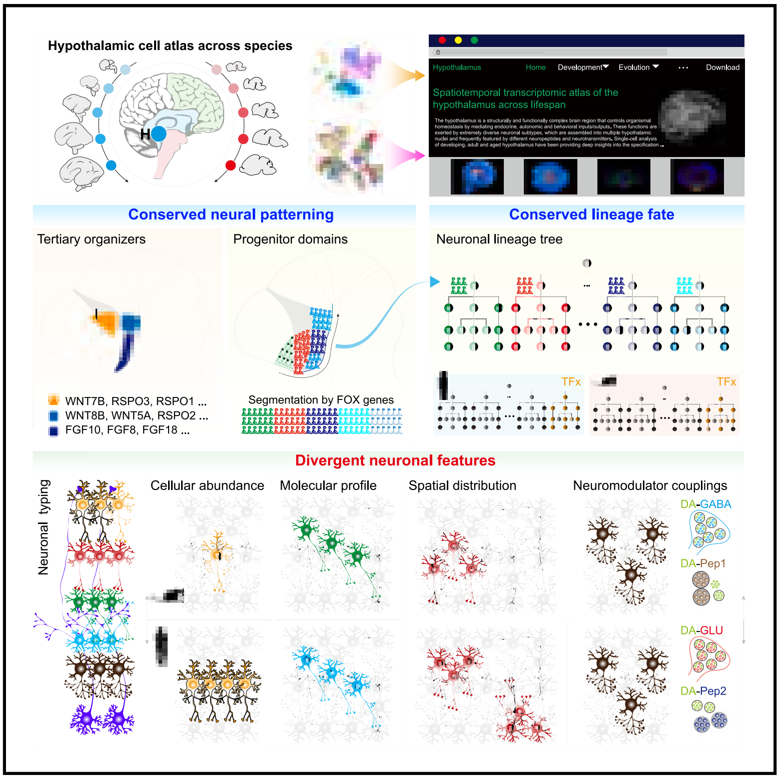
Resource

Developmental Cell

Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development

Graphical abstract



Spatiotemporal transcriptomic atlas of the hypothalamus across lifespan

The hypothalamus is a structurally and functionally complex brain region that controls organismal homeostasis by mediating endocrine, autonomic and behavioral inputs/outputs. These functions are exerted by extremely diverse neuronal subtypes, which are assembled into multiple hypothalamic nuclei and frequently featured by different neuropeptides and neurotransmitters. Single-cell analysis of developing, adult and aged hypothalamus have been providing deep insights into the specification ...

Highlights

- Neural patterning of the developing hypothalamus is evolutionarily conserved
- Neurogenic lineages are shared between human and mouse hypothalamus
- Human neurons feature a specific subtype and enhanced neuromodulation
- Neuronal distribution and neurochemical traits can vary across species

Authors

Zhen-Hua Chen, Taotao Bruce Pan, Yu-Hong Zhang, ..., Chun-Xia Yi, Longqi Liu, Qing-Feng Wu

Correspondence

liulongqi@genomics.cn (L.L.),
wu_qingfeng@genetics.ac.cn (Q.-F.W.)

In brief

Chen et al. demonstrate conserved neural patterning mechanisms across mammalian hypothalamus development, reconstruct a neurogenic lineage tree, and identify four adaptive evolutionary divergences in human developing neurons: a human-enriched subtype, enhanced neuromodulation, redistributed neuroendocrine neurons, and reconfigured neurochemistry in hypothalamic dopamine neurons.

Chen et al., 2025, Developmental Cell *60*, 1–15
July 7, 2025
https://doi.org/10.1016/j.devcel.2025.03.009

CellPress

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

Developmental Cell

CellPress

Resource

Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development

Zhen-Hua Chen,[1,2,15] Taotao Bruce Pan,[3,15] Yu-Hong Zhang,[1,4,15] Ben Wang,[5,15] Xue-Lian Sun,[1,2,15] Meixi Gao,[3,15] Yang Sun,[3] Mingrui Xu,[1,2] Shuhui Han,[3] Xiang Shi,[1,2] Felipe Correa-da-Silva,[6,7] Chenlu Yang,[3] Junfu Guo,[3,8] Haoda Wu,[1,2] Yu Zheng Li,[1] Xiu-Qin Liu,[9] Fei Gao,[10] Zhiheng Xu,[1,2] Shengjin Xu,[11] Xin Liu,[3] Ying Zhu,[12] Ziqing Deng,[3] Shiping Liu,[13] Yi Zhou,[11] Chun-Xia Yi,[6,7] Longqi Liu,[13,*] and Qing-Feng Wu[1,2,14,16,*]

[1]State Key Laboratory of Molecular Development Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China
[2]University of Chinese Academy of Sciences, Beijing 100101, China
[3]BGI Research, Beijing 102601, China
[4]Affiliated Brain Hospital of Guangzhou Medical University, Guangzhou 511436, China
[5]Department of Obstetrics and Gynecology, The Third Affiliated Hospital of Chongqing Medical University, Chongqing 401120, China
[6]Department of Endocrinology and Metabolism, Amsterdam University Medical Center, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, the Netherlands
[7]Netherlands Institute for Neuroscience, Meibergdreef 47, 1105 BA Amsterdam, the Netherlands
[8]BGI Research, Shenzhen 518083, China
[9]Department of Obstetrics and Gynecology, Baoding Second Central Hospital, Baoding 072750, China
[10]State Key Laboratory of Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China
[11]Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China
[12]Institutes of Brain Science, State Key Laboratory of Medical Neurobiology, Fudan University Shanghai, Shanghai 200032, China
[13]BGI Research, Hangzhou 310030, China
[14]Beijing Key Laboratory for Genetics of Birth Defects, Beijing 100045, China
[15]These authors contributed equally
[16]Lead contact
*Correspondence: liulongqi@genomics.cn (L.L.), wu_qingfeng@genetics.ac.cn (Q.-F.W.)

https://doi.org/10.1016/j.devcel.2025.03.009

SUMMARY

The hypothalamus, an “ancient” subcortical brain structure, maintains physiological homeostasis and controls native behaviors. The evolution of homeostatic regulation and behavioral control in mammals may rely on adaptable neuronal identity establishment but conserved neural patterning mechanisms during neurodevelopment. Here, we combined single-cell, single-nucleus, and spatial transcriptomic datasets to map the spatial patterning of diverse progenitor domains and reconstruct their neurogenic lineages in the developing human and mouse hypothalamus. While the regional organizers orchestrating neural patterning are conserved between primates and rodents, we identified a human-enriched neuronal subtype and found a substantial increase in neuromodulatory gene expression among human neurons. Furthermore, cross-species comparison demonstrated a potential redistribution of two neuroendocrine neuronal subtypes and a shift in inter-transmitter and transmitter-peptide coupling within hypothalamic dopamine neurons. Together, our study lays a critical foundation for understanding cellular development and evolution of the mammalian hypothalamus.

INTRODUCTION

The mammalian brain develops in an intricately orchestrated sequence of stages, including neural patterning, neurogenesis, neuritogenesis, synaptogenesis, and gliogenesis. While the anatomic and cellular architecture of fully developed brains is generally conserved across different mammalian species, there exist numerous species-specific differences between rodent and primate brains, such as variation in organ size and cellular complexity. Recently, single-cell comparative genomic studies have been conducted to elucidate the species difference in molecular design and cellular hierarchy of the developing neocortex.[1–4] However, the interspecies divergence of subcortical brain development remains largely enigmatic.

The hypothalamus is a structurally and functionally complex brain region that controls organismal homeostasis by mediating

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

endocrine, autonomic, and behavioral inputs/outputs, representing an ideal subcortical structure with which to address questions about cellular evolution in rodents and primates. In the hypothalamus, diverse neuronal subtypes are frequently featured by different neuropeptides and neurotransmitters.[5,6] Their identity, number, location, and connectivity are established during embryonic development, which lays the structural foundation for various hypothalamic functions.[6–9] Across the mouse developmental continuum, hypothalamic radial glial cells (RGCs), intermediate progenitor cells (IPCs), and neuronal precursors adopt a stepwise diversification strategy and precisely controlled lineage programs to generate neuronal diversity.[10,11] Developmental abnormalities of specific neural progenitors, neuronal subtypes, or circuits in the hypothalamus lead to metabolic disorders, reproductive defects, sleep disorders, growth retardation, anxiety, depression, and autism.[8,12] Although the hypothalamus has traditionally been considered as an ''ancient'' and evolutionarily conserved region of the mammalian brain,[5,13] bulk transcriptomic analysis recently implied a potential divergence between adult human and mouse hypothalamus.[14] This raises the question regarding the extent of conservation and divergence in the developmental programming of the mammalian hypothalamus.

Precise spatiotemporal regulation of gene expression is critical for all aspects of mammalian brain development, evolution, and function. Given the extraordinary progenitor and neuronal diversity in the subcortical brain structures, single-cell and spatially resolved transcriptomics have emerged as effective tools in understanding cellular development and evolution of the hypothalamus.[10,15,16] Here, we performed single-cell and single-nucleus RNA sequencing of the developing primate hypothalamus, found conservation in the spatial patterning of hypothalamic progenitor domains, computationally reconstructed various neuronal sublineages, and revealed unique features in human hypothalamic neurons compared with mice. We have designed an interactive platform that enables data exploration, which is accessible at https://hypoatlas.org (Video S1).

RESULTS

Conserved cell types with divergent molecular features between humans and mice

To better understand the cellular repertoire and developmental trajectory of the mammalian hypothalamus, we performed single-cell transcriptomic profiling of the human and macaque hypothalamus spanning different developmental stages (Figures 1A, 1B, and S1A), incorporating recent datasets from developing mouse studies.[10,17] Human samples were collected from 13 fetuses ranging from 5 to 20 weeks post-conception (PCW5–20), while macaque fetal specimens were dissected at three developmental time points (PCW5, 8, and 11). After stringent quality control, batch correction, and data integration, we collected 351,742 hypothalamic cells across three species and multiple stages during development, including 195,166 cells from humans (Figures 1C and S1B–S1H). Clustering analysis of the integrated and species-specific datasets using the Louvain algorithm identified 14 principal cell types, each characterized by specific expression of well-known marker genes (Figures 1C, S1I–S1K, S2A, and S2B; Table S1A): 3 neural progenitor types (RGCs and IPCs), 2 major neuronal types (glutamatergic and GABAergic neurons), 6 glial cell types (astrocytes, oligodendrocyte precursor cells, oligodendrocytes, ependymocytes, tanycytes, and microglia), and 3 non-neural cell types (vascular leptomeningeal cells, endothelial cells, and blood cells). We further validated the consistency and reliability of our results using publicly available developing and adult hypothalamus datasets (Figures S2C–S2J; Table S1A). Sex differences in these cell types were largely minimal at prenatal and perinatal stages (Figures S2K–S2O; Table S1B).

To examine cellular homology among different species, we measured transcriptomic similarity across species using two independent methods and found a robust cross-species conservation, irrespective of sex (Figures 1D and S3A–S3C). Among these homologous cell types, we found 1,163 cellular markers with cross-species conservation and further identified 1,685 cell-type-specific signature genes enriched in humans, 2,128 in macaques, and 733 in the mouse hypothalamus (Figures 1E and S3D; Table S1C; STAR Methods). To gain deeper insights into the molecular basis of hypothalamus evolution, we analyzed 340 previously identified primate-specific genes (PSGs)[18,19] and revealed the expression of 19 genes with neural cell-type specificity (Figure 1F; Table S1D). Notably, many of the identified PSGs were reported as risk factors for hypothalamus-related disorders. For instance, neural progenitor-specific *ZNF519* has been shown to associate with microcephaly and lissencephaly, while genetic defects in *RABL2A*, enriched in glutamatergic neurons, might lead to male infertility, growth retardation, and excessive weight gain in humans.[20,21]

Across neural cell types, we applied multiple trajectory inference methods to confirm a conserved differentiation pathway from RGCs to neurons and glial cells (Figures S4A and S4B). We also found that human hypothalamic neural progenitors employed a similar stepwise diversifying model for neurogenesis as observed in mice,[10] wherein RGCs generate two independent IPC subpopulations, and one subtype of them has the potential to differentiate into both glutamatergic and GABAergic neurons (Figure S4A). This finding suggests a common ontogenetic rule for neuronal lineage progression in the human and mouse hypothalamus.

As the primed expression of sublineage-specific marker genes predicted the fates of undifferentiated RGCs,[22] we assessed the timing of these neural progenitors entering the primitive state of lineage differentiation. We found that the priming of hypothalamic gliogenesis, including the generation of oligodendrocyte precursors (*PDGFRA*$^+$), astrocytes (*AQP4*$^+$), and ependymocytes (*FOXJ1*$^+$), occurred earlier in humans than in mice by aligning the neurodevelopmental time scales across species (Figures S4C–S4E). Cell composition analysis indeed showed that human glial cells emerged during the early-to-middle fetal stage, but mouse gliogenesis did not reach a detectable level until perinatal stage (Figure 1G). We next established correspondence between human and mouse developmental ages using a transcriptome age indexing approach (Figures S5A and S5B; STAR Methods) and found that the expression of cellular components involved in synapse formation was higher in humans than in mice at stage 3 (Figures 1H and S5C–S5G). These data suggest the temporal progression of gliogenesis and synaptogenesis, rather than neurogenesis, displays some degree of heterochrony

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

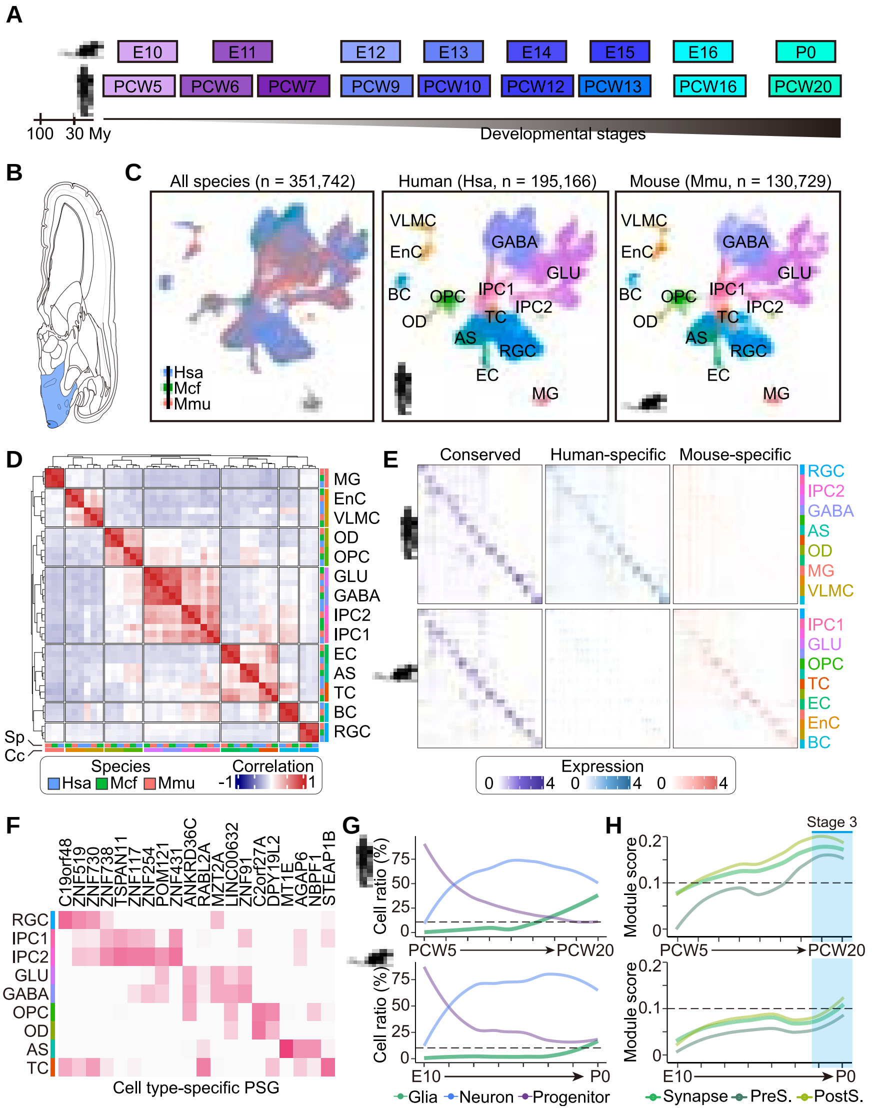


Figure 1. Conserved cell classes with divergent features in the developing human hypothalamus

(A) An overview of developmental time points sampled for the human and mouse hypothalamus. E, embryonic; P, postnatal; PCW, weeks post-conception; My, million years.

(B) Schematic diagram showing a coronal section of human brains at PCW16, with hypothalamus highlighted in blue.

(C) UMAP visualization of a total of 351,742 cells collected from the developing human (Hsa), macaque (Mcf), and mouse (Mmu) hypothalamus. RGC, radial glial cells; IPC, intermediate progenitor cells; AS, astrocytes; OPC, oligodendrocyte precursor cells; OD, oligodendrocytes; TC, tanycytes; EC, ependymal cells; GABA, GABAergic neurons; GLU, glutamatergic neurons; BC, blood cells; EnC, endothelial cells; VLMC, vascular and leptomeningeal cells.

(D) Pearson correlation of all cell types in different species. Sp, species; Cc, cell classes. (E) Heatmaps showing the conserved and species-specific molecular features in each cell type.

(F) Primate-specific genes (PSG) enriched in each neural cell type.

(G) Cell ratio dynamics of glia, neurons, and progenitors in the human and mouse hypothalamus.

(H) Module score dynamics showing the developmental changes in synaptic, presynaptic (PreS), and postsynaptic (PostS) gene expression.

between humans and mice, supporting that gliogenesis coordinates with the establishment of synapse.[23]

Conserved spatial patterning of neural progenitors

The hypothalamic primordium undergoes anteroposterior (AP) and dorsoventral (DV) regionalization to establish diverse progenitor domains, which are arranged into preoptic, anterior, tuberal, and mammillary zones (Figure 2A).[24,25] However, the molecular patterning of various progenitor domains in the human hypothalamus is still unclear. We thereby isolated RGCs at stages 1 and 2 and preserved both hypothalamic and prethalamic progenitors with spatial identities for subsequent analyses (Figure S6), given the critical role of prethalamus in hypothalamic patterning.[26] We identified 9 hypothalamic and 3 prethalamic progenitor domains (Figures 2B, S6A, and S7A), using the putative region-specific markers (e.g., *PAX6*, *NKX2-1*, *TBX3*, *SIX6*, and *BARHL2*) and a spatiotemporally resolved transcriptomic atlas of mouse organogenesis generated by spatial enhanced resolution omics-sequencing (Stereo-seq).[27] Compared with recent studies,[16,28] our work provides a more complete and comprehensive mapping of hypothalamic progenitor domains (Figures S7B–S7F). The majority of these domains were also identified in macaque datasets (Figures S7G–S7J). These domains were characterized by different molecular codes (Figure 2C; Table S1E), which emerged before the onset of neurogenesis and constantly defined distinct domains throughout development (Figures S7K–S7M). We further found that each domain was demarcated by a set of evolutionarily conserved transcription factors (TFs) that may dictate progenitor identities (Figures 2D and S8A–S8C). Importantly, forkhead factors *FOXG1*, *FOXD1*, *FOXD2*, *FOXB1*, *FOXA1*, and *FOXA2* were expressed in distinct domains in a segment-specific pattern (Figures 2E–2G and S8D–S8F), supporting previous studies demonstrating the critical role of Foxd1 and Foxb1 in specifying neurons for anterior nucleus and mammillary body.[29,30] The zonal segmentation along the AP axis by FOX gene family was shared between humans and mice (Figure 2E).

