dimensional UMAP (Figure S4a-c, Extended Figure 2: Cell types in 3D, Extended Figure 3: Oligodendrocyte pseudotime in 3D). This lineage is marked by expression of modules M2 and M54, as well as by individual genes that increase or decrease with oligodendrocyte development (Figure S4d-f). Lineage-defining marker genes included the key regulator transcription factors *SOX6* and *SOX10*, as well as *LRRK2* and *PCDH15*.

In contrast to oligodendrocytes, lineages leading to astrocytes, ependymocytes, and tanycytes comingled with neuronal lineages until they began to be resolved at the radial glia stage. Modules specific to mature astrocytes (M62, M4, M100), ependymocytes (M44, M65, M80), and tanycytes (M1, M9, M47) enabled us to investigate the molecular origins of each cell type, including both shared and unique key regulator TFs. Shared TFs included the canonical astroglial marker *SOX9*, as well as *GLIS3* and *TRPS1*. Astrocyte-specific TFs included *ARNTL*,

DBX2, *PRRX1*, *RORA*, *HES5*. Ependymocyte-specific TFs included *NFATC2*, *FOXJ1*, *TP73*, and *RFX3*. Tanycyte-specific TFs included *ATF3*, *RFX2*, and *ZSCAN1.* Many of these TFs were expressed starting early in each lineage, suggesting that the molecular identities of specific glial cell types are established very early in gliogenesis.

Cross-species comparisons reveal human-specific aspects of hypothalamic development

We compared patterns of early neurodevelopment in the human hypothalamus to previously described patterns in the mouse to identify shared principles of hypothalamic development, as well as human-specific features. First, we merged our CS22 data with scRNA-seq from the hypothalamus of E11-E13 prenatal mice[12] to gain insight into this critical period for the specification of nuclei. Homologous clusters of human and mouse neurons from most nuclei could readily be identified by integration of these data in a common UMAP space (Figure 5a,b), including near perfect overlap of ARC, DMH and PVH/SON clusters. An interesting exception were human VMH neurons, which at this stage clustered with mouse neuroprogenitors, suggesting delayed development of the VMH in humans.

A more detailed comparison of neuronal subtypes was possible for the arcuate nucleus, which has conserved functions in hunger and satiety. For this analysis, we compared our second trimester arcuate neurons to scRNA-seq of developing mouse arcuate neurons from E10 - P45 (Figure 5c)[12]. Clustering was bifurcated by excitatory/inhibitory class, then largely driven by neuron type. About half of the clusters were reciprocal top hits and consist of specific neuron types; ***OTP+NPY*** (Arc_4 / cluster 14), ***SST*** (Arc_5 / cluster 17), ***GHRH*** (Arc_9/ cluster 9), and ***B2M*** (Arc_12, cluster 12) (Supplementary Table 15: Reciprocal top hits between mouse and human).

Despite this overall similarity, we identified hundreds of gene expression differences between species, including many TFs that were utilized in a regional or developmental species-specific manner (Supplementary Table 16 - E11_E13_CS22_DEGs_by_region.xlsx). For example, the three members of the ONECUT family of TFs were expressed in different combinations in mouse vs. human hypothalamus. In mouse, *Onecut2* is highly expressed throughout development into adulthood, whereas *Onecut1* and *Onecut3* turn on later with low expression. In contrast, all three ONECUT homologs were expressed early and with similar expression patterns across human development, suggesting an expansion of the use of this family of TFs in human hypothalamic development.

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 vs. 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 individual at GW18. Co-embedding these samples in a shared UMAP space enabled direct comparison among the lineages (Figure 6a, Figures S5a,b). Radial glia and early neuroprogenitors from all three germinal zones co-clustered in UMAP space (Clusters 14, 12, 18 ,23 of Figure S5b). Later stage neuroprogenitors diverged toward distinct populations of mature excitatory and inhibitory neurons. 3D UMAP plots (Extended Figure 4) resolve diverging lineages, which overlapped spatially in the 2D UMAP plots.

Despite the overall similarities among radial glia and early neuroprogenitor populations across germinal zones, region-specific gene expression differences were apparent and may provide insight into unique regulatory programs that give rise to specialized neuronal populations. Region-specific signatures of radial glia populations (cluster 16) included genes that are active in developmental or differentiation processes, such as *ETV5* and *MDFI* in the cortex, *WWTR1* and *FGFR2* in the GE, or *GABBR2* and *MT1F* in hypothalamus (Figure S6b, Supplementary Table 17). Region-specific gene expression differences among early neuroprogenitors (cluster 14) included canonical cortex progenitor marker genes such as *EOMES* and *NEUROD4*, which were confirmed to be cortex specific and not expressed in the GE nor hypothalamus. Genes with similar expression patterns in cortex vs. hypothalamus included *GPC3*, a glypican which modulates cell signaling during neural pattering and is expressed within the hypothalamus late in mouse embryogenesis[32]. GRNs constructed with GENIE3 (Supplementary Table 18) predict TFs *FOS*, *HES1* and *EGR1* to regulate *GPC3* and all of these TFs had elevated expression in hypothalamic neuron progenitors compared to cortex and GE. Within the ganglionic eminence, *COL1A2*, a collagen, was uniquely expressed in the GE. While these genes are region specific markers within neuroprogenitors, 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[22]. However, in the GE and hypothalamus, *PPP1R17* is instead expressed in radial glia. *PENK* also showed temporal differences where it was enriched in progenitor populations in the cortex, but in the GE it was present in mature neurons and absent in hypothalamus.

The co-embedded neurogenic map retains well-known features of cortical neurogenesis. Excitatory neurons of the cortex separated into two lineages distinct already at the progenitor stage, with marker genes corresponding to neurons that project within the cortex (intratelencephalic; IT) or to brain regions outside of telencephalon (extratelencephalic; ET). ET neurons were further classified as projecting to thalamic (Clusters 26 and 27, Figure S5b) or subthalamic regions (Clusters 8 and 13, Figure S5b)[33]. The lineage toward IT neurons reflected the temporal ordering with which deep layer vs. superficial layer 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 older neurons at the lineage terminus primarily expressed markers of earlier-born deep layer 5/6 neurons such as *RORB*. Inhibitory neurons of the telencephalon have their origin in the ganglionic eminence, where throughout development maturing interneurons migrate from the ganglionic eminence into all layers of the cortex. This relationship is reflected in a lineage from neuroprogenitors and newborn inhibitory neurons in the GE to mature inhibitory neurons in cortex[34].

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 (Figure S5d). Consistent with our lineage-tracing analyses above, multiple hypothalamic neuronal lineages were apparent, including paths from progenitors that bifurcate into distinct inhibitory and excitatory populations (Figures 6a-c, Figures S5e,f). Unexpectedly, both the inhibitory and excitatory neuronal lineages arising from hypothalamic progenitors displayed a greater degree of overlap with GE compared to cortex, despite the fact that the GE does not give rise to excitatory neurons. Closer examination revealed two sub-groups of hypothalamic and GE progenitors. The hypothalamic inhibitory neuron lineage was adjacent to the lineage leading to GE-derived interneurons marked by *MAF*. The hypothalamic excitatory neuron lineage was adjacent to a distinct GE lineage

marked by *SIX3*, *ISL1*, and *EBF1*, which are most likely progenitors for spiny projection neurons and related progeny of the lateral ganglionic eminence. Therefore, the diversification of hypothalamic and GE progenitors may involve shared gene regulatory mechanisms that are distinct from those utilized within germinal zones of the cortex.

Discussion

Here, we have described the development and diversity of neuronal and non-neuronal cell types in the prenatal human hypothalamus. Our analysis greatly expands knowledge about both shared and human-specific aspects of hypothalamic development. Comparisons among matched 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. Then, specific neuronal subtypes differentiate later. This process was well underway by Carnegie Stage 22, just ~7.5 weeks of gestational age, at which time we were able to discern most nuclei by the expression of established markers. Notably, certain nuclei appeared to mature at different rates. For instance, arcuate neurons had a well-established nuclear identity at CS22.

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* has been detected by GW14[1,35,36] and *CRH* and *AVP* as early as GW12[37], although we observe *AVP* by at least GW7. 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 *CBLN2*, *KISS1,* and *PROK2*. The number of neuronal subpopulations detected across all nuclei (38 subtypes) is much lower than detected by scRNA-seq at postnatal timepoints in mice[5–12]. Although we cannot rule out that the number of sequenced cells in our studies was not large enough to obtain a more comprehensive neuronal sub-clustering, it is very likely that the populations we describe here will undergo further diversification and maturation. In this context the sparse expression of neuropeptides in second trimester might be interpreted as still ongoing differentiation into the various peptidergic subtypes.

As noted above, variation in the prenatal environment can have lasting consequences on many hypothalamic functions[13,14,16], 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 with 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 for different kinds of insults.

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[12,30,38–44]. However, hundreds of 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 sub-types are conserved from mice to humans. Nonetheless, results in mice should be interpreted cautiously when specific genes show divergent patterns across species.

Our comparison of forebrain neurogenic niches revealed that hypothalamic lineages leading to excitatory vs. inhibitory neurons differentiate early in neurodevelopment. Surprisingly, however, both of these lineages are more similar to ganglionic eminence lineages than to cortical lineages. These results suggest that the relationships among neurogenic lineages primarily reflect anterior-posterior spatial organization along the ventricle, rather than the neurotransmitters used in the neuronal progeny. In addition, while the organization and function of the hypothalamus is largely conserved across higher vertebrates, the cortex has greatly expanded in humans[45,46]. Therefore, the programs in the hypothalamus and ganglionic eminence may represent an ancestral program, whereas those in the cortex are more evolutionarily novel.

Several limitations should be noted, setting the stage for future work. First, our annotations of nuclei and neuronal subtypes are based on homology to the more thoroughly atlased mouse, so our analysis may under-represent human-specific features. This could be rectified through spatial transcriptomics studies. Second, while we sampled both first and second trimester time points, there is a substantial gap between 7.5 to 18 gestational weeks. For this reason, we annotated samples from each trimester separately, as to not confuse emerging nuclei of first trimester with the more established nuclei of second trimester. As with most human studies, information on the third trimester is also lacking. Additional scRNA-seq data from these time points could improve the resolution to characterization the maturation of nuclei and cell types. Characterization of human-specific patterns of gene expression in the hypothalamus will also enable researchers to more precisely evaluate genes as heritable causes of human diseases. 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 developmental consequences.

Acknowledgments

This work was supported by three awards from the national Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative (R24 MH114788, O. White, PI; R24 MH114815, R. Hertzano and O. White, PIs; U01 MH114825, A.K., PI). We thank Carlo Colantuoni for his assistance with CoGAPS and providing valuable feedback.

Data Availability

Data is available on the NeMO Archive: https://assets.nemoarchive.org/dat-37m82a9

Data can be explored on NeMO Analytics: https://nemoanalytics.org/p?l=ac0aba73&g=npy

Code Availability

https://github.com/brianherb/DevHumanHypothalamus

Author contributions

S.A.A. and B.R.H. designed the study and conducted the analysis. Experiments were performed by A.B. Data analysis was performed by B.R.H and H.J.G. Data deposition into the NeMO Repository and configuration of the NeMO Analytics profile were prepared by B.R.H. This manuscript was prepared by S.A.A. and B.R.H. with input from all authors.

Competing interests

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

Methods

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[21,22].

Dissociation. 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 15mins. Samples were inverted three times and continued incubating for another 15 mins. Next samples were triturated by manually pipetting with a glass pasteur pipette approximately ten times. Dissociated cells were spun down at 300g for 5mins and Papain removed.

scRNA-seq. 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).

Quality filtering and integration. All 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 20 percent of reads mapping to mitochondrial genes. Doublets were detected using scDblFinder and discarded from the analysis. For published datasets, cells passing quality filters defined in publication were used in this study. Unless otherwise noted, the Seurat package v4.0[47] 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. For each integrated dataset, Louvain clustering was performed and top marker genes for each cluster were determined to assist in cell type identification. To further assist cell type identification, a mouse hypothalamus reference dataset was compiled from 6 published studies[5–10] (Supplemental Table 19 – Marker genes from reference datasets). 2,000 cells were selected at random from each study and merged with the human dataset from this study. Integration of all datasets showed strong agreement among mouse samples and confirmed cell type assignments made based on marker genes. Hypothalamic nuclei present in CS22 were identified using canonical markers from mouse and human *in situ* data (Extended Figure 5). Trimester 2 neurons were assigned by sequential sub clustering. Cells were assigned first using well-established markers from mouse studies[3,5–12,24] which cross-reference across multiple datasets. Feature plots were generated for all genes, and genes with localized expression in unassigned cell populations were checked for localized expression on the Allen Brain Atlas[48].

Lineage analysis. Pseudotime for oligodendrocyte lineage in full human dataset and Excitatory/Inhibitory lineages in GW22 samples were calculated using the slingshot program[49] in R. For each lineage, cells belonging to a defined group of clusters based on developmental stages of that cell type, and pseudotime was calculated from the progenitor population to the most mature population.

Cross species cell type comparison. Subgroups of arcuate nucleus cells in human second trimester samples were directly compared to arcuate subgroups of the mouse using Metaneighbor[50].

Gene regulatory network and gene expression module detection. Gene regulatory networks (GRNs) were inferred using the random-forest based GENIE3 algorithm[31]. Prior to calculation, we identified 7425 variable genes across trimester two samples and smoothed gene counts with knn-smoothing (k=5) to improve gene-gene correlation structure. The top 500,000 TF-target gene pairs were retained and GRNs defined per TF. Gene modules were defined by k-means clustering. Starting with 10441 variable genes across trimester 1 and 2, we first grouped genes by k=100, then merged highly correlated modules (Pearson correlation > 0.30) and discarded modules containing fewer than 10 genes, resulting in 54 modules.

Identifying cortical neuron subtypes. Major classes of cortical cells were previously defined[21] and presented in figure S5c. Additional cell phenotyping was performed using the CoGAPS and projectR programs and the DeCoN dataset[33] to distinguish ET neurons projecting to thalamic or subthalamic regions. Briefly, the CoGAPS[51] program was used to perform non-negative matrix factorization (NMF) on the DeCoN dataset, which included sorted callosal, subcerebral and corticothalamic neurons of the neocortex in mouse. CoGAPS identified five patterns in the DeCoN dataset, where three patterns distinguished the sorted populations. We then used the projectR program to project these patterns into our merged GW18 dataset and noted that the callosal pattern was enriched in IT neurons, and ET neurons were split into two populations by subcerebral and corticothalamic patterns as described above.