Notably, brain morphogenesis is controlled by both intrinsic factors and extrinsic morphogens. While Spemann's organizer establishes the identity of neural plate, secondary organizers emerge from discrete domains and release diffusible signaling factors to refine regional patterning (Figure 2H).[31] In the prethalamic and hypothalamic primordium, diverse progenitor domains expressed different morphogen codes consisting of Sonic

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

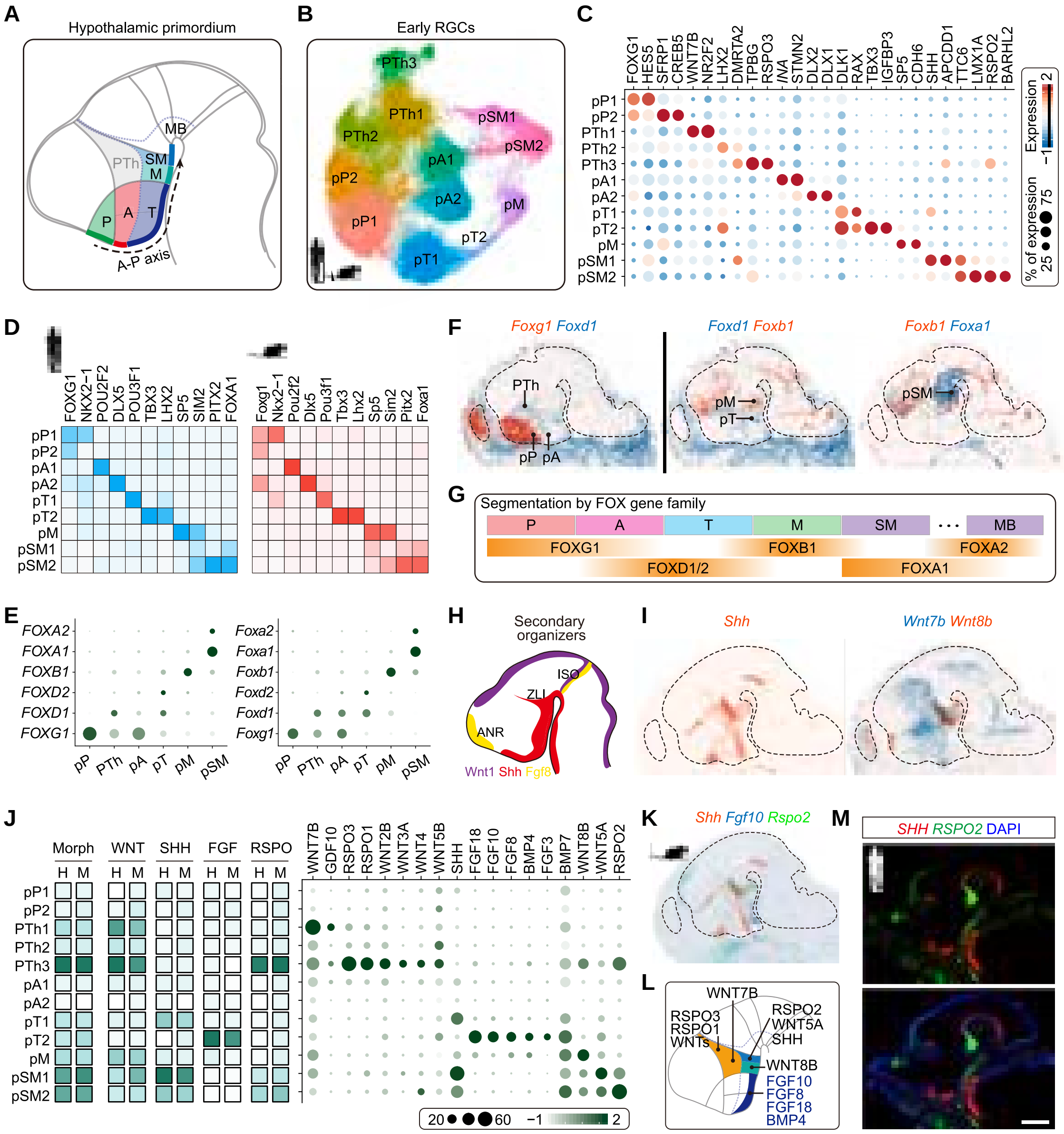


Figure 2. Spatial patterning of progenitor domains shared between humans and mice

(A) A diagram of classical segmentation of early developing hypothalamus into preoptic (P), anterior (A), tuberal (T), and mammillary (M) zones along the AP axis. PTh, prethalamus; MB, midbrain; SM, supramammillary zones.

(B) UMAP projection colored by diverse progenitor domains with distinct spatial code expression.

(C) Marker gene expression for each hypothalamic and prethalamic progenitor domains.

(D) Conserved TFs expressed in each human and mouse hypothalamic progenitor domains.

(E) Conserved expression of FOX gene family in different human and mouse hypothalamic zones.

(F) Spatial visualization of FOX gene expression patterns in E12.5 sagittal mouse brain sections.

(G) A summary of the segmentation of hypothalamic progenitor domains by FOX family genes.

(H) Schematic of the well-established secondary organizers: anterior neural ridge (ANR), zona limitans intrathalamica (ZLI), and isthmic organizer (ISO).

(I) Spatial visualization of morphogens in E12.5 mouse brain.

(J) Gene expression score (left) and profile (right) of morphogens including SHH, WNT, FGF, and RSPO families in diverse progenitor domains. Morph, morphogens; H, human; M, mouse.

(K) Spatial expression pattern of *Shh*, *Fgf10*, and *Rspo2*.

(L) Graphical summary of tertiary organizers in the developing hypothalamus.

(M) Sample images showing the spatial expression of *SHH* and *RSPO2* mRNA in a sagittal view of PCW6 human fetal brain. Scale bar, 1 mm.

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

Hedgehog (SHH), Wingless/Int-1 (WNT) family, fibroblast growth factors (FGFs), and R-spondin (RSPO) family (Figures 2I and 2J). Among the 12 progenitor domains, we identified PTh3, pT2, and pSM1/2 as tertiary organizers with differentially enriched expression of morphogens (Figures 2J–2L). Specifically, a subset of WNTs (e.g., *WNT7B*, *WNT2B*, and *WNT3A*) and RSPOs (e.g., *RSPO3* and *RSPO1*) defined anterodorsalizing organizer in early developing hypothalamus, whereas other WNT signals (*WNT8B*, *WNT5A*, and *RSPO2*) coordinated with *SHH* to establish caudalizing organizer. In the hypothalamic floor plate, FGF genes (*FGF18*, *FGF10*, and *FGF8*) were restricted in pT2 domain and mutually exclusive with *SHH* (Figure 2L). While Stereo-seq data demonstrated the spatial expression of morphogens in mice (Figure 2K), we further performed *in situ* hybridization to validate the expression of *SHH*, *FGF10*, and *RSPO2* in human brains at PCW6. Our human data confirmed the overlapping expression of *SHH* and *RSPO2* in pSM1/2 and the specific enrichment of *FGF10* in pT2 (Figures 2M and S8G). These results suggest that human and mouse brains share the common tertiary organizers and patterning mechanisms to induce TF expression and thereby establish the spatial identities of diverse hypothalamic progenitor domains.

Inference of neuronal sublineages for each progenitor domain

Distinct progenitor domains in the early developing hypothalamus presumably contribute to the extensive neuronal diversity through multiple sublineages.[32] However, current pseudotemporal ordering algorithms lack the power to simultaneously reconstruct individual sublineages spanning progenitors and their intermediate and differentiated states.[33,34] Here, we developed a computational strategy to infer the cellular fate transition of each progenitor domain by linking different cell types along the lineage hierarchy (Figures 3A and 3B; STAR Methods). We categorized IPCs and neurons into 14 and 22 subgroups, respectively (Figure 3B), and validated the strength and consistency of our results using publicly available datasets (Figures S9 and S10; Tables S1F and S1G). Importantly, we found that signature gene modules for each progenitor domain were potentially transmitted from progenitors to progeny (Figure S11A). We then reconstructed a neuronal lineage map by calculating the RGC-IPC and IPC-neuron relationships (Figures 3A, S11B, and S11C). To further demonstrate the robustness of our computational approach, we applied it to previously published single-cell datasets and were able to successfully reconstruct the cellular lineage (Figures S11D–S11H).

Next, we adapted Shannon's information theory for inferring lineage diversification spectrum with TF codes, which could define discrete progenitor domains and their progeny.[35,36] Our analysis indicated that T-zone (pT1, pT2, and their progeny) and M-zone (pM, pSM1, pSM2, and their progeny) lineage exhibited lower Shannon diversity index (Figure 3C), suggesting higher specialization in both lineages. To validate the inferred sublineages, we generated a high-resolution spatial transcriptomic atlas of serial coronal sections from E13.5 mouse brains using Stereo-seq (Figure S12A) and aligned our single-cell clusters to the spatial transcriptomic atlas with Tangram. The resulting probabilistic spatial profiles of specific cell subtypes not only demonstrated the spatial segregation of T-zone lineage from M-zone lineage but also corroborated the spatial proximity of each progenitor domain and its daughter cells (Figures 3D and S12B–S12E), supporting the robustness of our lineage inference.

To identify lineage-specific TFs that may direct neuronal specification, we employed a cosine similarity-based algorithm for specific gene identification along cellular hierarchy. Each neuronal sublineage was featured by distinct TF codes (Figures 3E and S13A; Table S1H), which showed functional diversification or redundancy for lineage specification (Figures S13B and S13C). These potential lineage factors were evolutionarily conserved to a large extent. For example, *FOXB1* and *BARHL1* specify two adjacent but distinct M-zone sublineages with sharp boundary, whereas *TBX3* defines a sublineage within T-zone in both humans and mice (Figures 3F and 3G). To verify the inferred neuronal sublineages, we labeled pT2 progenitor domain using *Tbx3-CreER*^{T2}*::Ai9* mice to track neuronal fate at postnatal stages (Figures 3H and S13D). Single-cell RNA sequencing at P14 confirmed 8 out of 9 neuronal subtypes arising from Tbx3$^+$ progenitors mapped onto T-zone lineage, including POMC, AgRP, and KNDy neurons (Figures 3I and S13E). Moreover, many neuronal subtypes from the inferred pT2 sublineage were lost in Tbx3-deficient brains (Figures 3J and S13F–S13J), substantiating our recent finding that Tbx3 depletion disrupts lineage specification.[8] These data collectively support the fidelity of our inferred lineage tree and the functional significance of lineage factors.

We further subdivided cells into 10 pseudotemporal bins, identified gene modules by non-negative matrix factorization (NMF), and annotated them with the top-ranked genes (Figures S14A–S14D; Tables S1I and S1J), aiming to uncover species-specific metagene programs that drive lineage progression. Clustering analysis revealed shared biological processes across species, with human neuronal lineages showing specific RNA splicing enrichment and mouse development featuring higher oxidative phosphorylation activity (Figures S14C–S14G). Increased RNA splicing in humans may enhance protein complexity, favoring energy-efficient information processing,[37] while lower mitochondria metabolism contributes to human brain neoteny.[38] Altogether, we computationally reconstructed neuronal lineage tree in the developing hypothalamus, highlighting conserved lineage factors and distinct metagene programs specific to humans.

Shared and divergent features of neuronal subtypes during development

Distinct neurogenic sublineages diverge across the developmental continuum and give rise to a multitude of neuronal subtypes. To reveal neuronal heterogeneity, we independently clustered human and mouse hypothalamic neurons to identify neuronal subtypes, followed by cross-species comparison to establish their homology (STAR Methods). We found that most of neuronal subtypes were conserved between humans and mice (Figure S15A). Given the conservation, we integrated human and mouse postmitotic neurons and identified 26 glutamatergic and 24 GABAergic subtypes with distinct molecular signatures and spatial distribution (Figures 4A, S15B, S15C, and S16A; Table S1K). Further comparison between human and macaque neurons supported their conservation to a large extent (Figures S16B–S16D), despite the number of macaque neurons

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

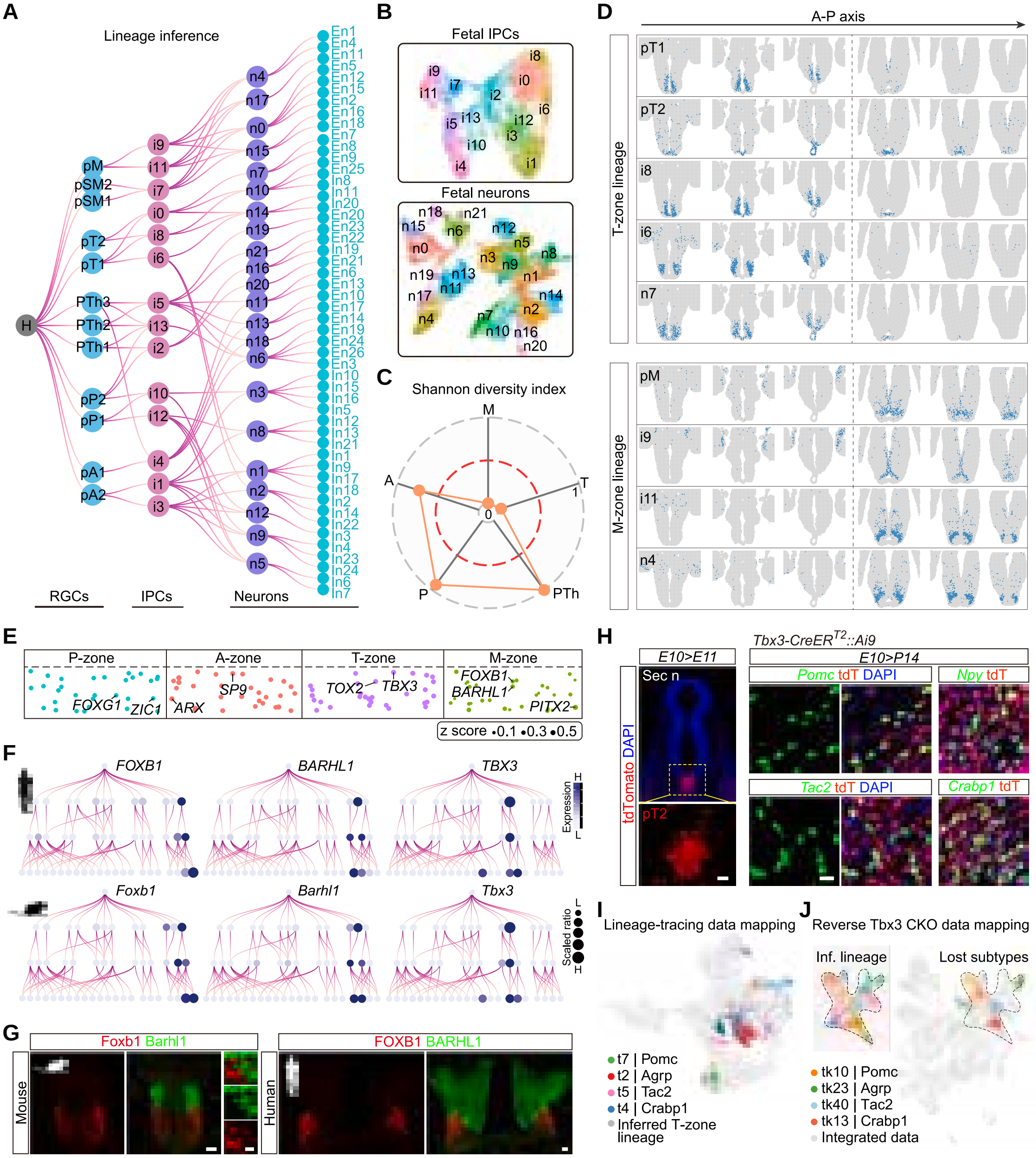


Figure 3. Inference of hypothalamic lineage tree and conserved lineage factors

(A) Computational reconstruction of hypothalamic neuronal lineage tree linking RGC progenitor domains (blue), IPCs (pink), and neurons (purple and cyan).

(B) Subclustering analysis of IPCs and neurons at embryonic and fetal stages for lineage inference.

(C) Shannon entropy analysis of the inferred PTh-, P-, A-, T-, and M-zone lineages based on TF expression diversity. A smaller Shannon diversity index indicates a higher fidelity of lineage identity during neuronal differentiation.

(D) Spatial transcriptomic analysis of signature gene modules for each progenitor and neuronal subtype within T- and M-zone lineages.

(E) Identification of lineage-specific TFs for P-, A-, T-, and M-zone lineages.

(F) Crown plots showing three conserved lineage factors in humans and mice.

(G) Immunostaining indicating the spatial segregation of M and SM zones by FOXB1 and BARHL1 in both E10.5 mouse and PCW6 human brains. Scale bars, 100 and 10 μm.

(H) Genetic labeling of pT2 progenitor domain with *Tbx3-CreER*^{T2}*::Ai9* mice at E10 and characterization of progeny neuron identities at P14 by single-molecule fluorescent *in situ* hybridization. Scale bars, 50 and 20 μm.

(I) Mapping of single-cell data from Tbx3-derived lineage onto UMAP plot that integrates T-zone neuronal lineage. Experimentally traced neuronal subtypes are color-coded.

(J) Reverse mapping of randomly sampled cells from inferred Tbx3 lineage onto UMAP plot that integrates 50,872 hypothalamic neurons from control and Tbx3 conditional knockout (CKO) mice.