Figures

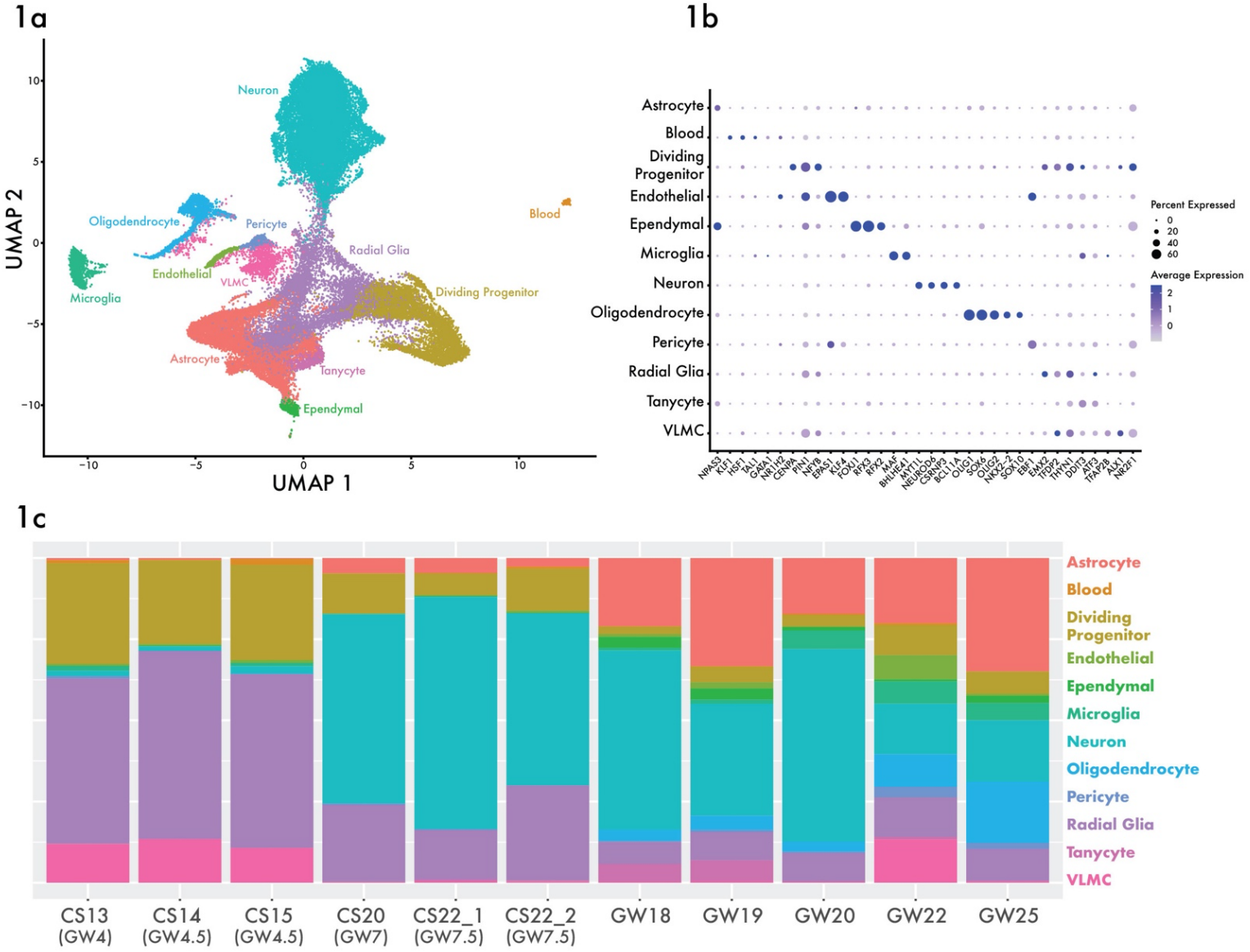


Figure 1. Cell type diversification in the hypothalamus during human development. a) Human samples across trimesters one and two were merged into the same reduced dimensional space, revealing a common progenitor pool and three distinct cell type lineages emerging. Actively dividing progenitors feed into glia, neuronal and other non- neuronal populations. **b)** Top transcription factors marking each cell type. **c)** Neuron populations begin to emerge at CS20 / GW7, with non-neuronal populations further diversifying around GW22.

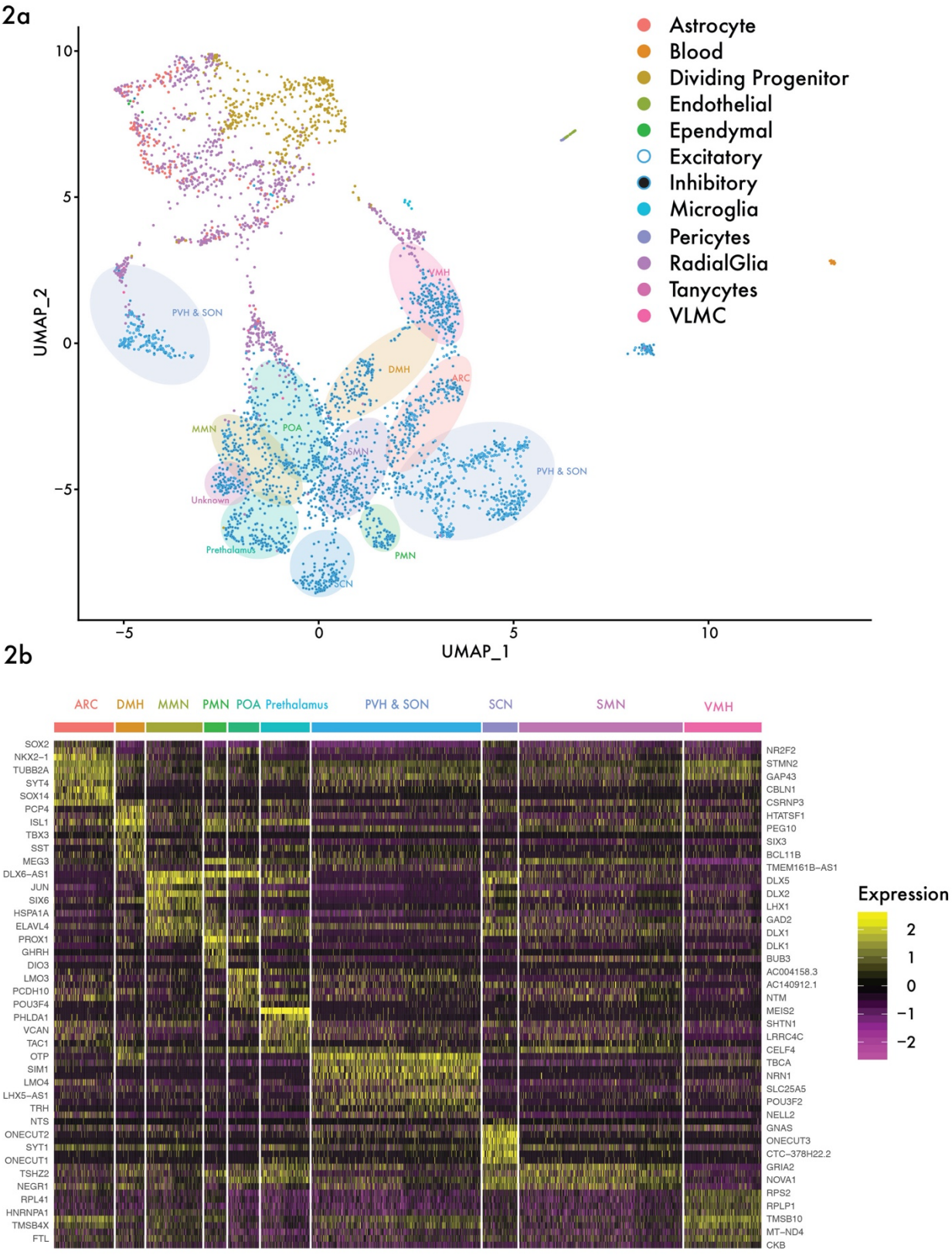
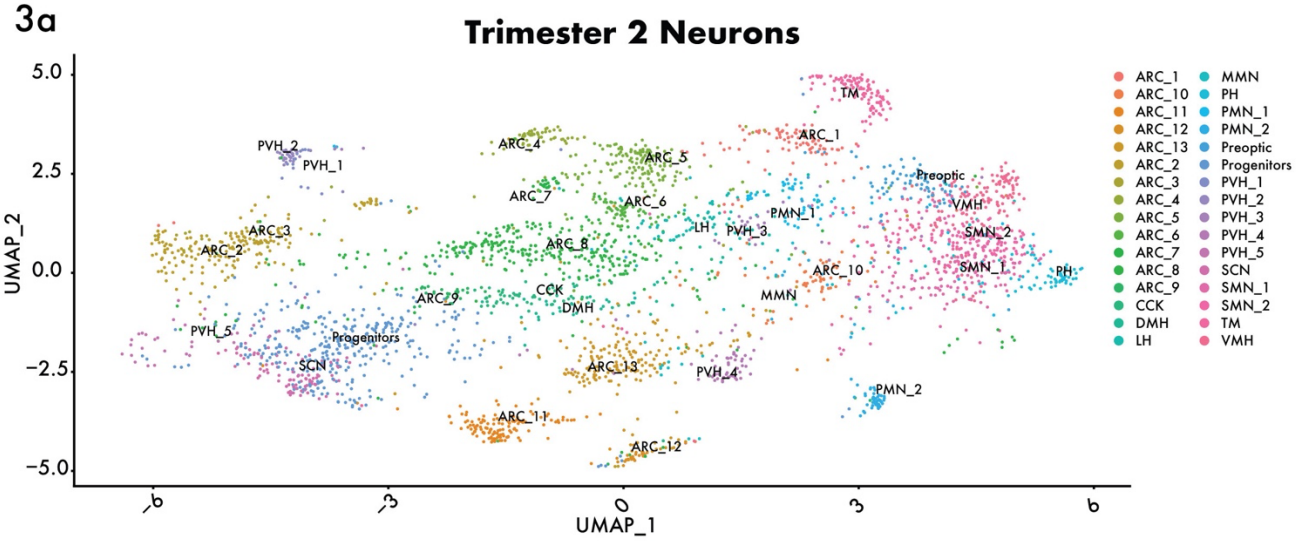
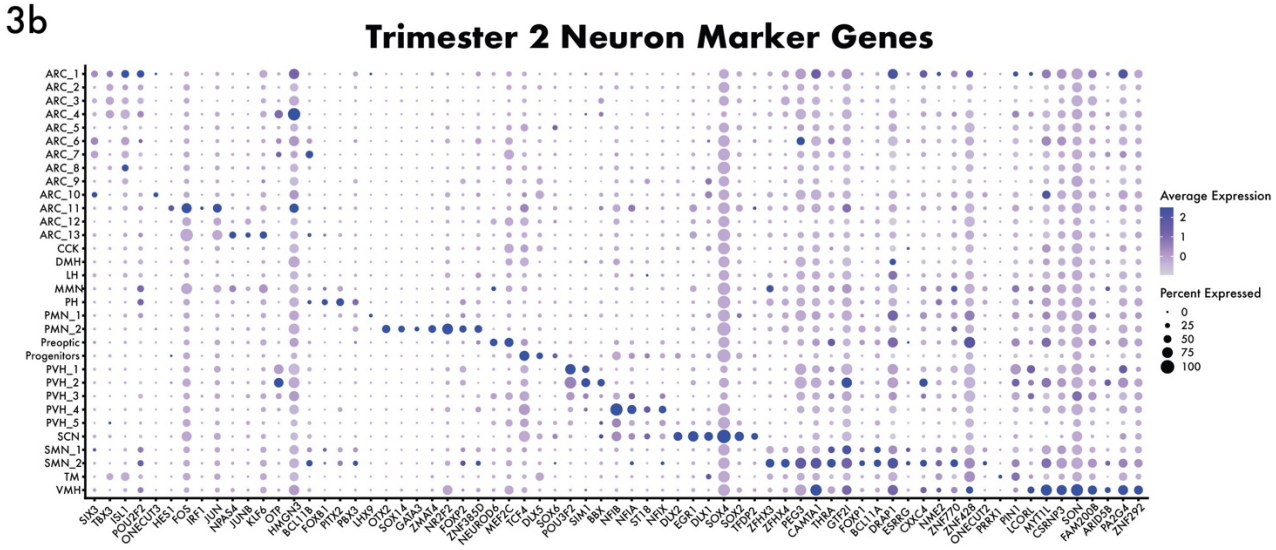


Figure 2. Hypothalamic nuclei are detectable by the end of trimester 1. a) The latest trimester 1 sample, CS22, contains maturing neuron populations that cluster by hypothalamic nuclei. Neuron clusters were assigned to nuclei by canonical markers. **b)** Region-specific marker genes include TFs and neuropeptides.





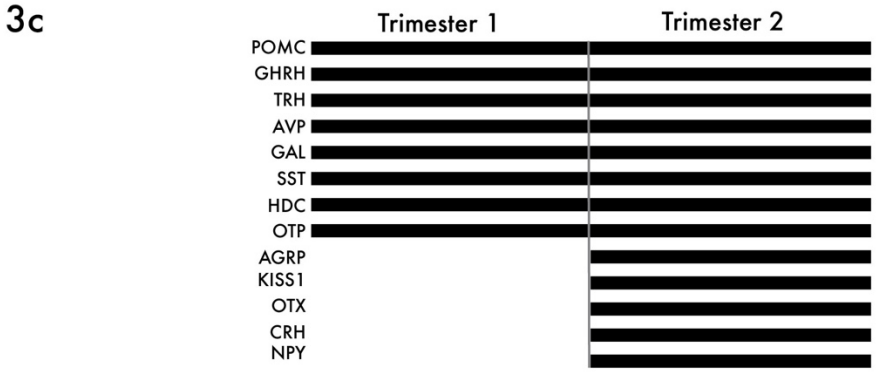


Figure 3. Derivation of neuronal subtypes and the transcription factors that drive their development. a) Neurons from trimester 2 were re-clustered and clusters with a clearly defined nuclei and neuron type are shown. **b)** Dot plot of top enriched transcription factors for each neuron subcluster. **c)** Estimated timing of neuropeptide production across trimester 1 and 2. Genes encoding certain neuropeptides such as *POMC* and *TRH* are detected in first trimester, whereas others like *NPY* and *CRH* are not expressed until second trimester.