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

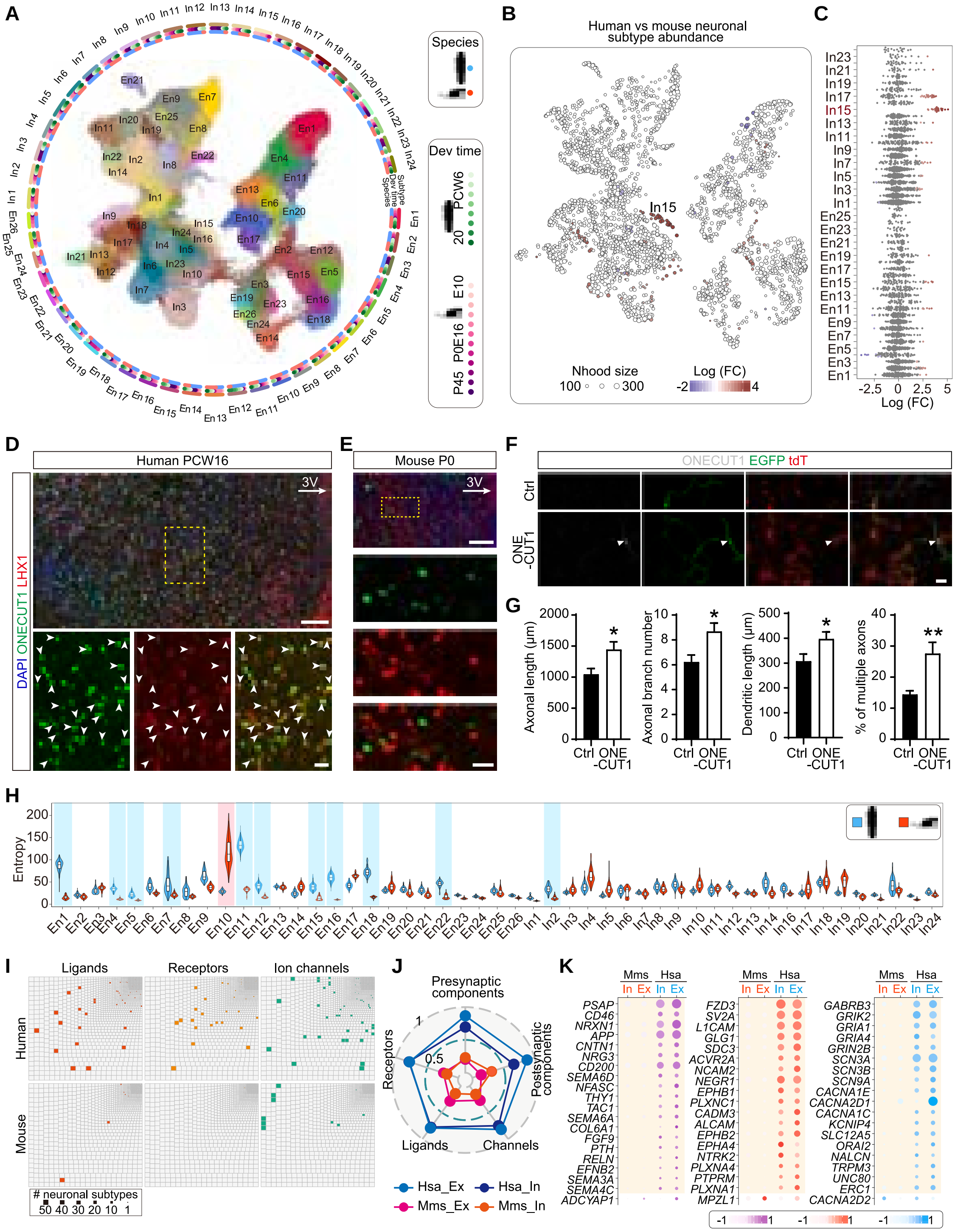


Figure 4. Transcriptional conservation and innovation of neuronal subtypes

(A) UMAP plot of transcriptionally defined neuronal taxonomy of the mouse and human developing hypothalamus, visualizing 155,089 neurons. Surrounding histograms from inner to outermost represent cell proportions for each species, developmental time points, and group, respectively.

(B) Projection of a neighborhood graph identified by Milo differential abundance testing onto UMAP plot. Nodes represent neighborhoods, colored by their $\log_2$ (fold change) across species.

(C) Beeswarm plot of the distribution of $\log_2$ (fold change) across species in neighborhoods containing cells from different neuronal subtypes.

(D and E) Sample images showing the species-specific abundance of ONECUT1$^+$LHX1$^+$ neurons in human PCW16 hypothalamus (D) compared with P0 mouse brains (E). 3V, third ventricle. Scale bars, 100 and 20 μm.

(legend continued on next page)

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

being limited. We identified 89 TFs with conserved and specific expression in different neuronal subtypes (Figure S17) and adapted single-cell regulatory network inference and clustering (SCENIC) to unveil a list of potential master regulators shaping neuronal identities (Figures S18 and S19; Tables S1L and S1M).

Given that voluntary control of hypothalamic subsystems is mediated by inputs from the medial prefrontal region and ventral temporal lobe of neocortex in humans,[39,40] we expected adaptive changes in neuronal composition, molecular expression, or spatial location across species due to neocortex expansion. Indeed, we identified a neuronal subtype (In15) that was much more abundant in the developing human hypothalamus compared with mice (Figures 4B and 4C), confirmed by neuronal subtype alignment analysis (Figures S20A and S20B). Further immunostaining revealed the presence of ONECUT1$^+$LHX1$^+$ In15 neurons across the primordium of anterior nucleus and dorsomedial nucleus in the human hypothalamus at PCW6 and PCW12 (Figures 4D, S20C, and S20D), in contrast to the minimal overlap between ONECUT1 (One Cut Homeobox 1) and LHX1 (LIM Homeobox 1) signals in prenatal and perinatal mouse brains (Figures 4E and S20E). To assess the role of ONECUT1 in neuronal morphogenesis, we ectopically expressed human ONECUT1 in the developing mouse hypothalamic neurons, demonstrating increased neurite outgrowth both *in vitro* and *in vivo* (Figures 4F, 4G, and S20F–S20I). Interspecies divergence arises from not only the gain or loss of specific neuronal subtypes but also the complexity of individual homologous cell subtypes. We adopted an entropy-based metric and revealed that the intercellular transcriptomic variability of 10 excitatory subtypes and one inhibitory subgroup was significantly augmented in the human hypothalamus (Figure 4H). These data imply the potential functional diversification of human hypothalamic excitatory neurons.

Differential gene expression profiles can also contribute to species-specific variations in the structure and function of each conserved cell subtype. Our analysis demonstrated a relatively limited human-mouse transcriptomic divergence across various neuronal subtypes (Figures S21A–S21C; Table S1N). While homologous neuronal subtypes in the human and mouse hypothalamus largely shared core regulatory complex TFs (Figure S21D; Table S1O), a substantial number of genes encoding ligands, receptors, ion channels, and synaptic components were upregulated in the human excitatory and inhibitory neurons in both sexes, as compared with mice (Figures 4I, 4J, S21E, and S21F; Table S1O). Specifically, human neurons displayed higher expression levels for axon guidance cues (e.g., *SEMA6D*, *SEMA6A*, and *EFNB2*), their receptors (e.g., *EPHB1*, *PLXNC1*, and *PLXNA4*), neurite outgrowth regulators (*FZD3* and *NTRK2*), postsynaptic channels (subunits of AMPA, NMDA, and kinase receptors), and clustered protocadherins (Figures 4K and S22A–S22C), which could facilitate extensive and intricate connectivity. Overall, we identified 50 unique neuronal subtypes with diverse molecular signatures in the developing hypothalamus, revealed a transcriptionally distinct neuronal subtype enriched in humans, and disclosed multiple sources of cross-species variation that may contribute to the increased complexity, connectivity, and modulation of human hypothalamic neurons.

Spatial redistribution of human GnRH and GHRH neuroendocrine neurons

An inherent attribute of the mammalian hypothalamus lies in its intricate neuroendocrine system (Figure S23A), which comprises a heterogeneous collection of neuropeptidergic neurons secreting oxytocin, vasopressin (AVP), corticotropin-releasing hormone (CRH), thyrotropin-releasing hormone (TRH), somatostatin, growth-hormone-releasing hormone (GHRH), and gonadotropin-releasing hormone (GnRH) either directly into the pituitary or indirectly via hypophyseal portal system.[41] To profile developing neuroendocrine neurons, we isolated them using canonical markers (e.g., neurohormones and *SCG2* encoding a neuroendocrine secretory protein) and revealed 13 subclusters within 7 main subtypes (Figures S23B–S23D; Table S1P). Pseudobulk differential expression analysis has revealed the molecular characteristics of these neuronal subtypes, such as enriched expression of genes associated with DNA damage and oxidative stress in CRH neurons (Figure S23E). We further identified a set of shared TFs within each human and mouse neuronal subtype (Figure S23F), implicating their roles in establishing cellular identities or functions. For instance, ONECUT family genes, known for specifying neuronal fate, were enriched in TRH neurons,[10] while NR3C1 expression, involved in negative feedback control of glucocorticoid release, emerged early in prenatal CRH neurons.[42]

Despite diverse spatial positions, neuroendocrine neurons project their axons uniformly toward the hypothalamic median eminence that encompasses hypophyseal portal vessels.[41] Our single-cell and spatial transcriptomic data suggested that Netrin-DCC, Slit2-Robo, and ephrinA5-Eph signaling may coordinate to mediate the precise pathfinding of neuroendocrine axons (Figure S24), extending previous studies showing that SHH, BMP7, and FGF10 regulate the orientation of hypothalamic axons toward median eminence.[43,44]

To reveal interspecies differences in neuroendocrine neurons, we first focused on cell subtype abundances, noting an

(F) Sample images showing the morphology of hypothalamic neurons (red) ectopically expressing human ONECUT1 (white) and EGFP (green). These neurons were dissected from E16.5 *vGat-Cre::Ai14* mice, cultured *in vitro* for plasmid transfection, and analyzed after 7 days. Scale bar, 75 μm.

(G) Quantification of neuronal morphology (control, $n = 24$; ONECUT1, $n = 35$). Data are shown as mean ± SEM ($^*p < 0.05$; $^{**}p < 0.01$ by two-tailed unpaired Student's t test.

(H) Transcriptomic heterogeneity among cells in each neuronal subtypes across species, reflected by Shannon entropy. A neuronal subtype (In15) with significant human-mouse difference (adjusted P value < 0.01 and $\log_2$ (fold change) > 1) is highlighted in color.

(I) Treemap showing the frequency of human- or mouse-enriched ligands, receptors, and channels, which are color-coded. The size of each box represents the number of neurons subtypes with species-specific gene enrichment.

(J) Normalized gene expression scores for ligands, receptors, channels, presynaptic, and postsynaptic components in both human and mouse neurons. Ex, glutamatergic neurons; In, GABAergic neurons.

(K) Comparative gene expression analysis of human and mouse ligands (left), receptors (middle), and channels (right).

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

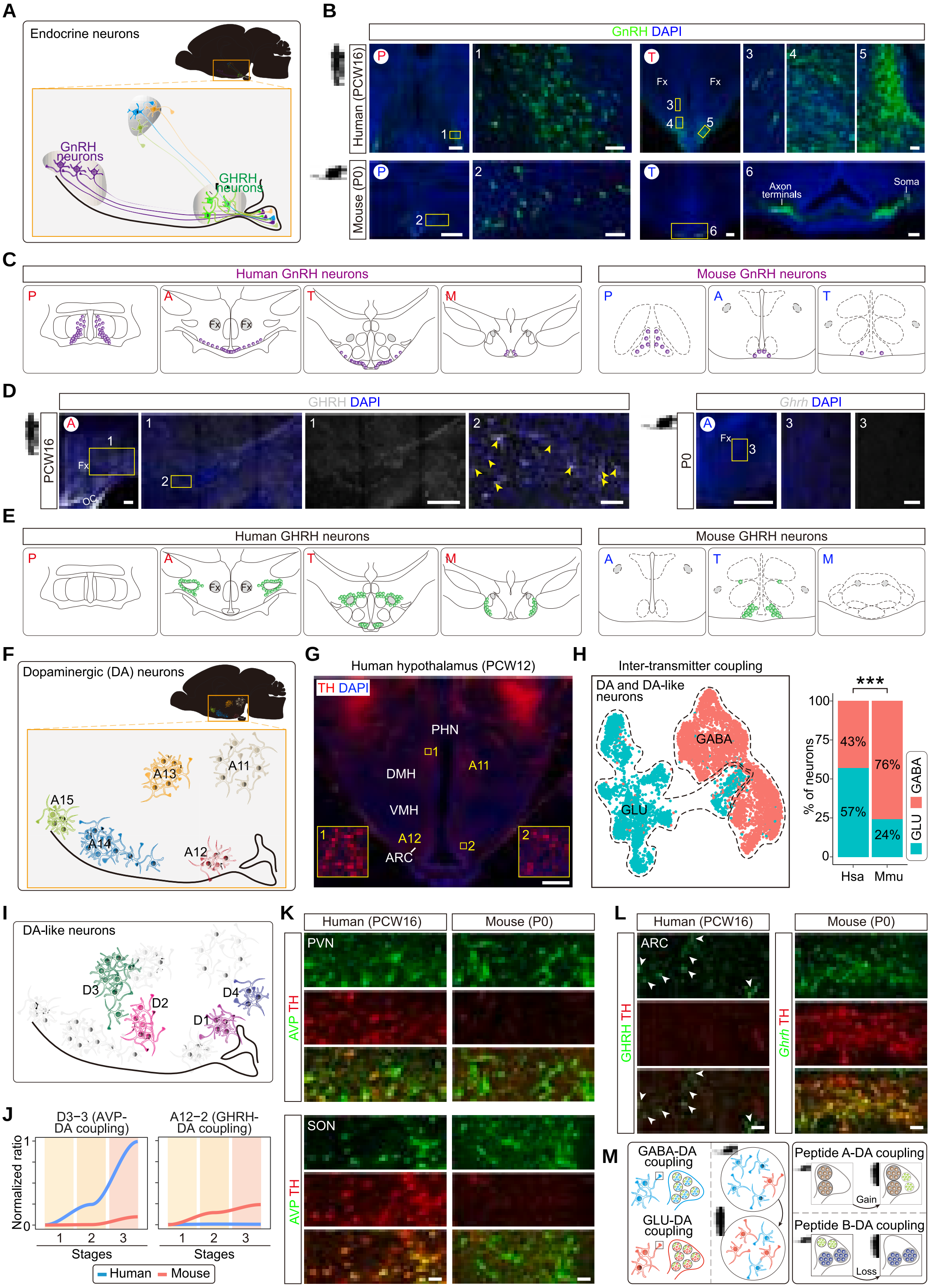


Figure 5. Species-specific variations in neuronal distribution and signal transduction

(A) A schematic showing spatial distribution and projection of neuroendocrine neurons in mouse brains.

(B) Representative images demonstrating the widespread distribution of GnRH neurons in the P- and T-zones of PCW16 human hypothalamus (top row), in contrast to their confinement primarily to preoptic area in mice (bottom row). Fx, fornix. Scale bars, 500 and 50 μm.

(C) Graphical summary of the spatial distribution patterns of human and mouse GnRH neurons.

(legend continued on next page)

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

increased proportion of GnRH neurons within the human developing hypothalamus (Figures S25A and S25B). GnRH neurons, controlling puberty onset and reproduction, originate from olfactory placode and migrate to the hypothalamus along olfactory-vomeronasal nerves in mammals.[45] Immunostaining of human PCW16 and mouse neonatal brains confirmed that the human hypothalamus accommodated a greater number of GnRH neurons (Figures 5A, 5B, and S25C–S25E). Next, we found a broader spatial distribution of human GnRH neurons, ranging from septal-preoptic region to infundibular nucleus (i.e., rodent arcuate nucleus) and mammillary body, whereas mouse GnRH neurons were predominantly distributed in the preoptic zone (Figures 5B, 5C, and S25F–S25I). Our results support previous reports on species differences in GnRH neuron distribution[46,47] but further show that human GnRH neurons deviated from hypothalamic midline and migrated along the ventrolateral margin after crossing septal-preoptic region (Figure S25C), implicating that they move along their own axonal guides upon detachment from vomeronasal nerves.

Furthermore, we unexpectedly revealed a species-specific variation in the spatial organization of GHRH-expressing neurons, which regulate growth and metabolism. In contrast to the spatial accumulation of mouse GHRH neurons in arcuate nucleus (Figure S25J), their human counterparts were not only distributed in infundibular nucleus but also dispersed into lateral hypothalamus, posterior hypothalamic nucleus, and lateral mammillary region (Figures 5D, 5E, and S25K). Notably, the spatial distribution of human GHRH neurons in lateral hypothalamus displayed a wreath-like shape (Figure 5E), encircling a densely aggregated subnucleus lateral to fornix. These findings suggest an early genesis and multiple origins of GHRH neurons in fetal human brains and also provide clues for the potential production of GHRH beyond arcuate nucleus as a neuromodulator or circulating hormone. Together, human GnRH and GHRH neurons are spatially redistributed during evolution, implying species-specific differences in their network organization and functional significance within the hypothalamus.

Reshuffled inter-transmitter and peptide-dopamine couplings across species

Another important attribute of hypothalamic neurons is their intricate interplay between neurotransmitters and neuropeptides, especially in the dopamine neuron system, which is traditionally subdivided into A11-A15 groups (Figures 5F, 5G, S26A, and S26B).[48,49] To investigate species differences in inter-transmitter and peptide-dopamine couplings, we subset hypothalamic dopamine neurons using the established criteria.[15] While it is hypothesized that they originate from GABA lineage and prevalently co-express GABA synthesis pathway in rodents,[15,50] we noticed an interesting coexistence of dopamine-glutamate and dopamine-GABA couplings (Figures 5H and S26C–S26E). This dual-transmitter dichotomy potentially stems from divergent TF code expression, as evidenced by DLX (Distal-Less Homeobox) factors associated with dopamine-GABAergic neurons and NHLH2/UNCX enriched in dopamine-glutamatergic neurons (Figure S26E). Notably, our ratio analysis unveiled that dopamine-glutamate coupling was more pronounced in the human dopamine system than in mouse neurons (Figures 5H and S26F). This divergence in inter-transmitter coupling may drive phenotypic differences between species, such as evolutionary alteration in reward learning and motivated behavior.

Among these dopamine and dopamine-like neurons, we identified 10 canonical subtypes (A11-A15) primarily coupled with GABA and 8 non-canonical subclusters (D1-D4) with glutamatergic feature (Figures 5I and S26G–S26J; Table S1Q). They exhibited distinct transcriptional profiles from midbrain dopaminergic neurons (Figures S27A–S27C), featured by En1 and En2 expression.[51] Genome-wide association mapping connected them to psychiatric conditions, metabolic disorders, and personal traits but not to Parkinson's disease (Figure S27D). Intriguingly, the D4 subtype displayed the highest transcriptional similarity to midbrain dopamine neurons (Figures S27E–S27H), implying potential value to alleviate Parkinson's disease. We further showed a prevalence of D3-3 subtype in human dopamine system as compared with mice, suggesting a pronounced colocalization between dopamine and AVP in humans (Figure 5J). Immunostaining confirmed that a subset of AVP neurons coexpressed tyrosine hydroxylase (TH) in human paraventricular and supraoptic nuclei (Figures 5K and S27I). By contrast, GHRH-dopamine coupling was more prominent in mice than in humans (Figures 5J and 5L). These data suggest that particular subtypes of dopamine neurons could either gain or lose interactions with specific neuropeptides, potentially influencing catecholamine-mediated neuronal signaling. Collectively, we revealed an increase in the dopamine-glutamate interaction within human hypothalamus compared with mice and a reshuffled dopamine-neuropeptide coupling across species (Figure 5M), which presumably serves as an important source of interspecific

(D) Immunostaining showing a wreath-shaped distribution of GHRH neurons in the lateral hypothalamus of humans but not in mice. OC, optic chiasm. Scale bars, 500, 200, and 50 μm.

(E) A graphic summary of the differences in GHRH neuron distribution between humans and mice.

(F) Spatial definition of canonical dopamine (DA) neurons in the mouse hypothalamus.

(G) A sample image showing the distribution pattern of TH-expressing DA neurons in the human hypothalamic T-zone at PCW12. PHN, posterior hypothalamic nucleus; DMH, dorsomedial nucleus; VMH, ventromedial nucleus; ARC, arcuate nucleus. Scale bar, 1 mm.

(H) UMAP visualization of hypothalamic DA and DA-like neurons co-expressing either glutamate or GABA transmitter, while cell ratio analysis highlights the differences in DA-glutamate and DA-GABA couplings between human and mouse on the right. ***$p < 0.001$ determined by permutation test.

(I) Identification and spatial annotation of non-canonical DA-like neurons in the mammalian hypothalamus.

(J) Dynamic curves of normalized cell ratio across species and developmental stages, highlighting the species difference in D3-3 and A12-2 neuronal subtypes.

(K) Sample images showing the co-expression pattern of AVP and TH in the PCW16 human and P0 mouse paraventricular nucleus (PVN) and supraoptic nucleus (SON). Scale bars, 20 μm.

(L) Sample images showing the coupled expression of TH with GHRH in mouse but not human ARC. Scale bars, 20 μm.

(M) Schematic diagrams summarizing a model of potential evolutionary changes of inter-transmitter and peptide-transmitter couplings in human DA and/or DA-like neurons.

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

variation in the behavioral effects of dopamine and distinct neuropeptides.

DISCUSSION

Subcortical brain structures substantially contribute to homeostatic regulation, behavioral control, and cognitive functions; however, our understanding of subcortical development and evolution in humans falls far behind that of neocortex. In this study, we present comprehensive datasets that reveal conserved neural patterning and neurogenic lineages in the developing human and mouse hypothalamus, alongside species-specific neuronal subtype abundance, molecular characteristics, spatial distribution, and neuromodulator couplings. While the human fetal hypothalamus features a unique neuronal subtype and increased neuromodulatory gene expression, we also observed spatial redistributions in two subtypes of human neuroendocrine neurons and a reconfiguration of neuromodulator coupling within human hypothalamic dopamine neurons compared with mice. Finally, we developed a web-based data resource that facilitates cross-species comparison of hypothalamic development.

Conserved spatial patterning of tertiary organizers and progenitor domains

During early neurodevelopment, mammalian brains are organized into spatially distinct progenitor domains, generating various neuronal subtypes through specific spatiotemporal specification processes.[52] We combined single-cell and spatial transcriptomics to define 9 hypothalamic progenitor domains and revealed the segmentation of hypothalamic primordium along the AP axis orchestrated by FOX gene family. Comparative analysis demonstrated that the spatial TF codes demarcating diverse progenitor domains were generally shared between human and mouse hypothalamus. It is well-known that spatial patterning of neural progenitors relies on morphogen gradients that induces TF code expression.[53,54] We analyzed morphogen (e.g., Shh, Wnt-, Fgf-, Bmp-, and Rspo-family genes) expression and proposed that the spatial organization of developing hypothalamus was shaped by tertiary organizers with evolutionary conservation. Specifically, *WNT7B*, *RSPO3*, and *RSPO1* delineated anterodorsalizing organizer in the prethalamus; *WNT8B*, *WNT5A*, and *RSPO2* were enriched in caudalizing organizer; and a complementary expression of SHH and FGF-family genes was orchestrated to establish a floor-plate organizer. The proposed tertiary organizers may instruct the self-organization of various progenitor domains and induce locally different neuronal fates. These results, yielded from single-cell and spatially resolved transcriptomic analyses, broaden our understanding of the compartmentalized expression of secreted ligands in conferring regional identity of diverse neural progenitors.

Developmental logics and conserved lineage factors directing lineage specification

Neural progenitors can give rise to distinct neuronal subtypes through lineage divergence and/or convergence.[55] This dynamic entails the production of a unique set of neuronal subtypes by each progenitor domain (divergence), while also allowing neural progenitors from different lineages to generate identical or closely related neuronal subtypes (convergence). To grasp the framework of lineage progression in the hypothalamus, we developed a computational tool to systematically map a hypothalamic neuronal lineage tree by reconstructing lineage relationships between progenitor domains and neuronal subtypes. This tool hinges on the established hypothesis that neurons acquire the molecular identity via inheriting TF codes from their ancestor cells. In contrast to embryonic spinal cord, wherein distinct progenitor domains arrayed along the DV axis produce neurons in a divergent manner,[52] both lineage divergence and convergence may occur among various hypothalamic progenitor domains. Of note, neural progenitors in T- and M-zones demonstrate a more restricted spatial segregation and lineage specialization. We further profiled the lineage-specific TFs that potentially mark, specify, and maintain cellular identities along individual hypothalamic sublineages and suggested that functional redundancy and diversification of lineage factors from the same gene family at least partially underlie the logic governing sublineage specification. Importantly, these lineage factors are largely shared between humans and mice, highlighting conserved control of lineage-specific neuronal production in the hypothalamus. Consistent with this finding, genetic ablation of Tbx3, a critical lineage factor specifying pT2 sublineage, not only disrupts the establishment of POMC and KNDy neuronal identities in mice but also mirrors the clinical manifestations (such as obesity and delayed puberty) in patients with *TBX3* mutations.[8,56,57] Collectively, our data establish a lineage tree for the developing hypothalamus and unveil a conserved mechanistic framework with multilayered cellular and molecular logic to ensure the stability of neuronal production during lineage progression.

Human-specific neuronal features in the developing hypothalamus

Despite the conserved cellular mechanisms governing neural patterning and neurogenic lineage specification, we explored species divergence of abundance, heterogeneity, molecular features, and spatial distribution of homologous neuronal subtypes. First, we subdivided the hypothalamic postmitotic neurons into 50 subgroups and identified cross-species homologies between most human and mouse neuronal subtypes. Notably, the human fetal hypothalamus contains a unique ONECUT1$^+$LHX1$^+$ neuronal subtype. Given the roles of cadherin in neuronal sorting and migration,[58] this evolutionarily divergent neuronal subtype, featured by enriched expression of cell adhesion (*CDH11*, *CDH9*, and *CDH8*) and axon guidance (*SEMA3C*, *PLXNA4*, and *DAB1*) molecules, may regulate chronological organization of neurons in an outside-in manner, control axonal fasciculation, and modify final target recognition of neuronal projections from prefrontal cortex. Second, we reveal that several excitatory neuronal subtypes in the human hypothalamus are transcriptionally more diversified than in mice, implying greater structural and functional diversification during human neuronal development. Although long-range GABAergic projection neurons exist, remote axonal projections are predominantly mediated by glutamatergic neurons.[59] The expanded diversity in human excitatory neurons suggests that a more complicated intra-hypothalamic and trans-regional connectome emerges to relay command information and maintain system homeostasis. Third, we observed a more pronounced expression of genes regulating axonal

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

growth and synaptic signaling in humans. Higher levels of axon guidance molecules, clustered protocadherins, and neurite outgrowth regulators would elongate axonal growth, steer axonal projections more precisely, and promote repulsion between homotypic neurites.[60] In parallel, increased expression of presynaptic and postsynaptic components in human hypothalamic neurons enables a tight control of native behaviors, adapting to the rapidly expanded neocortex. Fourth, human GnRH and GHRH neurons exhibit a broader distribution in the hypothalamus compared with mice. This change suggests either multiple origins or redirected migratory path of human neuroendocrine neurons that govern vitally important processes, including growth, reproduction, and metabolism. Given that GnRH and GHRH neurons have also been shown to improve cognition and promote slow wave activity,[61–63] their spatial redistribution observed in humans might not only cater to heightened hormone release demand but also potentially enhance primate cognition. Lastly, our analysis of hypothalamic dopamine neurons provides proof-of-concept evidence to demonstrate the potential shift in inter-transmitter and peptide-transmitter couplings across species. Neuropeptides are auxiliary messenger molecules that always co-exist in neurons with one or more neurotransmitters and serve as neuromodulators to regulate neuronal activity.[48] Further studies to elucidate the potential functions of such species-specific variation in peptide-neurotransmitter couplings are required.

In summary, our study of hypothalamus development and evolution suggests that human subcortical structures might adopt conserved neural patterning strategies but appropriately adapt neuronal composition, distribution, input sensitivity, and output robustness for advanced social cognition and behavioral flexibility. The computational reconstruction of hypothalamic lineage tree provides a framework for experimentally deconstructing lineage-specific neuronal fate determination and functional connection. The innovation among hypothalamic neurons found in this work will facilitate us to delve into the cellular mechanisms underlying human-specific physiological function and disease susceptibility.

Limitations of the study

Our work aligns with recent studies on major cell types and developmental trajectories of neurons in the developing human hypothalamus.[16,28] It similarly provides in-depth analyses of key regulons, molecular features, and the spatial distribution of neuronal subtypes. In stark contrast, we elucidate the molecular logic of neural patterning, propose the concept of tertiary organizers, reconstruct the neurogenic lineage tree, and reveal four aspects of neuronal innovation, including distinct neuronal distribution and neuromodulator couplings. There are several limitations to the interpretation of our results. First, although we collected approximately 200,000 cells from the human hypothalamus, a larger sample size that includes both sexes and multiple replicates at each time point would strengthen our findings. We have validated the conservation and consistency of our results using publicly available datasets. Second, while we have provided deep insights into neural patterning mechanisms and neuronal innovation, functional analyses of FOX gene swapping, human-specific neuronal subtypes, and reorganized neuromodulator couplings present challenges for future research. Lastly, although we identified transcriptomic features specific to human hypothalamic neurons, the hub genes that shape these distinctive profiles and enable the conversion of mouse neurons into human-like ones remain unclear.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Qing-Feng Wu (wu_qingfeng@genetics.ac.cn).

Materials availability

All animals and unique/stable reagents generated in this study are available from the lead contact with a completed materials transfer agreement.

Data and code availability

All data and computational code used in this work are available upon request. The sequencing datasets generated in this study are freely accessible in the CNGB Nucleotide Sequence Archive under accession code CNGBdb: CNP0004213. The code for bioinformatic analyses is available at https://doi.org/10.5281/zenodo.13761830.

ACKNOWLEDGMENTS

We thank X. Xu and P. Shu for comments. We thank donors and their families for brain tissue donation. The work was supported by the National Natural Science Foundation of China (32230031 and 32425025), the Beijing Municipal Science & Technology Commission (Z210010), and the Hundred-Talent Program (Chinese Academy of Sciences). This paper is from the Mesoscopic Brain Mapping Consortium.

AUTHOR CONTRIBUTIONS

Q.W. designed the study. Q.W. and L.L. supervised the project. Z.C., T.P., Y.-H.Z., and M.G. performed most of data analyses. B.W., X.-Q.L., F.G., and Z.X. collected human and macaque brain samples. Z.C., Q.W., S.H., and M.X. conducted scRNA-seq, snRNA-seq, and Stereo-seq experiments. Y.S., C.Y., J.G., and H.W. assisted in computational analysis. X.-L.S., X.S., F.C., and C.-X.Y. performed immunostaining, *in situ* hybridization, and imaging. S.X., Y.Z., X.L., Z.D., S.L., Y.Z.L., and L.L. contribute to interpretation of results. Q.W. and Z.C. developed concepts and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Human sample collection
  - Macaque fetal hypothalamus collection
  - Mice
- METHOD DETAILS
  - Immunohistochemistry
  - Single-molecule fluorescent in Situ hybridization (smFISH)
  - Plasmid construction, transfection and electroporation
  - Single-cell isolation and scRNA-seq
  - Single-nucleus isolation and snRNA-seq
  - Stereo-seq
  - Sequencing data preprocessing
  - Quality control and data integration
  - Dimensionality reduction and clustering
  - Differential gene expression analysis
  - Gene ontology (GO) analysis

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

- ○ Pearson correlation analysis
  - ○ Cross-species cell type correlation
  - ○ Developmental trajectory analysis
  - ○ Gene module scoring
  - ○ Stage correspondence analysis
  - ○ Inference of lineage tree and lineage factors
  - ○ Shannon entropy analysis
  - ○ Spatial transcriptomic analysis
  - ○ Functional validation of inferred sublineage
  - ○ Cell abundance and compositional analysis
  - ○ Non-negative matrix factorization (NMF)
  - ○ Identification of homologous cell types
  - ○ Label transfer
  - ○ Identification of regulons
  - ○ GWAS enrichment analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - ○ General statistical analysis
  - ○ Quantification of neuronal number and neurite length

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.devcel.2025.03.009.

Received: May 31, 2024
Revised: February 7, 2025
Accepted: March 14, 2025
Published: April 8, 2025

REFERENCES

1. Zemke, N.R., Armand, E.J., Wang, W., Lee, S., Zhou, J., Li, Y.E., Liu, H., Tian, W., Nery, J.R., Castanon, R.G., et al. (2023). Comparative single cell epigenomic analysis of gene regulatory programs in the rodent and primate neocortex. Preprint at bioRxiv *888*, 888. https://doi.org/10.1101/2023.04.08.536119.
2. Sousa, A.M.M., Zhu, Y., Raghanti, M.A., Kitchen, R.R., Onorati, M., Tebbenkamp, A.T.N., Stutz, B., Meyer, K.A., Li, M., Kawasawa, Y.I., et al. (2017). Molecular and cellular reorganization of neural circuits in the human lineage. Science *358*, 1027–1032. https://doi.org/10.1126/science.aan3456.
3. Shi, Y., Wang, M., Mi, D., Lu, T., Wang, B., Dong, H., Zhong, S., Chen, Y., Sun, L., Zhou, X., et al. (2021). Mouse and human share conserved transcriptional programs for interneuron development. Science *374*, eabj6641. https://doi.org/10.1126/science.abj6641.
4. Zhu, Y., Sousa, A.M.M., Gao, T., Skarica, M., Li, M., Santpere, G., Esteller-Cucala, P., Juan, D., Ferrández-Peral, L., Gulden, F.O., et al. (2018). Spatiotemporal transcriptomic divergence across human and macaque brain development. Science *362*, eaat8077. https://doi.org/10.1126/science.aat8077.
5. Watson, C. (2012). The Mouse Nervous System (Academic Press).
6. Romanov, R.A., Alpár, A., Hökfelt, T., and Harkany, T. (2019). Unified Classification of Molecular, Network, and Endocrine Features of Hypothalamic Neurons. Annu. Rev. Neurosci. *42*, 1–26. https://doi.org/10.1146/annurev-neuro-070918-050414.
7. Biran, J., Tahor, M., Wircer, E., and Levkowitz, G. (2015). Role of developmental factors in hypothalamic function. Front. Neuroanat. *9*, 47. https://doi.org/10.3389/fnana.2015.00047.
8. Shi, X., Zhuang, Y., Chen, Z., Xu, M., Kuang, J., Sun, X.L., Gao, L., Kuang, X., Zhang, H., Li, W., et al. (2022). Hierarchical deployment of Tbx3 dictates the identity of hypothalamic KNDy neurons to control puberty onset. Sci. Adv. *8*, eabq2987. https://doi.org/10.1126/sciadv.abq2987.
9. Peng, C.Y., Mukhopadhyay, A., Jarrett, J.C., Yoshikawa, K., and Kessler, J.A. (2012). BMP receptor 1A regulates development of hypothalamic circuits critical for feeding behavior. J. Neurosci. *32*, 17211–17224. https://doi.org/10.1523/JNEUROSCI.2484-12.2012.
10. Zhang, Y.H., Xu, M., Shi, X., Sun, X.L., Mu, W., Wu, H., Wang, J., Li, S., Su, P., Gong, L., et al. (2021). Cascade diversification directs generation of neuronal diversity in the hypothalamus. Cell Stem Cell *28*, 1483–1499.e8. https://doi.org/10.1016/j.stem.2021.03.020.
11. Benevento, M., Hökfelt, T., and Harkany, T. (2022). Ontogenetic rules for the molecular diversification of hypothalamic neurons. Nat. Rev. Neurosci. *23*, 611–627. https://doi.org/10.1038/s41583-022-00615-3.
12. Fyffe, S.L., Neul, J.L., Samaco, R.C., Chao, H.T., Ben-Shachar, S., Moretti, P., McGill, B.E., Goulding, E.H., Sullivan, E., Tecott, L.H., and Zoghbi, H.Y. (2008). 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. https://doi.org/10.1016/j.neuron.2008.07.030.
13. Lemaire, L.A., Cao, C., Yoon, P.H., Long, J., and Levine, M. (2021). The hypothalamus predates the origin of vertebrates. Sci. Adv. *7*, eabf7452. https://doi.org/10.1126/sciadv.abf7452.
14. Pembroke, W.G., Hartl, C.L., and Geschwind, D.H. (2021). Evolutionary conservation and divergence of the human brain transcriptome. Genome Biol. *22*, 52. https://doi.org/10.1186/s13059-020-02257-z.
15. Romanov, R.A., Tretiakov, E.O., Kastriti, M.E., Zupancic, M., Häring, M., Korchynska, S., Popadin, K., Benevento, M., Rebernik, P., Lallemend, F., et al. (2020). Molecular design of hypothalamus development. Nature *582*, 246–252. https://doi.org/10.1038/s41586-020-2266-0.
16. Zhou, X., Lu, Y., Zhao, F., Dong, J., Ma, W., Zhong, S., Wang, M., Wang, B., Zhao, Y., Shi, Y., et al. (2022). Deciphering the spatial-temporal transcriptional landscape of human hypothalamus development. Cell Stem Cell *29*, 328–343.e5. https://doi.org/10.1016/j.stem.2021.11.009.
17. Kim, D.W., Washington, P.W., Wang, Z.Q., Lin, S.H., Sun, C., Ismail, B.T., Wang, H., Jiang, L., and Blackshaw, S. (2020). The cellular and molecular landscape of hypothalamic patterning and differentiation from embryonic to late postnatal development. Nat. Commun. *11*, 4360. https://doi.org/10.1038/s41467-020-18231-z.
18. Ma, C., Li, C., Ma, H., Yu, D., Zhang, Y., Zhang, D., Su, T., Wu, J., Wang, X., Zhang, L., et al. (2022). Pan-cancer surveys indicate cell cycle-related roles of primate-specific genes in tumors and embryonic cerebrum. Genome Biol. *23*, 251. https://doi.org/10.1186/s13059-022-02821-9.
19. Shao, Y., Chen, C., Shen, H., He, B.Z., Yu, D., Jiang, S., Zhao, S., Gao, Z., Zhu, Z., Chen, X., et al. (2019). GenTree, an integrated resource for analyzing the evolution and function of primate-specific coding genes. Genome Res. *29*, 682–696. https://doi.org/10.1101/gr.238733.118.
20. Al-Naama, N., Mackeh, R., and Kino, T. (2020). C_2H_2-Type Zinc Finger Proteins in Brain Development, Neurodevelopmental, and Other Neuropsychiatric Disorders: Systematic Literature-Based Analysis. Front. Neurol. *11*, 32. https://doi.org/10.3389/fneur.2020.00032.
21. Ding, X., Fragoza, R., Singh, P., Zhang, S., Yu, H., and Schimenti, J.C. (2020). Variants in RABL2A causing male infertility and ciliopathy. Hum. Mol. Genet. *29*, 3402–3411. https://doi.org/10.1093/hmg/ddaa230.
22. Li, Z., Tyler, W.A., Zeldich, E., Baró, G.S., Okamoto, M., Gao, T.L.Y., Li, M.F., Sestan, N., and Haydar, T.F. (2020). Transcriptional priming as a conserved mechanism of lineage diversification in the developing mouse and human neocortex. Sci. Adv. *6*, eabd2068. https://doi.org/10.1126/sciadv.abd2068.
23. Allen, N.J., and Lyons, D.A. (2018). Glia as architects of central nervous system formation and function. Science *362*, 181–185. https://doi.org/10.1126/science.aat0473.
24. Díaz, C., Morales-Delgado, N., and Puelles, L. (2014). Ontogenesis of peptidergic neurons within the genoarchitectonic map of the mouse hypothalamus. Front. Neuroanat. *8*, 162. https://doi.org/10.3389/fnana.2014.00162.
25. Ferran, J.L., Puelles, L., and Rubenstein, J.L.R. (2015). Molecular codes defining rostrocaudal domains in the embryonic mouse hypothalamus. Front. Neuroanat. *9*, 46. https://doi.org/10.3389/fnana.2015.00046.
26. Kim, D.W., Place, E., Chinnaiya, K., Manning, E., Sun, C., Dai, W., Groves, I., Ohyama, K., Burbridge, S., Placzek, M., and Blackshaw, S. (2022).