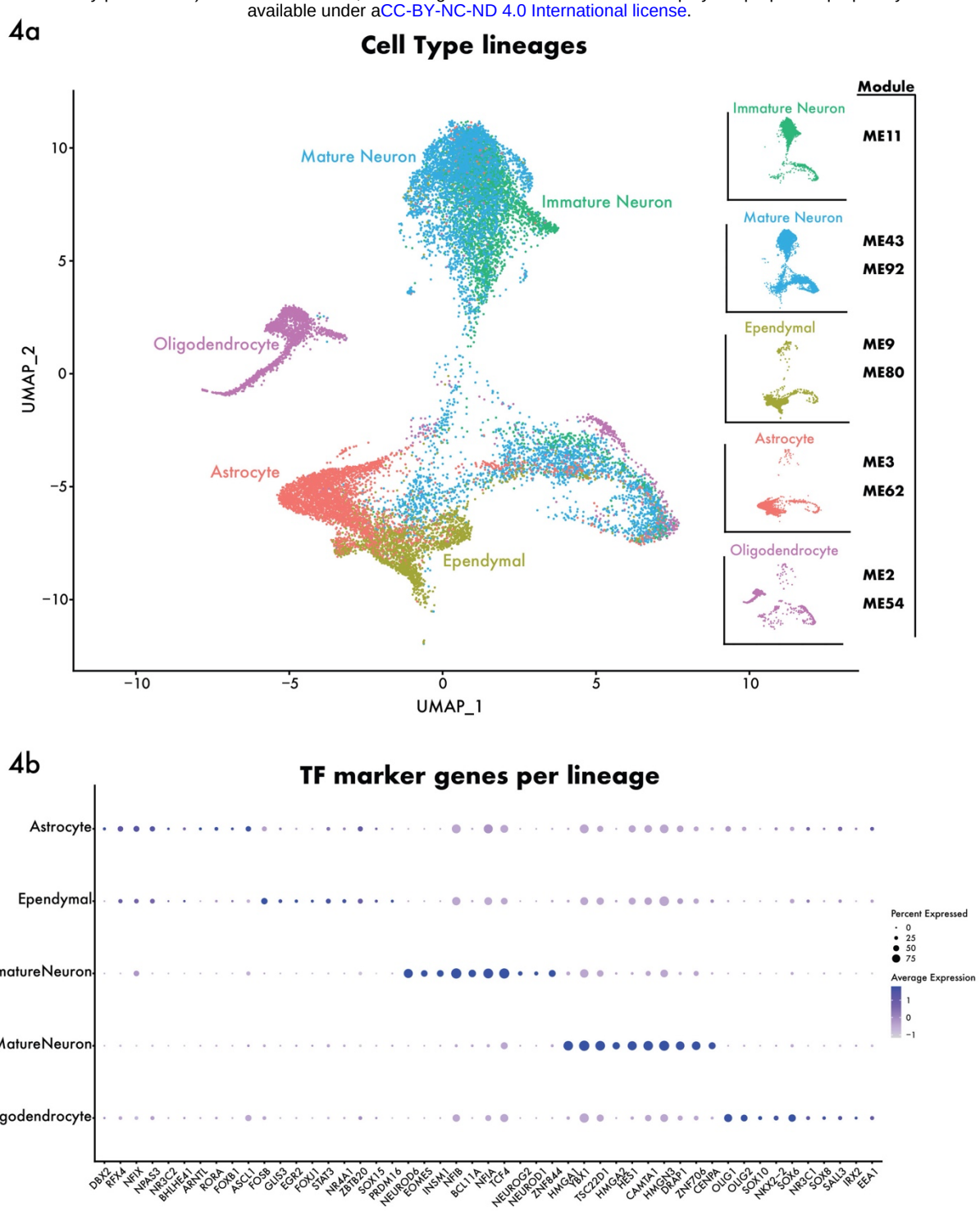
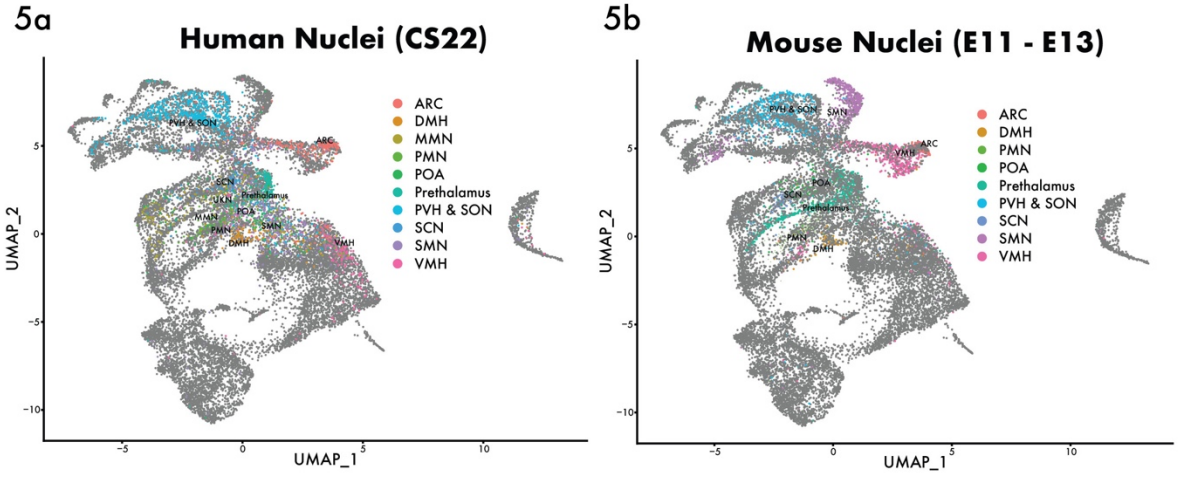


Figure 4. Gene regulatory modules reveal developmental lineages. a) Gene regulatory modules active in both mature neuronal and non-neuronal populations as well as dividing progenitor cells were used to identify developmental lineages. The main figure assigns cells to a module based on highest eigengene value of the modules listed on the right-hand column. Subplots in the right column isolate cells from each lineage. Two distinct neuronal lineages were detected. **b)** Top transcription factor marker genes were found across lineages within the dividing progenitor population, indicating early drivers of neuronal and non-neuronal lineages.