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

Single-cell analysis of early chick hypothalamic development reveals that hypothalamic cells are induced from prethalamic-like progenitors. Cell Rep. *38*, 110251. https://doi.org/10.1016/j.celrep.2021.110251.

27. Chen, A., Liao, S., Cheng, M., Ma, K., Wu, L., Lai, Y., Qiu, X., Yang, J., Xu, J., Hao, S., et al. (2022). Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. Cell *185*, 1777–1792.e21. https://doi.org/10.1016/j.cell.2022.04.003.
28. Herb, B.R., Glover, H.J., Bhaduri, A., Colantuoni, C., Bale, T.L., Siletti, K., Hodge, R., Lein, E., Kriegstein, A.R., Doege, C.A., and Ament, S.A. (2023). Single-cell genomics reveals region-specific developmental trajectories underlying neuronal diversity in the human hypothalamus. Sci. Adv. *9*, eadf6251. https://doi.org/10.1126/sciadv.adf6251.
29. Newman, E.A., Kim, D.W., Wan, J., Wang, J., Qian, J., and Blackshaw, S. (2018). Foxd1 is required for terminal differentiation of anterior hypothalamic neuronal subtypes. Dev. Biol. *439*, 102–111. https://doi.org/10.1016/j.ydbio.2018.04.012.
30. Radyushkin, K., Anokhin, K., Meyer, B.I., Jiang, Q.H., Alvarez-Bolado, G., and Gruss, P. (2005). Genetic ablation of the mammillary bodies in the Foxb1 mutant mouse leads to selective deficit of spatial working memory. Eur. J. Neurosci. *21*, 219–229. https://doi.org/10.1111/j.1460-9568.2004.03844.x.
31. La Manno, G., Siletti, K., Furlan, A., Gyllborg, D., Vinsland, E., Mossi Albiach, A., Mattsson Langseth, C., Khven, I., Lederer, A.R., Dratva, L.M., et al. (2021). Molecular architecture of the developing mouse brain. Nature *596*, 92–96. https://doi.org/10.1038/s41586-021-03775-x.
32. Briscoe, J., Pierani, A., Jessell, T.M., and Ericson, J. (2000). A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. Cell *101*, 435–445. https://doi.org/10.1016/s0092-8674(00)80853-3.
33. Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics *19*, 477. https://doi.org/10.1186/s12864-018-4772-0.
34. Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H., and Trapnell, C. (2017). Reversed Graph Embedding Resolves Complex Single-Cell Developmental Trajectories (Cold Spring Harbor Laboratory).
35. Butler, S.J., and Bronner, M.E. (2015). From classical to current: Analyzing peripheral nervous system and spinal cord lineage and fate. Dev. Biol. *398*, 135–146. https://doi.org/10.1016/j.ydbio.2014.09.033.
36. Martínez, O., and Reyes-Valdés, M.H. (2008). Defining diversity, specialization, and gene specificity in transcriptomes through information theory. Proc. Natl. Acad. Sci. USA *105*, 9709–9714. https://doi.org/10.1073/pnas.0803479105.
37. Wright, C.J., Smith, C.W.J., and Jiggins, C.D. (2022). Alternative splicing as a source of phenotypic diversity. Nat. Rev. Genet. *23*, 697–710. https://doi.org/10.1038/s41576-022-00514-4.
38. Iwata, R., Casimir, P., Erkol, E., Boubakar, L., Planque, M., López, I.M.G., Ditkowska, M., Gaspariunaite, V., Beckers, S., Remans, D., et al. (2023). Mitochondria metabolism sets the species-specific tempo of neuronal development. Science *379*, eabn4705. https://doi.org/10.1126/science.abn4705.
39. Risold, P.Y., Thompson, R.H., and Swanson, L.W. (1997). The structural organization of connections between hypothalamus and cerebral cortex. Brain Res. Brain Res. Rev. *24*, 197–254. https://doi.org/10.1016/S0165-0173(97)00007-6.
40. Rempel-Clower, N.L., and Barbas, H. (1998). Topographic organization of connections between the hypothalamus and prefrontal cortex in the rhesus monkey. J. Comp. Neurol. *398*, 393–419. https://doi.org/10.1002/(Sici)1096-9861(19980831)398:3<393::Aid-Cne7>3.0.Co;2-V.
41. Alvarez-Bolado, G. (2019). Development of neuroendocrine neurons in the mammalian hypothalamus. Cell Tissue Res. *375*, 23–39. https://doi.org/10.1007/s00441-018-2859-1.
42. Zhou, J.N., and Fang, H. (2018). Transcriptional regulation of corticotropin-releasing hormone gene in stress response. IBRO Rep. *5*, 137–146. https://doi.org/10.1016/j.ibror.2018.08.003.
43. Liu, F., Placzek, M., and Xu, H. (2013). Axon guidance effect of classical morphogens Shh and BMP7 in the hypothalamo-pituitary system. Neurosci. Lett. *553*, 104–109. https://doi.org/10.1016/j.neulet.2013.08.027.
44. Liu, F., Pogoda, H.M., Pearson, C.A., Ohyama, K., Löhr, H., Hammerschmidt, M., and Placzek, M. (2013). Direct and indirect roles of Fgf3 and Fgf10 in innervation and vascularisation of the vertebrate hypothalamic neurohypophysis. Development *140*, 1111–1122. https://doi.org/10.1242/dev.080226.
45. Duittoz, A.H., Forni, P.E., Giacobini, P., Golan, M., Mollard, P., Negrón, A.L., Radovick, S., and Wray, S. (2022). Development of the gonadotropin-releasing hormone system. J. Neuroendocrinol. *34*, e13087. https://doi.org/10.1111/jne.13087.
46. Silverman, A.J., Krey, L.C., and Zimmerman, E.A. (1979). A comparative study of the luteinizing hormone releasing hormone (LHRH) neuronal networks in mammals. Biol. Reprod. *20*, 98–110. https://doi.org/10.1093/biolreprod/20.1.98.
47. Casoni, F., Malone, S.A., Belle, M., Luzzati, F., Collier, F., Allet, C., Hrabovszky, E., Rasika, S., Prevot, V., Chédotal, A., and Giacobini, P. (2016). Development of the neurons controlling fertility in humans: new insights from 3D imaging and transparent fetal brains. Development *143*, 3969–3981. https://doi.org/10.1242/dev.139444.
48. Svensson, E., Apergis-Schoute, J., Burnstock, G., Nusbaum, M.P., Parker, D., and Schiöth, H.B. (2018). General Principles of Neuronal Co-transmission: Insights From Multiple Model Systems. Front. Neural Circuits *12*, 117. https://doi.org/10.3389/fncir.2018.00117.
49. Björklund, A., and Dunnett, S.B. (2007). Dopamine neuron systems in the brain: an update. Trends Neurosci. *30*, 194–202. https://doi.org/10.1016/j.tins.2007.03.006.
50. Romanov, R.A., Zeisel, A., Bakker, J., Girach, F., Hellysaz, A., Tomer, R., Alpár, A., Mulder, J., Clotman, F., Keimpema, E., et al. (2017). Molecular interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes. Nat. Neurosci. *20*, 176–188. https://doi.org/10.1038/nn.4462.
51. Simon, H.H., Saueressig, H., Wurst, W., Goulding, M.D., and O'Leary, D.D.M. (2001). Fate of midbrain dopaminergic neurons controlled by the engrailed genes. J. Neurosci. *21*, 3126–3134. https://doi.org/10.1523/JNEUROSCI.21-09-03126.2001.
52. Ge, M., Sheikhshahrokh, A., Shi, X., Zhang, Y.H., Xu, Z., and Wu, Q.F. (2023). A Spacetime Odyssey of Neural Progenitors to Generate Neuronal Diversity. Neurosci. Bull. *39*, 645–658. https://doi.org/10.1007/s12264-022-00956-0.
53. Vieira, C., Pombero, A., García-Lopez, R., Gimeno, L., Echevarria, D., and Martínez, S. (2010). Molecular mechanisms controlling brain development: an overview of neuroepithelial secondary organizers. Int. J. Dev. Biol. *54*, 7–20. https://doi.org/10.1387/ijdb.092853cv.
54. Sousa, V.H., and Fishell, G. (2010). Sonic hedgehog functions through dynamic changes in temporal competence in the developing forebrain. Curr. Opin. Genet. Dev. *20*, 391–399. https://doi.org/10.1016/j.gde.2010.04.008.
55. Barrière, A., and Bertrand, V. (2020). Neuronal specification in C. elegans: combining lineage inheritance with intercellular signaling. J. Neurogenet. *34*, 273–281. https://doi.org/10.1080/01677063.2020.1781850.
56. Quarta, C., Fisette, A., Xu, Y., Colldén, G., Legutko, B., Tseng, Y.T., Reim, A., Wierer, M., De Rosa, M.C., Klaus, V., et al. (2019). Functional identity of hypothalamic melanocortin neurons depends on Tbx3. Nat. Metab. *1*, 222–235. https://doi.org/10.1038/s42255-018-0028-1.
57. Galazzi, E., Duminuco, P., Moro, M., Guizzardi, F., Marazzi, N., Sartorio, A., Avignone, S., Bonomi, M., Persani, L., and Bonati, M.T. (2018). Hypogonadotropic hypogonadism and pituitary hypoplasia as recurrent features in Ulnar-Mammary syndrome. Endocr. Connect. *7*, 1432–1441. https://doi.org/10.1530/EC-18-0486.

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

58. Price, S.R., De Marco Garcia, N.V.D., Ranscht, B., and Jessell, T.M. (2002). Regulation of motor neuron pool sorting by differential expression of type II cadherins. Cell *109*, 205–216. https://doi.org/10.1016/S0092-8674(02)00695-5.

59. Isaacson, J.S., and Scanziani, M. (2011). How inhibition shapes cortical activity. Neuron *72*, 231–243. https://doi.org/10.1016/j.neuron.2011.09.027.

60. Grueber, W.B., and Sagasti, A. (2010). Self-avoidance and Tiling: Mechanisms of Dendrite and Axon Spacing. Cold Spring Harb. Perspect. Biol. *2*, a001750. https://doi.org/10.1101/cshperspect.a001750.

61. Zhang, J.Y., Obál, F., Zheng, T., Fang, J.D., Taishi, P., and Krueger, J.M. (1999). Intrapreoptic microinjection of GHRH or its antagonist alters sleep in rats. J. Neurosci. *19*, 2187–2194. https://doi.org/10.1523/JNEUROSCI.19-06-02187.1999.

62. Liao, F., Taishi, P., Churchill, L., Urza, M.J., and Krueger, J.M. (2010). Localized Suppression of Cortical Growth Hormone-Releasing Hormone Receptors State-Specifically Attenuates Electroencephalographic Delta Waves. J. Neurosci. *30*, 4151–4159. https://doi.org/10.1523/JNEUROSCI.6047-09.2010.

63. Manfredi-Lozano, M., Leysen, V., Adamo, M., Paiva, I., Rovera, R., Pignat, J.M., Timzoura, F.E., Candlish, M., Eddarkaoui, S., Malone, S.A., et al. (2022). GnRH replacement rescues cognition in Down syndrome. Science *377*, eabq4515. https://doi.org/10.1126/science.abq4515.

64. Haider, S., Ballester, B., Smedley, D., Zhang, J., Rice, P., and Kasprzyk, A. (2009). BioMart Central Portal—unified access to biological data. Nucleic Acids Res. *37*, W23–W27. https://doi.org/10.1093/nar/gkp265.

65. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

66. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Camb) *2*, 100141. https://doi.org/10.1016/j.xinn.2021.100141.

67. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. *10*, 1523. https://doi.org/10.1038/s41467-019-09234-6.

68. Dai, M., Pei, X., and Wang, X.-J. (2022). Accurate and fast cell marker gene identification with COSG. Brief. Bioinform. *23*, bbab579. https://doi.org/10.1093/bib/bbab579.

69. Crow, M., Paul, A., Ballouz, S., Huang, Z.J., and Gillis, J. (2018). Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. Nat. Commun. *9*, 884. https://doi.org/10.1038/s41467-018-03282-0.

70. Bergen, V., Lange, M., Peidli, S., Wolf, F.A., and Theis, F.J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. Nat. Biotechnol. *38*, 1408–1414. https://doi.org/10.1038/s41587-020-0591-3.

71. Farrell, J.A., Wang, Y., Riesenfeld, S.J., Shekhar, K., Regev, A., and Schier, A.F. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. Science *360*, eaar3131. https://doi.org/10.1126/science.aar3131.

72. Biancalani, T., Scalia, G., Buffoni, L., Avasthi, R., Lu, Z., Sanger, A., Tokcan, N., Vanderburg, C.R., Segerstolpe, Å., Zhang, M., et al. (2021). Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. Nat. Methods *18*, 1352–1362. https://doi.org/10.1038/s41592-021-01264-7.

73. Dann, E., Henderson, N.C., Teichmann, S.A., Morgan, M.D., and Marioni, J.C. (2022). Differential abundance testing on single-cell data using k-nearest neighbor graphs. Nat. Biotechnol. *40*, 245–253. https://doi.org/10.1038/s41587-021-01033-z.

74. Van de Sande, B., Flerin, C., Davie, K., De Waegeneer, M., Hulselmans, G., Aibar, S., Seurinck, R., Saelens, W., Cannoodt, R., Rouchon, Q., et al. (2020). A scalable SCENIC workflow for single-cell gene regulatory network analysis. Nat. Protoc. *15*, 2247–2276. https://doi.org/10.1038/s41596-020-0336-2.

75. Wang, J., Li, T., Wang, J.L., Xu, Z., Meng, W., and Wu, Q.F. (2020). Talpid3-Mediated Centrosome Integrity Restrains Neural Progenitor Delamination to Sustain Neurogenesis by Stabilizing Adherens Junctions. Cell Rep. *33*, 108495. https://doi.org/10.1016/j.celrep.2020.108495.

76. Tosches, M.A., Yamawaki, T.M., Naumann, R.K., Jacobi, A.A., Tushev, G., and Laurent, G. (2018). Evolution of pallium, hippocampus, and cortical cell types revealed by single-cell transcriptomics in reptiles. Science *360*, 881–888. https://doi.org/10.1126/science.aar4237.

77. Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature *566*, 496–502. https://doi.org/10.1038/s41586-019-0969-x.

78. Di Bella, D.J., Habibi, E., Stickels, R.R., Scalia, G., Brown, J., Yadollahpour, P., Yang, S.M., Abbate, C., Biancalani, T., Macosko, E.Z., et al. (2021). Molecular logic of cellular diversification in the mouse cerebral cortex. Nature *595*, 554–559. https://doi.org/10.1038/s41586-021-03670-5.

79. Workman, A.D., Charvet, C.J., Clancy, B., Darlington, R.B., and Finlay, B.L. (2013). Modeling Transformations of Neurodevelopmental Sequences across Mammalian Species. J. Neurosci. *33*, 7368–7383. https://doi.org/10.1523/JNEUROSCI.5746-12.2013.

80. Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A.M.M., Pletikos, M., Meyer, K.A., Sedmak, G., et al. (2011). Spatio-temporal transcriptome of the human brain. Nature *478*, 483–489. https://doi.org/10.1038/nature10523.

81. Ma, S., Skarica, M., Li, Q., Xu, C., Risgaard, R.D., Tebbenkamp, A.T.N., Mato-Blanco, X., Kovner, R., Krsnik, Ž., de Martin, X., et al. (2022). Molecular and cellular evolution of the primate dorsolateral prefrontal cortex. Science *377*, eabo7257. https://doi.org/10.1126/science.abo7257.

82. Hodge, R.D., Bakken, T.E., Miller, J.A., Smith, K.A., Barkan, E.R., Graybuck, L.T., Close, J.L., Long, B., Johansen, N., Penn, O., et al. (2019). Conserved cell types with divergent features in human versus mouse cortex. Nature *573*, 61–68. https://doi.org/10.1038/s41586-019-1506-7.

83. Hook, P.W., and McCallion, A.S. (2020). Leveraging mouse chromatin data for heritability enrichment informs common disease architecture and reveals cortical layer contributions to schizophrenia. Genome Res. *30*, 528–539. https://doi.org/10.1101/gr.256578.119.

84. Finucane, H.K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P.-R., Anttila, V., Xu, H., Zang, C., Farh, K., et al. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat. Genet. *47*, 1228–1235. https://doi.org/10.1038/ng.3404.

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Goat polyclonal anti-Foxb1	Abcam	Cat#ab5274; RRID: AB_304806
Rabbit polyclonal anti-Barhl1	Novusbio	Cat#NBP1-86513; RRID: AB_11034569
Sheep polyclonal anti-ONECUT1	R&D Systems	Cat#AF6277; RRID: AB_10640363
Mouse monoclonal anti-Lhx1	DSHB	Cat#4F-2; RRID: AB_531784
Rabbit monoclonal anti-GnRH	Abcam	Cat#ab281844; RRID: AB_3678596
Rabbit polyclonal anti-GHRH	Abcam	Cat#ab187512; RRID: AB_3678595
Rat polyclonal anti-TH	Oasis Biofarm	Cat#OB-PGP064; RRID: AB_2938883
Rabbit polyclonal anti-TH	Merk-Millipore	Cat#T8700-1VL; RRID: AB_1080430
Rabbit polyclonal anti-Avp	Immunostar	Cat#20069; RRID: AB_572219
Rat polyclonal anti-Avp	Oasis Biofarm	Cat#OB-PRB033; RRID: AB_2938914
Rabbit polyclonal anti-Nkx2.1	Abcam	Cat#ab76013; RRID: AB_1310784
Rabbit polyclonal anti-mCherry/tdTomato	Oasis Biofarm	Cat#OB-PRB013; RRID: AB_2934227
Cy2 AffiniPure Donkey Anti-Mouse IgG (H+L)	Jackson ImmunoResearch Labs	Cat#715-225-150; RRID: AB_2340826
Cy2 AffiniPure Donkey Anti-Sheep IgG (H+L)	Jackson ImmunoResearch Labs	Cat# 713-005-003; RRID: AB_2340703
Cy3 AffiniPure Donkey Anti-Goat IgG (H+L)	Jackson ImmunoResearch Labs	Cat#705-165-003; RRID: AB_2340411
Cy3 AffiniPure Donkey Anti-Rat IgG (H+L)	Jackson ImmunoResearch Labs	Cat#712-007-003; RRID: AB_2340634
Cy3 AffiniPure Donkey Anti-Rabbit IgG (H+L)	Jackson ImmunoResearch Labs	Cat#711-165-152; RRID: AB_2307443
Cy5 AffiniPure Donkey Anti-Rabbit IgG (H+L)	Jackson ImmunoResearch Labs	Cat#711-175-152; RRID: AB_2340607
Goat Anti-Mouse IgG Antibody (H+L), Biotinylated	Vector Laboratories	Cat#BA-9200-1.5; RRID: AB_2336171
Chemicals, Peptides, and Recombinant Proteins		
Tamoxifen	Sigma-Aldrich	Cat#T5648
DAPI	Sigma-Aldrich	Cat#D9542
Papain	Worthington	Cat#LK003178
DNaseI	Worthington	Cat#LK003172
RNase inhibitor	Clontech	Cat#2313A
Glutamax	Life Technologies	Cat#35050061
B27 supplement	Thermo Fisher Scientific	Cat#17504044
Hibernate-E media	Life Technologies	Cat#A1247601
Trypsin inhibitor	Sigma-Aldrich	Cat#T9128
Poly-D-lysine hydrobromide	Sigma-Aldrich	Cat#27964-99-4
Dulbecco's Modified Eagle Medium (DMEM)	Gibco	Cat#11965092
FBS	Gibco	Cat#10099141
2,2,2-tribromoethanol	Macklin	Cat#T903147
Target retrieval solution	DAKO	Cat#s1699
Tween 20	Scientific Chemical	Cat#9005-64-5
SDS	Amresco	Cat#O227
Polyvinyl alcohol	Sigma-Aldrich	Cat#P8136
40-µm cell strainer	BD Bioscience	Cat#352340
Ampure XP beads	Beckman-Counlter	Cat#A63881
DMEM/F12 medium	Gibco	Cat#11320033
Exonuclease I	NEB	Cat#M0293L
Tissue-Tek O.C.T. Compound	Sakura	Cat#4583
Critical Commercial Assays		
QubitTM ssDNA Assay Kit	Thermo Fisher Scientific	Cat#Q10212

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Agilent high-sensitivity DNA Kit	Agilent Technologies	Cat#5067-4626
DNBelab C4 Single-Cell Library Prep Set	MGI	Cat#1000021082
Qubit dsDNA HS assay kit	Invitrogen	Cat#Q32854
Hot Start DNA Polymerase	QIAGEN	Cat#203603
Single-cell 3’ Library and Gel Bead Kit V3	10X Genomics	Cat#1000075
The STOmics Gene Expression kit S1	BGI	Cat#1000028493
Deposited Data		
Single-cell RNA-seq and single-nuclei RNA-seq data	This study	CNGBdb: CNP0004213
Spatial transcriptomic data	This study	https://hypoatlas.org
Experimental Models: Cell lines		
HEK293T	ATCC	CVCL_0063
Experimental Models: Organisms/Strains		
Mouse: Rax-CreER^{T}2 line	The Jackson Laboratory	RRID: IMSR_JAX:025521
Mouse: Ai9 reporter line	The Jackson Laboratory	RRID: IMSR_JAX:007909
Mouse: Ai14 reporter line	The Jackson Laboratory	RRID: IMSR_JAX:007914
Mouse: vGat-IRES-Cre line	The Jackson Laboratory	RRID: IMSR_JAX:028862
Mouse: Tbx3-CreER^{T}2 line	See Shi et al.[8] Qing-Feng Wu’s lab, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences	N/A
Recombinant DNA		
Plasmid: pCDH-EF1α	Addgene	Cat#72266
Plasmid: pCDH-EF1α-EGFP	This paper	N/A
Plasmid: pCDH-EF1α-mScarlet	This paper	N/A
Plasmid: pCDH-EF1α-ONECUT1-HA-T2A-EGFP	This paper	N/A
Oligonucleotides		
Oligonucleotides used for smFISH	Table S2	N/A
Software and Algorithms		
Leica Application Suite X	Leica	N/A
ImageJ	National Institute of Health	https://imagej.net/; RRID: SCR_003070
GraphPad Prism 9	GraphPad Software	http://www.graphpad.com/; RRID: SCR_002798
Imaris v9.0.1	Bitplane	http://www.bitplane.com/imaris/imaris; RRID: SCR_007370
Microsoft Excel	Microsoft	https://www.microsoft.com/en-gb/; RRID: SCR_016137
R v4.0.2	R Foundation	https://www.r-project.org; RRID: SCR_001905
Slingshot v2.12.0	See Street et al.[33]	https://github.com/kstreet13/slingshot
BiomaRt v0.9.1	See Haider et al.[64]	https://bioconductor.org/packages/release/bioc/html/biomaRt.html
Seurat v4.0.1	See Hao et al.[65]	https://github.com/satijalab/seurat
SeuratWrappers v0.3.1	Rahul Satija’s Lab, New York University	https://github.com/satijalab/seurat-wrappers
STAR v2.7.4a	Alexander Dobin’s Lab, Cold Spring Harbor Laboratory	https://github.com/alexdobin/STAR
clusterProfiler v4.0	See Wu et al.[66]	https://github.com/YuLab-SMU/clusterProfiler
Metascape	See Zhou et al.[67]	http://metascape.org/
COSGR v0.9.0	See Dai et al.[68]	https://github.com/genecell/COSGR
MetaNeighbor v1.9.1	See Crow et al.[69]	https://github.com/gillislab/MetaNeighbor
Monocle 3 v1.3.0	Cole Trapnell’s Lab, University of Washington	https://cole-trapnell-lab.github.io/monocle3/

(*Continued on next page*)

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
scVelo v0.2.4	See Bergen et al.[70]	https://github.com/theislab/scvelo
URD v1.1.1	See Farrell et al.[71]	https://github.com/farrellja/URD
ggraph v2.1.0	See Thomas Lin Pedersen	https://github.com/thomasp85/ggraph
Tangram v1.0.4	See Biancalani et al.[72]	https://github.com/broadinstitute/Tangram
miloR v0.99.18	See Dann et al.[73]	https://github.com/MarioniLab/miloR
scCODA v0.1.9	Fabian Theis's Lab, Institute of Computational Biology, Helmholtz Center Munich	https://github.com/theislab/scCODA/tree/master
Cacoa v 0.4.0	Peter Kharchenko's Lab, Harvard Medical School	https://github.com/kharchenkolab/cacoa
Speckle v0.03	Alicia Oshlack's lab, Peter MacCallum Cancer Centre	https://github.com/Oshlack/speckle
motifmatchr v1.22.0	William Greenleaf's Lab, Stanford University	https://github.com/GreenleafLab/motifmatchr
SCENIC v0.12.0	See Van et al.[74]	https://github.com/aertslab/SCENIC
PISA	BGI-Shenzhen	https://github.com/shiquan/PISA
Sambamba v0.7	Pjotr Prins's labs, St. Petersburg State University	https://github.com/biod/sambamba
Cell ranger v3.0.2	10xGenomics	https://support.10xgenomics.com/
Adobe Photoshop	Adobe	https://www.adobe.com/products/photoshop.html; RRID: SCR_014199
Adobe Illustrator	Adobe	http://www.adobe.com/products/illustrator.html; RRID: SCR_010279
Other		
Lecia M205 fluorescence stereomicroscope	Leica	N/A
Leica SP8 confocal microscope	Leica	N/A
Leica CM3050 S Cryostat	Leica	N/A
Illumina NovaSeq 6000	Illumina Inc	N/A
DNBSEQ™ technology platform	MGI	N/A
BD FACSAria II flow cytometer	BD Bioscience	N/A

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human sample collection

Human fetuses at post-conception weeks (PCW) 5, 6, 7, 9, 10, 12, 13, 16 and 20 were collected at Baoding Second Central Hospital and quickly delivered on ice to the Institute of Genetics and Developmental Biology, Chinese Academy of Sciences. Each developmental stage of the human hypothalamus is represented by 1-2 samples only, with key stages (e.g. PCW6 and PCW10) each including two samples. Among these, PCW6 samples include both sexes and both PCW10 samples are male, while the remaining samples are female. The fetal hypothalamus was freshly microdissected for single-cell RNA sequencing (scRNA-seq) or frozen in liquid nitrogen to perform single-nucleus RNA sequencing (snRNA-seq). Prior to sequencing, samples were screened for RNA quality by isolating total RNA from a small piece of tissue and measuring the RNA Integrity Number (RIN) using the Agilent 2100 Bioanalyzer. Only samples with a RIN $\geq$ 6 were selected for single-cell or single-nucleus RNA sequencing. An informed consent document was signed by each pregnant female before human fetus collection. All experimental procedures were reviewed and approved by the Medical and Institutional Ethics Committee (IGDB-2020-IRB-001). Human adult hypothalamic samples were acquired from the Netherlands Brain Bank (NBB), with consent from donors or their next of kin for brain autopsy, access to medical records and research use of brain tissues. Permission to collect human adult brain material was granted by the Medical Ethics Committee of the VU University Medical Center, Amsterdam, the Netherlands.

Macaque fetal hypothalamus collection

Macaque samples at PCW5, 8, 11 from natural conception were collected at the Beijing Institute of Xieerxin Biology Resource under the supervision of a veterinarian. Due to resource limitations during COVID-19 pandemic, macaque hypothalamic tissues included only one sample per developmental stage, with PCW5 being the only female specimen and the rest being male. The cesarean section was expeditiously performed, with a delivery time of fewer than 5 min, following the general anesthesia of the dam under sterile conditions. After the delivery of fetus, the dam received uterine and skin closure in accordance with international standard operating

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

procedures, and postoperative care including pain management. The collected brains were maintained in ice-cold DMEM/F12 medium (Gibco, 11320033) during the microdissection process. Under the microscope, the hypothalamus was meticulously isolated according to the Brain Maps Site (http://www.brainmaps.org). All procedures were conducted in compliance with prior approval by the Institutional Animal Care and Use Committee of Xieerxin Biology Resource (XEX20212025).

Mice

To collect prenatal and perinatal brains from wild-type mice, we ordered male and female C57BL/6N mice from SPF Biotechnology Co. Ltd (Beijing, China), bred them and checked the vaginal plug to determine the time of pregnancy. The noon of the day when the vaginal plug was detected was designated as embryonic day 0.5, and the mouse embryos at the age of E10.5, E12.5, E13.5 and P0 were obtained by cesarean section.

To conduct lineage tracing of Tbx3$^+$ progenitor domain, we crossed *Tbx3-CreER*^{T2} male mice with *Ai9* (Stock No. 007909) female mice, and intraperitoneally injected a single dose of tamoxifen (132 mg/kg body weight) into timed pregnant females at E9 or E10. TdTomato-labeled embryos at E11 were collected for direct fixation with cold 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS). To obtain postnatal brains at P14 and P30, we recovered live embryos at E18-E19 by cesarean section, cared for the pups with foster female animals, and sacrificed the mice for further analyses.

Animals were maintained on a 12-hour (hr) light-dark cycle with *ad libitum* access to food and water. All animal procedures in this study were performed according to protocols approved by the Institutional Animal Care and Use Committee at the Institute of Genetics and Developmental Biology, Chinese Academy of Sciences.

METHOD DETAILS

Immunohistochemistry

Mice were anesthetized by hypothermia and then transcardially perfused with saline followed by 4% PFA in PBS. Mouse brains were immediately dissected, post-fixed for 4-8 hrs in 4% PFA at 4°C, and subsequently cryo-protected in 20% sucrose in PBS for 12 hrs followed by 30% sucrose for 12 hrs. Human fetal samples were fixed for 1-3 days in 4% PFA at 4°C and subsequently cryo-protected in 20% sucrose 24 hrs followed by 30% sucrose for 24-48 hrs. Tissue blocks were prepared by embedding samples in Tissue-Tek O.C.T. Compound (Sakura 4583). Brain sections (20-40 μm in thickness) were prepared using a cryostat microtome (Leica, CM3050S) and stored in a -80°C freezer. For immunostaining, hypothalamic tissue sections were washed with 1×TBS (pH=7.4, containing 3 mM KCl, 25 mM Trizma base, and 137 mM NaCl). To unmask epitopes, retrieved using Target Retrieval Solution (DAKO, S1699) at 95°C and pre-blocked with 1×TBS++ (TBS containing 5% donkey serum and 0.3% Triton X-100) for 1 hr at room temperature (RT), followed by incubation with primary antibodies diluted in TBS++ overnight at 4°C. Human adult postmortem brain tissues were immersed in 10% phosphate-buffer formalin at RT after autopsy. Then, hypothalamic tissue was embedded in paraffin and sectioned in a rostral caudal orientation (6 μm in thickness), using a microtome. Anatomical orientation of the PVN and SON was determined by Nissl staining of every 100th section available, and further validated by evaluation of AVP-ir. For histological procedure, sections were mounted in Superfrost+ slides (Thermo Scientific), and dried on a 37°C heating plate for 48 hrs. Removal of paraffin was achieved by immersion of sections in 100% xylene, followed by rehydration in grading ethanol (100%–50%), and rinsed in distilled water. Heat induced antigen retrieval was performed using microwave treatment (10 min at 700 W) in citrate buffer (82.5 mM sodium citrate dihydrate and 17.5 mM citric acid; pH 6.0). After cooling, sections were washed in 1×TBS, treated with 3% hydrogen peroxide in SUMI buffer (0.25% gelatine, 0.5% Triton X-100 in TBS (pH 7.6)). Sections were then washed in 1×TBS, and incubated with their primary antibodies for 1 hr at RT, followed by overnight incubation at 4°C.

The primary antibodies used in this study included goat anti-Foxb1 (Abcam; ab5274; 1:500), rabbit anti-Barhl1 (Novusbio; NBP1-86513; 1:500), sheep anti-ONECUT1 (R&D; AF6277; 1:300), mouse anti-LHX1 (DSHB; 4F-2; 1:10), rabbit anti-GnRH (Abcam; ab281844; 1:500), rabbit anti-GHRH (Abcam; ab187512; 1:300), Rat anti-TH (Oasis Biofarm; OB-PGP064; 1:500), rabbit anti-TH (Merck, T8700-1VL) (adult humans exclusively), rabbit anti-Avp (Immunostar; 20069; 1:500), rabbit anti-Avp (Oasis Biofarm; OB-PRB033; 1:500), mouse anti-Avp (produced by Netherlands Institute for Neuroscience, a gift from dr. F.W. Leuween, OT-A-I-28) (adult humans exclusively), rabbit anti-Nkx2.1 (Abcam; ab76013; 1:300), rabbit anti-mCherry/tdTomato (Oasis Biofarm; OB-PRB013; 1:1000). After the primary antibody incubation, the brain sections were washed three times with 1×TBS and incubated with the following secondary antibodies for 2 hrs at RT: anti-mouse Cy2, anti-sheep Cy2, anti-rat Cy3, anti-goat Cy3, anti-rabbit Cy3 and anti-rabbit Cy5 (Donkey; Jackson ImmunoResearch; 1:500). For adult human hypothalamic sections, sections were washed and incubated with biotinylated secondary antibody and avidin–biotin complex (1:400 goat anti-mouse IgG, BA-9200-1.5, Vector Laboratories) for 1 hr. Sections were then rinsed and incubated with a corresponding fluorescent secondary anti-rabbit antibody (against TH) and streptavidin-fluorescence for 1 hr. Sections were once again rinsed with TBS, followed by a DAPI counterstaining (1:5000, 62248, ThermoFischer).

Single-molecule fluorescent in Situ hybridization (smFISH)

To detect the mRNA expression pattern of *SHH*, *RSPO2* and *FGF10* in human brains as well as *Pomc*, *Npy*, *Tac2*, *Crabp1*, *Ghrh*, *Slc17a6* and *Slc32a1* in mouse hypothalamus, we used the hybridization chain reaction (HCR) approach and designed mRNA probes targeting their coding sequence and 3′ untranslated region as previously described.[8,10] The sequences of all HCR probes are listed in Table S2 and all of them were synthesized by Sangon Biotech, China. The brain sections were permeabilized in 70% ethanol for

4~12 hrs at 4°C, treated with 0.5% Triton X-100 in 1×PBS for 1 hr, and digested with 10 μg/mL Protease K at 37°C for 2~3 min to improve mRNA accessibility. After two washes with 1×PBS at RT, the sections were pre-hybridized in 30% probe hybridization buffer for 1 hr and then incubated in the same buffer containing HCR probes (10 μM for each) at 37°C for 12~16 hrs. After mRNA *in situ* hybridization, the brain sections were rinsed with 30% probe hybridization buffer and washed with 5×SSCT [containing 5×sodium chloride citrate (SSC) buffer and 0.1% (v/v) Tween-20] at RT. Simultaneously, we prepared fluorescently labeled hairpins (30 pmol) by snap cooling 10 μL of 3 μM stock in hairpin storage buffer (heat at 95°C for 90s and cool to RT in darkness for 30 min). After pre-amplification at RT for 30 min, we added the working hairpin solution to the brain sections and incubated overnight (8~12 hrs) in the dark, followed by immunofluorescence staining with antibodies.

Plasmid construction, transfection and electroporation

Full-length human ONECUT1 was amplified, fused with a 1×HA tag and T2A-EGFP at the C-terminus, and cloned into the AvrII/SalI sites of the pCDH-EF1α (Plasmid No. 72266) vector to generate the ONECUT1-overexpressing plasmid. The pCDH-EF1α-EGFP and pCDH-EF1α-mScarlet vectors were used as negative controls for *in vitro* transfection and *in utero* electroporation, respectively. To validate the expression of ONECUT1-EGFP, we transfected HEK293T cells, collected cells at 3 days post-transfection and performed immunoblotting.