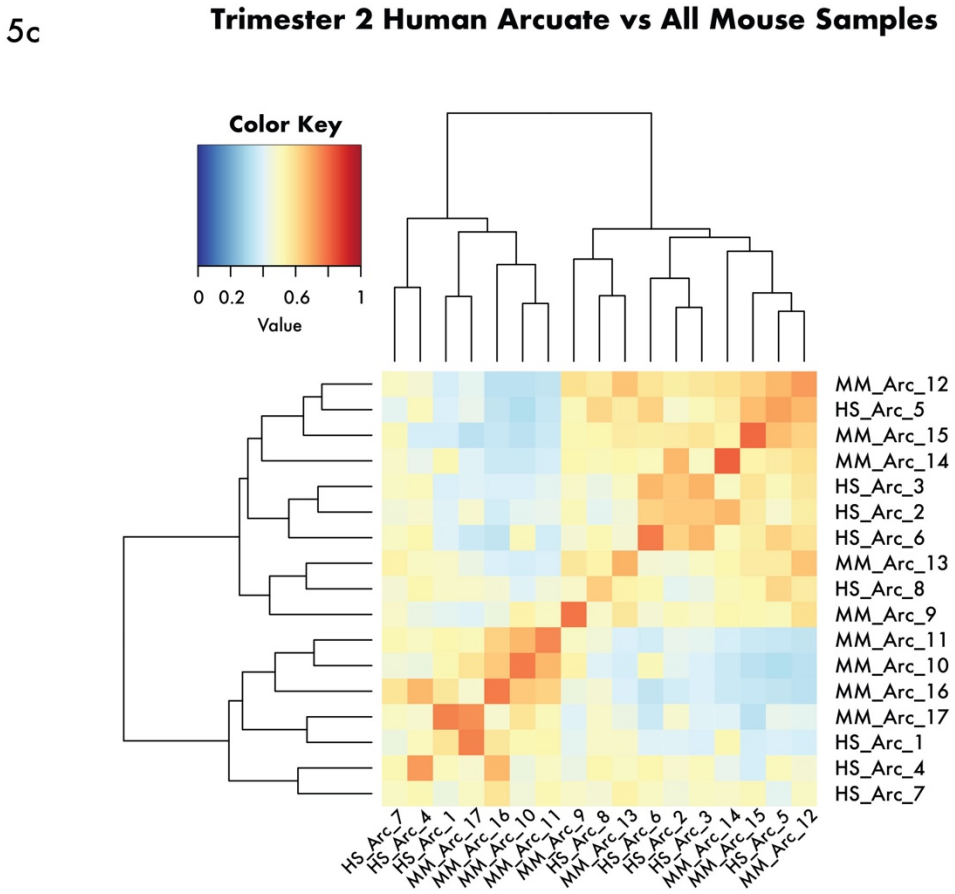


Figure 5. Comparison of human and mouse nuclei across development. We compared our human developmental dataset against a mouse developmental dataset from Kim et al. 2020 that spans E10 - P45. **ab)** Merge of the human CS22 sample with E11 - E13 mouse samples. Human nuclei (a) largely overlap mouse nuclei (b), however VMH, and to a lesser degree SMN exhibit a delayed development compared to mouse. **c)** MetaNeighbor comparison of sub clusters of Arcuate from human trimester 2 and all mouse samples included in Kim et al. 2020. Clustering is bifurcated by excitatory/inhibitory class, then largely driven by neuron type.

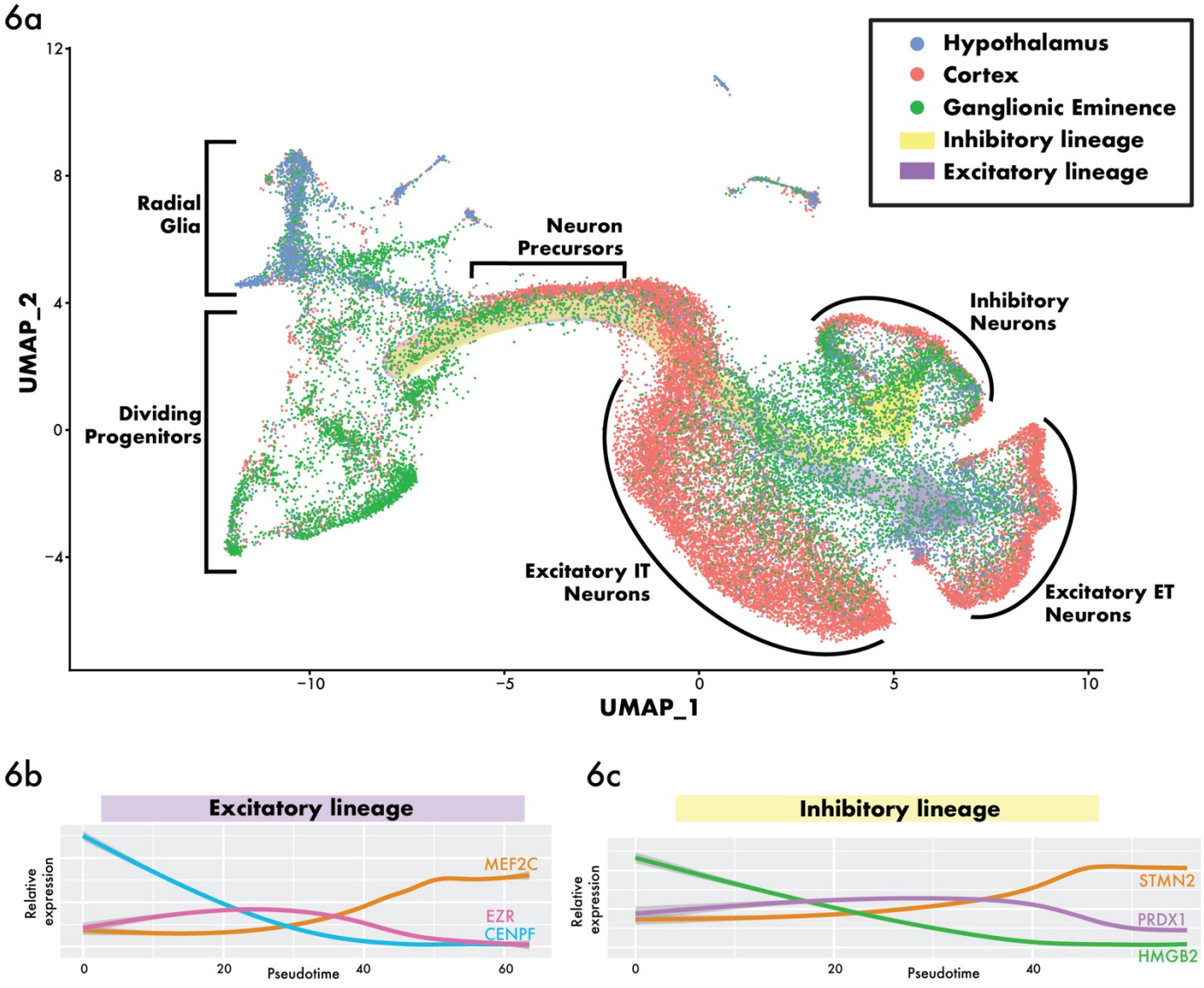


Figure 6. Neurogenesis in hypothalamus shares gene expression profiles with the ganglionic eminence that are distinct from lineages in the cortex. a) While similarities exist in excitatory and inhibitory neurons across the brain, their development may differ across regions. We found that the hypothalamus was more like neuron development in the ganglionic eminence versus the cortex. Cortex neurons were grouped into ET and IT types, and projecting neurons of the hypothalamus and GE cluster proximal to ET neurons of the cortex, suggesting a common gene signature of projection neurons. Excitatory and inhibitory lineages share an origin in neuron precursors and branch later in development. **b,c)** Examples of genes that are expressed in early, middle, and late stages of neuronal development.

Supplementary Figure 1. Sample mixing and cell cycle state within common UMAP space. **a)** Distribution of individual samples within the common UMAP space for trimester 1 and 2 samples. First trimester samples dominate the right arm containing the dividing progenitor population, whereas second trimester samples constitute the majority of mature neuronal and non-neuronal populations. **b)** Cell cycle stage status for each cell.

Supplementary Figure 2. Trimester 1 and 2 cellular phenotypes presented in UMAP space. **a)** Spatial distribution of first trimester samples in common UMAP space. **b)** Major cell types as determined across all samples. **c)** Cell cycle status for each cell. **d)** Hypothalamic nuclei as determined in CS22 samples. **e)** The two CS22 samples were integrated using Seurat and UMAP plot shows samples overlap well. **f)** Neurons clustered but nuclei and their identity were determined using canonical markers. Inhibitory neurons are marked by GAD2 (g) and excitatory neurons by SLC17A6 (h). **i)** Spatial distribution of second trimester samples in common UMAP space. **j)** Major cell types as determined across all samples.

Supplementary Figure 3. Drivers of neuron maturation. a) Hierarchical clustering of neurons identified transcription factors potentially involved in the development of hypothalamic nuclei and neuron differentiation. Right panel shows top expressed transcription factors stratified by neuronal subtype.