For *in vitro* experiments, the hypothalamus was micro-dissected from E16.5 *vGAT-Cre::Ai14* mouse brains and treated with a digestion buffer containing 10 U/mL papain, 200 U/mL DNaseI, 1×GlutaMAX, and 1×B27 supplement in Hibernate-E media. After 5 minutes of digestion at 37°C, digestion was halted with 1× trypsin inhibitor, and the cell suspension was filtered through a 40-μm cell strainer to remove aggregates. The neurons were then cultured on pre-coated cover glasses at 37°C with 5% CO_2 in Neurobasal Medium supplemented with 1×GlutaMAX, 1×B27 supplement, 100 U/mL penicillin, and 10 mg/mL streptomycin. The cover glasses were pretreated in 24-well plates with 10 μg/mL Poly-D-lysine hydrobromide overnight, followed by treatment with 5 μg/mL laminin for 2 hrs at 37°C.

We performed *in utero* electroporation as previously described.[75] Briefly, E13.5 pregnant C57BL/6N mice were anesthetized with 2,2,2-Tribromoethanol at 30 mg/kg of body weight. After cleaning the abdomen with 70% ethanol, a midline laparotomy of ~1.5 cm was performed and the uterine horns were exposed. DNA plasmids (2 mg/mL) with Fast Green (2 mg/mL; Sigma) were microinjected into a lateral ventricle of each embryo using polished glass micropipettes (Drummond). The plasmids were electroporated by delivering five square electric pulses (30 mV, 50 ms, 950 ms interval) using an electroporator (ECM-830 BTX, Harvard Apparatus). Subsequently, embryos were placed back in the abdominal cavity and the abdomen wall and skin were sutured. The brain tissue was collected at E18.5 and sliced by Vibrating Blade Microtomes for neural morphology analysis.

Single-cell isolation and scRNA-seq

To isolate single cells from fresh tissues, we collected human (PCW6, 10 and 12) and macaque (PCW5, 8 and 11) fetuses, sectioned the fetal brains in iced Hibernate-E buffer with vibratome and microdissected the hypothalamus under a stereoscope (Stemi 305, Carl Zeiss, Germany). The mouse hypothalamic tissues were isolated from *Rax-CreER*^{T2}*::Ai14* mice at E11, E14, P0 and P7 followed by fluorescence-activated cell sorting to obtain tdTomato^{+} cells or directly extracted from wild type mice at E10, E11, E12, E13, E14, E15, E16, E18, P4, P8, P14, and P45. The hypothalamic tissues across different species were transferred into a 5-mL Eppendorf tube and incubated with digestion buffer for 1.5 hrs at 37°C. The digestion was terminated by replacing the buffer with Hibernate E buffer containing 1×GlutaMAX, 0.2×B27, 0.01 mM Y27632 dihydrochloride, and 1% FBS. The tissues were then gently triturated through Pasteur pipettes with finely polished tips and washed once with Hibernate E buffer to generate single-cell suspension.

We constructed scRNA-seq libraries with the single-cell 3′ Library and Gel Bead Kit V3 (10× Genomics, 1000075) according to the instructions provided by 10× Genomics. Single-cell suspensions (300 to 600 living cells/μL determined by Count Star) were loaded onto the Chromium Single Cell Controller to generate single-cell gel beads within emulsion (GEM). The captured cells were lysed to release mRNA, which were then barcoded by reverse transcribing individual GEMs. Reverse transcription was carried out using the S1000TM Touch Thermal Cycler (Bio-Rad, USA) with a program that included incubation at 53°C for 45 min, followed by 85°C for 5 min and hold at 4°C. Subsequently, the complementary DNA (cDNA) library was generated, amplified, and subjected to quality control assessment using Agilent 4200. The scRNA-seq was further performed on the Illumina NovaSeq 6000 sequencer with a sequencing depth of at least 100,000 reads per cell and 150-base pair (bp) paired-end reads.

Single-nucleus isolation and snRNA-seq

To isolate single nuclei from frozen human samples (PCW5, 6, 7, 9, 10, 13, 16, 20), we immersed the tissues in ice-cold homogenization buffer containing 250 mM sucrose, 25 mM KCl, 5 mM $MgCl_2$, 10 mM Tris-HCl, 0.1 mM DTT, 1% BSA, 0.1% NP-40 and ribonuclease inhibitor (0.4 U/μl), followed by mechanical grinding with a glass homogenizer for 10 times and collection of single-nucleus suspension for snRNA-seq.

We took advantage of DNBSEQ technology platforms (BGI Genomics, China) and DNBelab C4 Single-Cell Library Prep Set (MGI Tech, #1000021082) for library preparation. Briefly, single-nucleus suspensions were pumped into microfluidic device for droplet generation, followed by emulsion breakage, bead collection, reverse transcription and cDNA amplification to generate barcoded libraries. The cDNA products were then sheared to short fragments with a length of 250 to 400 bp, and indexed sequencing libraries were constructed according to the manufacturer's protocol. The quality of sequencing libraries was examined by the Qubit ssDNA

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

Assay Kit (Thermo Fisher Scientific, #Q10212). DNBs were loaded into the patterned nano arrays and sequenced on the ultrahigh-throughput DIPSEQ T-series sequencer.

Stereo-seq

To acquire spatial transcriptomic data, we collected E13.5 mouse brains, embedded them in Tissue-Tek O.C.T. compound and performed serial sectioning at a coronal thickness of 10 μm using a cryostat microtome. The tissue sections were mounted on Stereo-seq chip (1×1 cm), dried at 37°C for 3 min and fixed in precooled methanol for 30 min at -20°C. Before tissue permeation, we stained the sections with a nucleic acid reagent (Invitrogen, Q10212) for 5 min, washed them with RNase-free 0.1×SSC buffer and captured the images using a Motic PA53 FS6 microscope.

Subsequently, we permeated the tissue sections by incubating them with 0.1% pepsin in HCl buffer (0.01 M) for 10 min at 37°C to expose mRNA, performed reverse transcription for 3 hrs at 42°C and then digested with tissue removal buffer (10 mM Tris-HCl, 25 mM EDTA, 100 mM NaCl, 0.5% SDS) at 55°C for 10 min. Furthermore, the chip was treated with Exonuclease I (NEB, M0293L) for 16 hrs at 37°C to release barcoded cDNAs and washed once with 0.1×SSC buffer. The cDNAs were amplified using Hot Start DNA Polymerase (QIAGEN, 203603), purified using 0.6× VAHTSTM DNA Clean Beads and quantified using the Qubit dsDNA HS assay kit (Invitrogen, Q32854). We next tagmented the cDNAs with Tn5 transposases at 55°C for 10 min, amplified the fragmented cDNAs and purified them with AMPure XP Beads. The purified products were used for generating DNA nano balls and sequenced on the DNBSEQ-T series sequencing platform (MGI Tech, Shenzhen, China) using 50 bp for read 1 and 100 bp for read 2.

Sequencing data preprocessing

For scRNA-seq data sequenced with 10× Genomics Chromium platform, we used the Cellranger's pipeline (v3.0.2) to demultiplex raw sequencing data, assign barcodes and quantify unique molecular identifiers (UMIs). Using a prebuilt annotation package, we mapped the reads to human (GRCh38/hg38) and macaque (Macaca_fascicularis_6.0) reference genomes to generate gene expression matrices. For human snRNA-seq data produced by DNBSEQ platform, we filtered and demultiplexed the raw sequencing reads using PISA (https://github.com/shiquan/PISA) pipeline. Reads were further aligned to GRCh38/hg38 human genome using STAR (v2.7.4a) and sorted by Sambamba (v0.7.0).

For Stereo-seq data, the coordinate identity (CID) sequences on the forward reads were first mapped to the designed coordinates of Stereo-seq chip. Reads with molecular identifiers (MID) containing N bases or more than 2 bases with a quality score below 10 were filtered out. CID and MID associated with each read were included in the read header and the remaining reads were then aligned to GRCm38/mm10 mouse genome using STAR (v2.7.4a). These procedures allowed the generation of a CID-containing expression profile matrix.

Quality control and data integration

We first executed data quality control with the following criteria: 1) for the human and mouse datasets, we retained cells with less than 5∼10% mitochondrial gene counts, more than 600 genes and fewer than 6000∼8,000 genes; 2) macaque cells with less than 15% mitochondrial gene counts, more than 600 gene and fewer than 5000 genes were preserved for data integration; 3) exclusion of doublets for all datasets. Next, we aligned the orthologous transcripts across species with BiomaRt (v0.9.1) package,[64] employed canonical-correlation analysis to correct batch effects and integrated data from different species and across distinct developmental stages in a common embedding using a standard Seurat integration workflow.[65] We used the "FindIntegrationAnchors" function to identify "anchors" and the "IntegrateData" function to merge filtered datasets. To ascertain the robustness and consistency of our integrated results, we employed multiple integration methods, including RPCA, fastMNN, and Harmony algorithms, via the SeuratWrappers (v0.3.1) package. After data integration, we removed cells from adjacent brain regions such as thalamus and midbrain by detecting the expression of region-specific genes (e.g. *TCF7L2*, *OLIG3*, *GBX2*, *GATA3*, *FOXA2* and *SLC6A3*). Using the integrated single-cell data matrix, we performed principal components analysis and computed the first 50 principal components for dimensionality reduction with uniform manifold approximation and projection (UMAP).

Dimensionality reduction and clustering

Using the integrated single-cell data matrix that encompasses 416,536 qualified cells, we identified 2,000 highly variable feature genes, conducted principal component analysis (PCA) and selected the top 50 principal components for dimensionality reduction with uniform manifold approximation and projection (UMAP). A shared nearest neighbor (SNN) graph-based algorithm was applied for unsupervised clustering analysis. These clusters were compared pairwise to pinpoint cell type-specific genes and the cell identities were ascertained by cross-referencing their marker genes. Next, we further excluded neural progenitors and neurons from the developing thalamus and midbrain using specific marker gene expressions. For example, *OLIG3* exhibits specific expression in thalamic progenitor cells, while *TCF7L2*, *GBX2* and *GATA3* are notably enriched in the developing thalamus. After the cell filtering process, a final dataset comprising 351,742 cells was retained to comprehensively map the developmental landscape of human and mouse hypothalamus. The subclustering process focused on RGCs, IPCs, all postmitotic neurons, neuroendocrine neurons and dopamine neurons, which were subsetted and reclustered at an optimal resolution.

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

Differential gene expression analysis

To identify differentially expressed genes (DEGs) that are enriched in distinct cell types or neuronal subtypes, we performed pairwise comparisons and used "FindAllMarkers" function with Wilcoxon rank-sum test in Seurat, which returned logarithmic fold-changes (FC) of the average gene expression among groups and adjusted *P* values for each gene. Our criteria for designating cell type-specific DEGs included $\log_2\text{FC} > 0.5$, adjusted P value < 0.05 and expression ratio > 10%. For normalization, we utilized z-score scaling in Seurat to standardize the data to a Gaussian distribution. Within the identified DEG list, we further curated transcription factors (TFs) and primate-specific genes (PSGs), defining them as cell type-specific TFs or PSGs.

The human, macaque and mouse datasets were split to ascertain conserved and species-specific signature genes for each cell class. We randomly sampled 1,000 cells from each cell class to reduce the effect of varying cell numbers between species, applied log-normalization for each dataset and employed "FindAllMarkers" function to identify cell class-specific signature genes for each species. To differentiate conserved from divergent marker genes, we used the formula $CG_c = (H_c \cap M_c \cap N_c)$ for designating conserved marker genes in cell class c and $DG(S_{1d}) = \complement_{(G(S_{2d}) \cup G(S_{3d}))} G(s_{1d})$ for divergent marker genes specific to cell class d in species $S1$. Here, H, M and N represent human, macaque and mouse, respectively, while $G(s_{1d})$, $G(s_{2d})$ and $G(s_{3d})$ represent species-specific genes for d cell class.

Gene ontology (GO) analysis

The most significantly differentially expressed gene sets were selected for GO enrichment analysis. In this study, we employed either clusterProfiler software (v4.0)[66] or Metascape,[67] a web-based gene set analysis toolkit, to annotate transcriptional programs specifically enriched in various cell types and developmental stages. Functional annotation of cell type-specific gene sets was achieved by applying the hypergeometric test, with a significance threshold set at $P < 0.05$. Human- and mouse-specific gene modules were subjected to GO enrichment analysis, followed by functional protein association network analysis using STRING (https://string-db.org/) that is a biological database and web resource specializing in predicted protein-protein interactions.

Pearson correlation analysis

To assess the conservation of cell types across different species, we first identified the top 2000 highly variable genes that are one-to-one orthologs across different species, generated pseudo-bulk matrices "gene set × cell type" for each species and used a gene-specificity index to calculate cross-species pairwise correlation between cell clusters.[76] Within a set of cell types (C), the specificity index ($S_{g,c}$) of a gene (g) for a cell type (c) is defined as the ratio between the expression level of g within c (g_c) and the mean expression of g across C:

$$S_{g,c} = \frac{g_c}{\frac{1}{N}\sum_{i \in C} g_i}$$

where N is the number of cell types in C. We further calculated pairwise Spearman rank order correlations with the gene-specificity matrices determined as above, identifying correlated clusters across datasets.

To determine the RGC-IPC and IPC-neuron kinships, we used cosine similarity algorithm[68] to designate cell subtype-specific genes, computed the mean expression of these signature genes for each cell subtype and thereby yielded pairwise matrices (gene set × cell subtype) with a common gene set. We transformed these matrices into gene-specificity matrices and then calculated pairwise cluster correlations. Significance for correlation coefficients was determined with a permutation test.

Cross-species cell type correlation

To assess cross-species cell similarity, we employed MetaNeighbor (v1.9.1), a computation tool that quantifies the degree to which cell types replicate across datasets, to identify clusters with high transcriptomic similarity.[69] MetaNeighbor provides a performance score, which is the mean area under the receiver operator characteristic curve (AUROC), to quantify the similarity between cell-type pairs. We identified 2,000 variable features genes (1:1:1 orthologs) from the integrated dataset and computed AUROC scores using the "MetaNeighbourUS" function with these genes. For a given gene set, this approach generates a cell-cell similarity network by calculating Spearman correlation between all cells across the genes within the set. The network is then ranked and standardized, ensuring that all values fall within the range of 0 to 1. Each pairwise comparison between cell clusters is assigned an AUROC score, which varies from 1 (signifying a high similarity) to 0 (indicating a low similarity).

Developmental trajectory analysis

We performed pseudotemporal analyses to decipher the developmental trajectory of hypothalamic cell lineages and sublineages using Monocle 3 (v1.3.0),[77] Slingshot (v2.12.0),[33] scVelo(v0.2.4)[70] and URD (v1.1.1)[71] softwares. To confirm the cascade diversifying model in the developing human hypothalamus, we subsetted human RGCs, IPCs and neurons to generate a specialized dataset, performed gene expression normalization by log-transformation and then employed Monocle 3 to assign UMAP coordinates from the Seurat object to the 'cell_data_set' (CDS) object for trajectory inference. For Slingshot analysis, we first used "as.SingleCellExperiment" function to transform Seurat objects into SingleCellExperiment objects. We then inferred cell lineages with default parameters and visualized the trajectories by projecting them onto a UMAP plot. For scVelo analysis, we constructed spliced and unspliced count matrices, loaded the input data into Python, pre-processed it with the "scv.pp.moments" function and then

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

computed RNA velocity using the stochastic model. Genes with fewer than 20 counts in both spliced and unspliced matrices were excluded from the analysis. The RNA velocity streams across all cells were visualized by projecting a vector field on the UMAP embedding.

To estimate pseudotime for T-zone and M-zone lineages, we employed the "learn_graph" function in Monocle 3 to construct a principal graph, designated neural progenitors as the root node and subsequently applied the "order_cells" function with default parameters to compute pseudotime values and predict lineage progression path. Alternatively, we took advantage of URD,[71,78] a tool designed for reconstructing transcriptional trajectories, to infer differentiation paths. In brief, we computed transition probabilities based on transcriptomic distance between cells, generated a diffusion map using "calcDM" function in URD and specified 'root' and 'tip' cells to infer trajectory. Using the "floodPseudotime" function, we determined the distance of each cell from the root and assigned a pseudotime value to each cell within the lineage.

Gene module scoring

Scoring gene module represents a simple yet powerful approach for assessing the intensity of biological signals within a transcriptome, typically linked to a specific cell type, cellular component or biological process. As single-cell atlases provide specialized gene sets to define the transcriptional features of distinct cell types, public databases (e.g. MSigDB) have likewise supplied an array of well-recognized gene modules associated with diverse cellular components and biological processes. To evaluate the expression robustness of a specialized gene set in a cell group, we calculated the gene module scores with the following formula:

$$\mathrm{MS}_j(i) \ = \ \mathrm{mean}[\mathrm{Er}(\mathrm{G}_j, i)] - \mathrm{mean}\left[\mathrm{Er}\left(\mathrm{G}_j^{Ctrl}, i\right)\right].$$

where MS is the module score for individual cells, j is the input gene, i is the individual cell, Er is the relative expression of genes, G_j is the module gene set and G_j^{Ctrl} is the control gene set. Specifically, the $\mathrm{MS}_j(i)$ for i cell was determined by subtracting the average relative expression of genes in G_j from the average relative expression of a control gene-set (G_j^{Ctrl}). To define the control gene set, we binned all genes into 25 groups based on their aggregate expression levels and randomly selected 100 genes from the same expression bin for each gene in G_j. The gene modules used in this study were curated from either public databases (e.g. synaptic components or oxidative phosphorylation process) or defined using signature genes (e.g. RGC subtype markers) specific to each cell type. Using the "AddModuleScore" function in Seurat, we quantified the gene module scores for various cell groups divided by developmental timing or cell types.

Stage correspondence analysis

We established the correspondence between human and mouse developmental stages based on the transcriptome of hypothalamic cells using the recently published *TranscriptomeAge* method.[4] First, we performed a linear transformation on the human (PCW5-20) and mouse (E10-P0) datasets across different developmental stages using Min-Max normalization approach, given public databases suggesting that mid-fetal stage at PCW20 in human best matched mouse perinatal (P0) stage.[79] Second, the transformed ages of human and mouse hypothalamic cells were regressed against the average expression of orthologous genes to identify feature genes using a ridge regression model. We applied this regression model with the glmnet R package (v4.1.6), with $\alpha = 0$ and optimal λ value identified by ten-fold cross-validation using the "cv.glmnet" function. Genes with coefficients of 0 in the human and mouse cells were excluded and ranked based on the absolute value of the coefficient. Third, we conducted lasso regression analysis between the human ages (x) and the average expression matrix of feature genes (y), which consist of the overlapped top-ranking 12,000 genes identified by the ridge model in human and mouse datasets. The value of λ was similarly selected through ten-fold cross-validation. Lastly, we predicted the transcriptome age for both human and mouse hypothalamic cells and established their correspondence across multiple developmental stages. We further subdivided human ages into three stages, as described in Kang et al.[80]: stage 1 (embryonic stage), stage 2 (early-fetal stage) and stage 3 (mid-fetal stage).