Supplementary Figure 4. Oligodendrocyte maturation across trimester 1 and 2. a) 3-dimensional UMAP plotting of all samples reveal a continuous oligodendrocyte lineage that was broken in 2-dimensional UMAP. **b)** Oligodendrocyte lineage pseudo time in 3D space and **c)** 2D space calculated using the slingshot program. **de)** Gene expression module ME2 increases in expression across the entire oligodendrocyte lineage, whereas module ME54 highlights the latest stages of oligodendrocyte maturation. **f)** Examples of genes expressed early or late across the lineage.

Supplementary Figure 5. Regional comparison in sample GW18 revealed distinct neuronal developmental lineages between hypothalamus, ganglionic eminence, and cortex. a) Locations of the nine individual brain region samples in common UMAP space for GW18. Hypothalamus overlapped most with GE, whereas Cortex samples clustered together, especially in mature populations. **b)** Louvain clusters of all samples. Cluster 14 consisted of a mix of all regions, marking an early population of neural progenitors. Maturation progresses from left to right across plot. **c)** Cell types identified in Eze et al. 2020. Excitatory neurons form groups of IT and ET neurons, and IT neurons are stratified by cortical layer. **d)** Hypothalamic cell types as identified in this study. **e)** Excitatory lineage and **f)** inhibitory lineage calculated by the slingshot program.

Extended Figure 1. Expression of 54 gene expression modules across all samples. For each gene expression module, eigengene values were plotted from yellow (low values) to red (high values).

Extended Figures 2-3. UMAP plotting of all samples in 3-dimensional space. Cell types identified across all samples (Fig E2) and oligodendrocyte lineage (Fig E3) are presented in 3-dimensional UMAP space.

Extended Figure 4. UMAP plotting of all GW18 brain regions in 3-dimensional space. Cell clusters identified across all samples are presented in 3-dimensional UMAP space. Note that cluster 10, which contains the majority of hypothalamus and GE neural progenitors, is a continuous cluster that begins at the dividing progenitors and ends at the start of maturing neuronal populations.

Extended Figure 5. In situ data from Allen brain atlas. Marker genes used in the identification of human hypothalamic nuclei were checked for spatial expression across the adult mouse hypothalamus. In situ data is presented alongside corresponding brain atlas diagrams.

Supplementary Table 1. Cells included in study per sample. Counts of high-quality cells per sample after filtering.

Supplementary Table 2. Gene expression differences across major cell types. Differentially expressed genes across major cell types in trimesters 1 and 2 were calculated using the Seurat program. Table includes a dedicated sheet per cell type and positive fold changes indicate higher expression in that cell type compared to all other cell types combined.

Supplementary Tables 3-5. GO analysis of major cell type markers across all samples. GO category enrichments were calculated for upregulated genes in each major cell type. Tables are organized by biological process (Table S3), molecular function (Table S4), and cellular compartment (Table S5). Each table has a dedicated sheet per cell type.

Supplementary Table 6. Gene expression differences across hypothalamic nuclei in sample CS22. Differentially expressed genes across hypothalamic nuclei in sample CS22 were calculated using the Seurat program. Nuclei are indicated in the 'cluster' column and positive fold changes indicate higher expression in that nuclei compared to all other nuclei combined.

Supplementary Tables 7-9. GO analysis of hypothalamic nuclei in sample CS22. GO category enrichments were calculated for upregulated genes in each hypothalamic nucleus in sample CS22. Tables are organized by biological process (Table S7), molecular function (Table S8), and cellular compartment (Table S9). Each table has a dedicated sheet per nucleus.

Supplementary Tables 10-12. GO analysis of major cell type markers across trimester 2 samples. GO category enrichments were calculated for upregulated genes in each major cell type within trimester 2. Tables are organized by biological process (Table S10), molecular function (Table S11), and cellular compartment (Table S12). Each table has a dedicated sheet per cell type.

Supplementary Table 13. Gene regulatory networks enriched in gene expression modules. Hypergeometric tests were used to identify enrichment of GRNs within the 54 gene expression modules found across all samples.

Supplementary Table 14. Top TF markers of early developmental lineages. Gene regulatory modules were used to identify 5 lineages among dividing progenitors, leading to neuronal and glial populations. The Seurat program was used to identify differentially expressed TFs per

lineage among the dividing progenitor population. Lineages are indicated in the 'cluster' column and positive fold changes indicate higher expression in that lineage compared to all other lineages combined.

Supplementary Table 15. MetaNeighbor reciprocal top matches between human and mouse arcuate neuron clusters. The MetaNeighbor program was used to calculate correlations between the 9 clusters of human arcuate neurons and the 9 clusters of mouse arcuate neurons identified in Kim et al. 2020. Clusters with highest correlations are presented in table.

Supplementary Table 16. Cross-species gene expression differences in the developing hypothalamus. Early (E11-E13) hypothalamic nuclei in developing mice were compared to human nuclei from late trimester 1 (CS22). Table includes a dedicated sheet per nucleus and positive fold changes indicate higher expression in the human sample.

Supplementary Table 17. Gene expression differences across multiple brain regions in GW18 sample. Hypothalamus, ganglionic eminence, and cortex were compared within the same early trimester 2 sample. Cells were merged across regions and differential gene expression was calculated within clusters. Table includes a dedicated sheet per cell cluster and positive fold changes indicate higher expression for brain region noted in 'cluster' column.

Supplementary Table 18. Gene regulatory networks calculated by GENIE3 in Trimester 2. Gene regulatory networks (GRNs) were inferred using the random-forest based GENIE3 algorithm[31]. Prior to calculation, we identified 7425 variable genes across trimester two samples and smoothed gene counts with knn-smoothing (k=5) to improve gene-gene correlation structure. The top 500,000 TF-target gene pairs were retained and GRNs defined per TF. Each sheet contains one GRN and its target genes.

Supplemental Table 19 – Marker genes from reference datasets. Top marker genes for published neuron clusters were recalculated using 5000 most variable genes per dataset. Only neurons were considered in this analysis.

References Cited

1. Swaab, D. F. Development of the human hypothalamus. *Neurochemical Research* **20**, (1995).
2. Xie, Y. & Dorsky, R. I. Development of the hypothalamus: Conservation, modification and innovation. *Development (Cambridge)* vol. 144 1588–1599 (2017).
3. Bedont, J. L., Newman, E. A. & Blackshaw, S. Patterning, specification, and differentiation in the developing hypothalamus. *Wiley Interdisciplinary Reviews: Developmental Biology* vol. 4 445–468 (2015).
4. Saper, C. B. & Lowell, B. B. The hypothalamus. *Current Biology* vol. 24 R1111–R1116 (2014).
5. Campbell, J. N. *et al.* A molecular census of arcuate hypothalamus and median eminence cell types. *Nature Neuroscience* **20**, 484–496 (2017).
6. Chen, R., Wu, X., Jiang, L. & Zhang, Y. Single-Cell RNA-Seq Reveals Hypothalamic Cell Diversity. *Cell Reports* **18**, 3227–3241 (2017).

7. Moffitt, J. R. *et al.* Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* **362**, (2018).
8. Kim, D. W. *et al.* Multimodal Analysis of Cell Types in a Hypothalamic Node Controlling Social Behavior. *Cell* **179**, 713-728.e17 (2019).
9. Mickelsen, L. E. *et al.* Single-cell transcriptomic analysis of the lateral hypothalamic area reveals molecularly distinct populations of inhibitory and excitatory neurons. *Nature Neuroscience* **22**, 642–656 (2019).
10. Wen, S. *et al.* Spatiotemporal single-cell analysis of gene expression in the mouse suprachiasmatic nucleus. *Nature Neuroscience* **23**, 456–467 (2020).
11. Romanov, R. A. *et al.* Molecular design of hypothalamus development. *Nature* **582**, 246–252 (2020).
12. Kim, D. W. *et al.* The cellular and molecular landscape of hypothalamic patterning and differentiation from embryonic to late postnatal development. *Nature Communications* **11**, (2020).
13. Chang, G. Q., Gaysinskaya, V., Karatayev, O. & Leibowitz, S. F. 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).
14. Franke, K. *et al.* "Programming" of orexigenic and anorexigenic hypothalamic neurons in offspring of treated and untreated diabetic mother rats. *Brain Research* **1031**, 276–283 (2005).
15. Nishi, M. Effects of Early-Life Stress on the Brain and Behaviors: Implications of Early Maternal Separation in Rodents. *International journal of molecular sciences* vol. 21 (2020).
16. Inoue, F. *et al.* Genomic and epigenomic mapping of leptin-responsive neuronal populations involved in body weight regulation. *Nature Metabolism* **1**, 475–484 (2019).
17. Tessmar-Raible, K. *et al.* Conserved Sensory-Neurosecretory Cell Types in Annelid and Fish Forebrain: Insights into Hypothalamus Evolution. *Cell* **129**, 1389–1400 (2007).
18. Zhou, X. *et al.* Cellular and molecular properties of neural progenitors in the developing mammalian hypothalamus. *Nature Communications* **11**, (2020).
19. Nowakowski, T. J. *et al.* Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* **358**, (2017).
20. Pollen, A. A. *et al.* Establishing Cerebral Organoids as Models of Human-Specific Brain Evolution. *Cell* **176**, 743-756.e17 (2019).
21. Bhaduri, A. *et al.* Cell stress in cortical organoids impairs molecular subtype specification. *Nature* **578**, (2020).
22. Eze, U. C., Bhaduri, A., Haeussler, M., Nowakowski, T. J. & Kriegstein, A. R. Single-cell atlas of early human brain development highlights heterogeneity of human neuroepithelial cells and early radial glia. *Nature Neuroscience* **24**, 584–594 (2021).
23. Koutcherov, Y., Mai, J. K., Ashwell, K. W. S. & Paxinos, G. Organization of human hypothalamus in fetal development. *Journal of Comparative Neurology* **446**, 301–324 (2002).
24. Shimogori, T. *et al.* A genomic atlas of mouse hypothalamic development. *Nature Neuroscience* **13**, 767–775 (2010).

25. Padilla, S. L., Carmody, J. S. & Zeltser, L. M. Pomc-expressing progenitors give rise to antagonistic neuronal populations in hypothalamic feeding circuits. *Nature Medicine* **16**, (2010).
26. Prevot, V. *et al.* The versatile tanycyte: A hypothalamic integrator of reproduction and energy metabolism. *Endocrine Reviews* **39**, 333–368 (2018).
27. Pfrieger, F. W. & Ungerer, N. Cholesterol metabolism in neurons and astrocytes. *Progress in Lipid Research* vol. 50 357–371 (2011).
28. McCarthy, M. M., Amateau, S. K. & Mong, J. A. Steroid modulation of astrocytes in the neonatal brain: Implications for adult reproductive function. *Biology of Reproduction* vol. 67 691–698 (2002).
29. Funa, K. & Sasahara, M. The roles of PDGF in development and during neurogenesis in the normal and diseased nervous system. *Journal of Neuroimmune Pharmacology* vol. 9 168–181 (2014).
30. Kim, D. W. *et al.* Gene regulatory networks controlling differentiation, survival, and diversification of hypothalamic Lhx6-expressing GABAergic neurons. *Communications Biology* **4**, (2021).
31. Huynh-Thu, V. A., Irrthum, A., Wehenkel, L. & Geurts, P. Inferring Regulatory Networks from Expression Data Using Tree-Based Methods. *PLoS ONE* **5**, (2010).
32. Luxardi, G. *et al.* Glypicans are differentially expressed during patterning and neurogenesis of early mouse brain. *Biochemical and Biophysical Research Communications* **352**, 55–60 (2007).
33. Molyneaux, B. J. *et al.* DeCoN: Genome-wide analysis of invivo transcriptional dynamics during pyramidal neuron fate selection in neocortex. *Neuron* **85**, 275–288 (2015).
34. Mayer, C. *et al.* Developmental diversification of cortical inhibitory interneurons. *Nature* **555**, 457–462 (2018).
35. Schubert, F., George, J. M. & Bhaskar Rao, M. Vasopressin and oxytocin content of human fetal brain at different stages of gestation. *Brain Research* **213**, (1981).
36. Mengod, G., Goudsmit, E., Probst, A. & Palacios, J. M. Chapter 4 In situ hybridization histochemistry in the human hypothalamus. in (1992). doi:10.1016/S0079-6123(08)64562-8.
37. MASTORAKOS, G. & ILIAS, I. Maternal and Fetal Hypothalamic-Pituitary-Adrenal Axes During Pregnancy and Postpartum. *Annals of the New York Academy of Sciences* **997**, (2003).
38. Orquera, D. P. *et al.* 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).
39. Nasif, S. *et al.* Islet 1 specifies the identity of hypothalamic melanocortin neurons and is critical for normal food intake and adiposity in adulthood. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E1861–E1870 (2015).
40. de Souza, F. S. J. *et al.* Identification of Neuronal Enhancers of the Proopiomelanocortin Gene by Transgenic Mouse Analysis and Phylogenetic Footprinting. *Molecular and Cellular Biology* **25**, 3076–3086 (2005).
41. Sanz, E. *et al.* Fertility-regulating kiss1 neurons arise from hypothalamic pomc-expressing progenitors. *Journal of Neuroscience* **35**, 5549–5556 (2015).
42. Li, M. M. *et al.* The Paraventricular Hypothalamus Regulates Satiety and Prevents Obesity via Two Genetically Distinct Circuits. *Neuron* **102**, 653-667.e6 (2019).

43. Alvarez-Bolado, G., Paul, F. A. & Blaess, S. Sonic hedgehog lineage in the mouse hypothalamus: From progenitor domains to hypothalamic regions. *Neural Development* **7**, (2012).
44. Morales-Delgado, N. *et al.* Regionalized differentiation of CRH, TRH, and GHRH peptidergic neurons in the mouse hypothalamus. *Brain Structure and Function* **219**, 1083–1111 (2014).
45. Smaers, J. B., Gómez-Robles, A., Parks, A. N. & Sherwood, C. C. Exceptional Evolutionary Expansion of Prefrontal Cortex in Great Apes and Humans. *Current Biology* **27**, 714–720 (2017).
46. Hill, J. *et al.* Similar patterns of cortical expansion during human development and evolution. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 13135–13140 (2010).
47. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, (2021).
48. Lein, E. S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, (2007).
49. van den Berge, K. *et al.* Trajectory-based differential expression analysis for single-cell sequencing data. *Nature Communications* **11**, (2020).
50. Crow, M., Paul, A., Ballouz, S., Huang, Z. J. & Gillis, J. Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. *Nature Communications* **9**, (2018).
51. Fertig, E. J., Ding, J., Favorov, A. v., Parmigiani, G. & Ochs, M. F. CoGAPS: an R/C++ package to identify patterns and biological process activity in transcriptomic data. *Bioinformatics* **26**, (2010).