Inference of lineage tree and lineage factors

To reconstruct the neurogenic lineage tree in the developing hypothalamus, we developed a computational tool integrating cosine similarity algorithm, gene specificity indexing and Spearman correlation analysis. During hypothalamus neurogenesis, RGCs adopt a conserved strategy for multipotential differentiation and hierarchically give rise to IPCs and nascent neurons. Along the lineage hierarchy, we isolated RGCs, IPCs and neurons from the integrated dataset and performed reclustering analysis for each major cell class. This process resulted in the subdivision of RGCs into 12 subtypes, IPCs into 14 subtypes and neurons into 22 subtypes, each characterized by distinct molecular signatures. To more accurately profile the transcriptional feature of each cell subtype, we applied a cosine similarity-based approach to identify gene signatures for each subtype of RGCs, IPCs and neurons. Cosine similarity compares the orientation of two n-dimensional vectors using the cosine value of the angle between the vectors in the vector space. Within the cell space, each dimension corresponds to a cell, a gene's vector comprises n-basis and the basis coordinates reflect the gene expression level in each cell. The cosine similarity between two genes is equivalent to the cosine value of the angle between their representative vectors in the cell space. To identify signature genes for a dataset of *N* cells (clustered into *K* groups) with *M* expressed genes, this algorithm creates an artificial marker gene λ_k for group k ($G_k, k \in \{1,\ldots,K\}$). The representative vector for each gene ($g_i, i \in \{1,\ldots,M\}$) will be compared with the representative vector of λ_k:

$$\cos(g_i,\lambda_k) = cosine(\theta) = \frac{g_i \cdot \lambda_k}{\|g_i\| \times \|\lambda_k\|} = \frac{\sum_{j=1}^{N} x_{ji}\lambda_{jk}}{\sqrt{\sum_{j=1}^{N} x_{ji^2}} \times \sqrt{\sum_{j=1}^{N} \lambda_{jk^2}}}$$

where x_{ji} is g_i's expression value in the cell j ($c_j, j \in \{1,\ldots,N\}$) and λ_{jk} is λ_k's expression value in c_j, $\lambda_{jk} = 1$ if $c_j \in G_k$ and $\lambda_{jk} = 0$ if $c_j \notin G_k$. Given that greater similarity in gene expression patterns results in a smaller angle, we evaluated the scores for each gene in each cell group with the following formula:

$$cos\ score(g_i,G_K) = \cos(g_i,\lambda_k) * \frac{\cos(g_i,\lambda_k)^2}{\cos(g_i,\lambda_k)^2 + \mu \times \sum_{t \in \{1,\ldots,K\}, t \neq k} \cos(g_i,\lambda_t)^2}$$

where $\mu \geq 0$ is the penalty factor and by default $\mu = 1$. We defined the genes with cosine scores greater than 0.1 as signature features for each cell subtype. A gene-specificity index as aforementioned was further used to calculate pairwise correlation between cell clusters across lineage hierarchy. After generating the gene-specificity matrices for each major cell classes, we merged the matrices that would be compared pairwise and quantified the Spearman correlation coefficients (SCC) for each pair of cell subtypes from two different datasets, which largely reflect the RGC-IPC and IPC-neuron kinships. To infer lineage relationships, we set a threshold for correlation coefficients: SCC > 0.37 & P value < 0.05 to link RGC and IPC subtypes, and SCC > 0.52 & P value < 0.05 for linking IPC and neuronal subtypes. The hypothalamic lineage tree was visualized by ggraph (v2.1.0).

To identify lineage-specific TFs that potentially mark, specify and maintain cellular identities, we split the cells in the hypothalamic lineage tree into P-, A-, T- and M-zone lineages (vertical partition), with each lineage containing RGCs, IPCs and neurons. We employed cosine similarity algorithm to identify lineage-specific genes, screened TFs enriched in each lineage, and selected the top TFs with COSG scores greater than 0.1 as candidate lineage factors. Simultaneously, we used the same computational tool to identify gene signatures for each subtype of RGCs, IPCs and neurons along the lineage hierarchy (horizontal partition). A final list of lineage factors was obtained by intersecting lineage-specific TFs with specific marker genes for different subtypes of RGCs, IPCs and neurons within individual lineages.

Shannon entropy analysis

Single-cell entropy explains much of the global variability in gene expression and serves as a simple metric that can reflect the transcriptomic diversity of a cell type or the differentiation diversity of a cellular lineage.[81] To assess the inter-cell transcriptomic heterogeneity in homologous neuronal subtypes across species, we computed and compared single-cell entropy for each subtype of human and mouse postmitotic neurons. We first identified 2000 highly variable genes (HVGs) in each species and performed random sampling of cells from each neuronal subtype to have an equivalent cell number between humans and mice. After conducting UMAP dimensionality reduction with the down-sampled data, we subdivided the cells from each subtype into m bins (m<20) based on their one-dimensional UMAP values, calculated the average expression of HVGs in each bin, and further split the bins into n groups (n<15) as per gene expression levels. Next, we computed Shannon entropy for each HVG in each bin and subtracted the entropy obtained from randomly permutated data. To ensure robustness, we performed 30 rounds of random permutations and analyzed the average value of entropy difference (>0.5). The down-sampling process was repeated 100 times, and the distribution of entropy differences for each neuronal subtype was visualized by violin plots.

Given that progeny cells are apt to acquire the molecular identity via inheriting TF codes from their progenitor cells in a cellular lineage, we determined to compute Shannon entropy for each TF enriched in individual lineages. We divided the hypothalamic neuronal lineages into P-, A-, T- and M-zones and identified all TFs from the top 1,000 HVGs specifically enriched in each lineage. To assess the spectrum of lineage diversification, we analyzed Shannon diversity index for each lineage:

$$H(X_j) = -\sum_{i=1}^{n} P(x_{ij}) \log P(x_{ij})$$

Wherein the expression of gene j was divided into n bins based on its expression level, $P(x_{ij})$ denotes the probability of gene j with expression in bin i. We normalized the Shannon diversity index and visualized the results using a radar map.

Spatial transcriptomic analysis

Scanpy provides a solution for spatial transcriptomic data analysis and visualization. Using the Stereo-seq data from serial coronal sections of the E13.5 mouse hypothalamus, we examined the spatial expression pattern of region-specific marker genes (e.g. *Nkx2-1*, *Arx*, *Pax6*, *Six6*, *Tbx3*, and *Barhl1*) using the "sc.pl.spatial" function in Scanpy, enabling the generation of scatter plots in spatial coordinates. To visualize the spatial distribution of different cell subtypes, we took advantage of Tangram (v1.0.4)[72] to transfer the gene expression pattern of each cell type from sc/snRNA-seq data to spatial profiling data. Prior to spatial alignment, we identified the molecular signatures specific to each subtype of RGCs, IPCs and neurons by COSG software, with a COSG score threshold set at > 0.1. Next, we employed these molecular signatures as input training genes for Tangram, and applied the "map_cells_to_space" function with the mode set to "clusters" to obtain a cell-type localization prediction. This procedure simulates the spatial correlation between each training gene in the sc/snRNA-seq data and the Stereo-seq data, and rearranged the single-cell

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

expression profiles of each cell type to spatial dimensions. While Tangram generated normalized probabilities of cell types in the Stereo-seq data, we selected cells with probabilities exceeding the 98th percentile to ensure data quality.

Functional validation of inferred sublineage

To corroborate the inferred neuronal sublineages, we genetically labeled Tbx3$^+$ progenitor domain using *Tbx3-CreER^{T2}::Ai9* mice at E9.5∼E10.5 and traced the fate of their progeny cells at P14. Our recent study profiled the identities of traced neurons by scRNA-seq at P14,[8] which was used as query dataset for lineage tracing data mapping. We further subset cells within the T-zone lineage from our hypothalamic lineage tree and defined it as a reference dataset. After log-normalization and variable feature identification of both datasets, we identified anchors between datasets using the ''FindIntegrationAnchors'' function and employed ''MapQuery'' function to project query cells onto reference UMAP feature plot. To investigate the potential role of lineage factors (e.g. Tbx3) in specifying neuronal fate, we integrated and used single-cell datasets from the hypothalamus of control and Tbx3-deficient mice as a reference[8]. The cells from the inferred T-zone lineage were randomly down-sampled and reversely mapped onto the reference UMAPs using ''MapQuery'' function within the Seurat workflows.

Cell abundance and compositional analysis

Cross-species comparison

To identify human-specific or human-enriched neuronal subtypes, we conducted comparative analysis using Milo, a versatile and scalable statistical framework that employs differential abundance testing by assigning cells to partially overlapping neighborhoods on a k-nearest neighbor (KNN) graph.[73] We first projected the gene expression matrix for each cell onto the leading principal components after log-transformation, and computed Euclidean distances between cells in this reduced space to identify the most similar cells for each cell. Second, we defined cell neighborhoods by grouping sets of interconnected cells on the KNN graph using default parameters. To counteract the differences in sampling depth, we normalized neighborhoods with trimmed mean of M-values. Lastly, we quantified neighborhood counts and fit a negative binomial generalized linear model to the counts for each neighborhood. Significance for each neighborhood was assessed by quasi-likelihood F-test. In summary, we quantified the abundance of each neuronal type across species by constructing KNN graph, defining cell neighborhoods, counting cells in neighborhoods and differential abundance test in neighborhoods.

Cross-sample comparison

To reliably assess changes in cell composition due to Tbx3 deletion in the hypothalamus, we employed three bioinformatic approaches for statistical analysis of cell type abundance changes as previously described[8]: scCODA (v0.1.9), Cacoa (v0.4.0), and Speckle (v0.0.3). scCODA selects the cell type with the least change as a reference and infers shifts in the relative proportions of all other cell types across samples with an FDR < 0.05 threshold. Cacoa determines significance based on random resampling of cells and replicates across the dataset in each sample. Speckle evaluates differences in cell type proportions between different samples by multiple hypothesis testing with Bonferroni correction. Neuronal subtypes with significantly reduced proportion in the Tbx3-deficient hypothalamus compared to controls were identified as lost subtypes.

Non-negative matrix factorization (NMF)

To decipher the transcriptional programs directing lineage progression in the developing hypothalamus, we adapted a NMF procedure to analyze the single-cell data of human and mouse neuronal lineages. We first performed trajectory analysis of P-, A-, T-, M- and PTh-zone lineages using Monocle 3 and subdivided the cells in each putative lineage into 10 pseudotemporal bins. To dissect the gene expression dynamics during lineage progression, we applied the NMF framework, available at https://github.com/YiqunW/NMF, to each pseudotemporal bin. This process involved decomposing a gene expression matrix (with dimensions M genes × N cells) into two matrices: a matrix G (M genes × K modules), and a matrix C (K modules × N cells). Matrix G assigns a weight to each gene for each gene module, whereas matrix C associates each gene module with a level for each cell. We determined the optimal number of gene modules (K value) independently for each dataset as follows: we tested K values within the range of 5 to 13 and repeated the NMF process 5 times for each K value with random initial conditions. To ensure the data reliability, we evaluated the consistency among results from repeated NMF runs and considered the cophenetic coefficient. A K value with low inconsistency and a high cophenetic coefficient was selected for each pseudotemporal bin matrix. Next, we integrated gene modules from different lineages and pseudotemporal bins, computed the Spearman correlation between gene modules and performed hierarchical clustering analysis, which identified 10 meta-programs for both human and mouse NMF results. For each metagene-program, genes were ranked based on their NMF scores within gene modules. We selected the top 30 genes from each gene module and combined these genes together to define each meta-program. Furthermore, we conducted hypergeometric testing on the top 40 genes with the highest frequency in each meta-program (significance defined by FDR-adjusted $P < 0.05$) to indicate the distinct features of each meta-program, followed by GO functional annotation.

Identification of homologous cell types

To estimate cell-type homology, we independently clustered human and mouse hypothalamic neurons across various developmental stages using the Louvain algorithm. We clustered the aligned embedding output from Seurat and identified human and mouse neuronal subtypes that co-clustered.[82] Specifically, we first constructed a weighted graph based on the Jaccard similarity of the nearest neighbors for each neuronal subtype. Next, we employed Louvain community detection to identify clusters by optimizing

Please cite this article in press as: Chen et al., Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development, Developmental Cell (2025), https://doi.org/10.1016/j.devcel.2025.03.009

the global modularity of the partitioned graph. For each pair of human and mouse neuronal subtypes, the overlap was defined as the sum of the minimum proportion of cells in each original subtype that overlapped within each aligned cluster. This approach consistently identified pairs of human and mouse clusters that co-clustered within one or more aligned clusters. The degree of cluster overlap, ranging from 0 to 1, was visualized as a heatmap. Finally, cell-type homologies were defined based on the overlap patterns observed between clusters.

Label transfer

To determine cell identities across different datasets, we used label transfer from annotated reference data. Specifically, each query and reference dataset were preprocessed using Seurat's SCTransform method, and the top 3,000 highly variable orthologous genes were selected. Label transfer was then performed using the first 30 principal components, with anchor points between the reference and query datasets identified by the "transferData" function. This allowed for the projection of PCA structure from the reference data. Finally, we generated a UMAP representation with the "RunUMAP" function and mapped both UMAP coordinates and annotations from the reference to the query data using the "MapQuery" function.

Identification of regulons

Cell type-specific regulatory network inference (CESNI)

Cell type-specific gene expression patterns results from the activity of transcriptional gene regulatory networks (GRNs), which establishes connections between master TFs, signaling proteins and target genes. To reveal potential GRNs for each neuronal subtype, we first conducted differential gene expression analysis to identify cluster biomarkers with "FindAllMarkers" function in Seurat and selected the top 50 marker genes of each subtype for subsequent analysis. Among these biomarkers, we considered TFs as potential regulators and retrieved DNA sequences encompassing 2000 bp upstream and 500 bp downstream of the transcriptional start site for each selected marker gene. Using TF binding motifs inferred by the CHIP-seq database (https://github.com/GreenleafLab/chromVARmotifs), we employed the "matchMotifs" function within motifmatchr package (v1.22.0) to establish connections between the regulators and target genes based on accessible sequence-specific motif matches (https://github.com/GreenleafLab/motifmatchr). The motif matching score was calculated by:

$$Score(u) = \sum_{j=1}^{m} log \frac{\Pi(\mathrm{j}, u_j)}{q_{u_j}}$$

where m is the positionally weighted pattern, j is the position, u is the DNA sequence, u_j is sequence position, q is background probability, q_{u_j} is the background probability in specific sequence position. We set $Score(u) > 0.5$ as a threshold to establish the regulatory relationship between TF and target genes.

Single-Cell Regulatory Network Inference (SCENIC)

We further identified GRNs (i.e. regulons) for each neuronal subtype with the SCENIC (v0.12.0) pipeline.[74] In brief, we generated the gene co-expression modules that encompass TFs and potential target genes for each cell, inferred regulons by leveraging a motif database with genome-wide rankings, scored the activity of each regulon using the area under the recovery curve and determined the regulon's cell-type specificity with regulon specificity score (RSS). We selected the top 6 regulons for each neuronal subtype to create a regulon network and quantified their species specificity.

GWAS enrichment analysis

To evaluate the enrichment of genes associated with human diseases and traits in each dopamine neuronal subtype, we performed linkage disequilibrium score (LDSC) regression analysis as described previously (https://github.com/bulik/ldsc/wiki/LD-Score-Estimation-Tutorial). We obtained GWAS (genome-wide association studies) summary statistics for quantitative traits related to neural and metabolic disorders from previous literatures,[83] including epilepsy, schizophrenia, neuroticsm, major depressive disorder, bipolar disorder, psychiatric disorders, schizophrenia, autism, Parkinson's disease, body mass index (BMI), height, waist-to-hip ratio (WHR), high cholesterol, basal metabolic rate, fasting glucose, high density lipoprotein (HDL), low density lipoprotein (LDL), triglyceride, diabetes, intelligence and smoking habits. We pre-processed summary statistics to the standard format using the LDSC pipeline,[84] followed by 1) identification of DEGs with adjusted P < 0.05 and log2FC > 0.5 in each neuronal subtype; 2) creation of annotation files for 22 chromosomes in each cell subtype with the "make_annot.py" function and using the options "–bed-file, –bimfile 1000G.EUR.QC.bim" and "–annot-file"; 3) calculation of LD scores with the "ldsc.py" function and using the options "–l2, –bfile 1000G.EUR.QC, –ld-wind-cm 1, –annot, –thin-annot, and –print-snps"; and 4) performance of regression using the "ldsc.py" script with the "–h2-cts flag". The association of each neuronal subtype with the traits was determined by reporting the coefficient *P*-values.

QUANTIFICATION AND STATISTICAL ANALYSIS

General statistical analysis

All statistical details of the experiments can be found in the figure legends and/or text. Statistical analyses were performed using GraphPad Prism, Microsoft Excel, and R software. Data collection and analysis were blinded in this study. Grouped data are presented as mean ± standard error of the mean (SEM). Statistical significance was assessed using unpaired two-tailed Student's *t*-test or permutation test. Statistical significance is indicated as follows: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. All experiments were independently repeated at least three times.

Quantification of neuronal number and neurite length

To quantify the number of GnRH neurons, human and mouse brain tissues were serially sectioned into 40-μm slices and immunostained. The total number of GnRH neurons on each slice was counted by LAS X (v3.0.11). Each group included two human fetal brains and at least three mouse brains. Axonal and dendritic length of cultured neurons expressing EGFP was quantified using ImageJ (v1.53c). To analyze the morphology of neurons transfected with ONECUT1 *in vivo*, mouse brains were sectioned into 150-μm thick slices with vibratome and imaged with a Nikon A1R MP microscope (Japan). Neurons were reconstructed using the "Surface" and "Filaments" functions in Imaris (v9.2), which automatically analyzed neurite length. Each group included at least three mouse brains.

Developmental Cell, Volume *60*

Supplemental information

Transcriptional conservation and evolutionary divergence of cell types across mammalian hypothalamus development

Zhen-Hua Chen, Taotao Bruce Pan, Yu-Hong Zhang, Ben Wang, Xue-Lian Sun, Meixi Gao, Yang Sun, Mingrui Xu, Shuhui Han, Xiang Shi, Felipe Correa-da-Silva, Chenlu Yang, Junfu Guo, Haoda Wu, Yu Zheng Li, Xiu-Qin Liu, Fei Gao, Zhiheng Xu, Shengjin Xu, Xin Liu, Ying Zhu, Ziqing Deng, Shiping Liu, Yi Zhou, Chun-Xia Yi, Longqi Liu, and Qing-Feng Wu

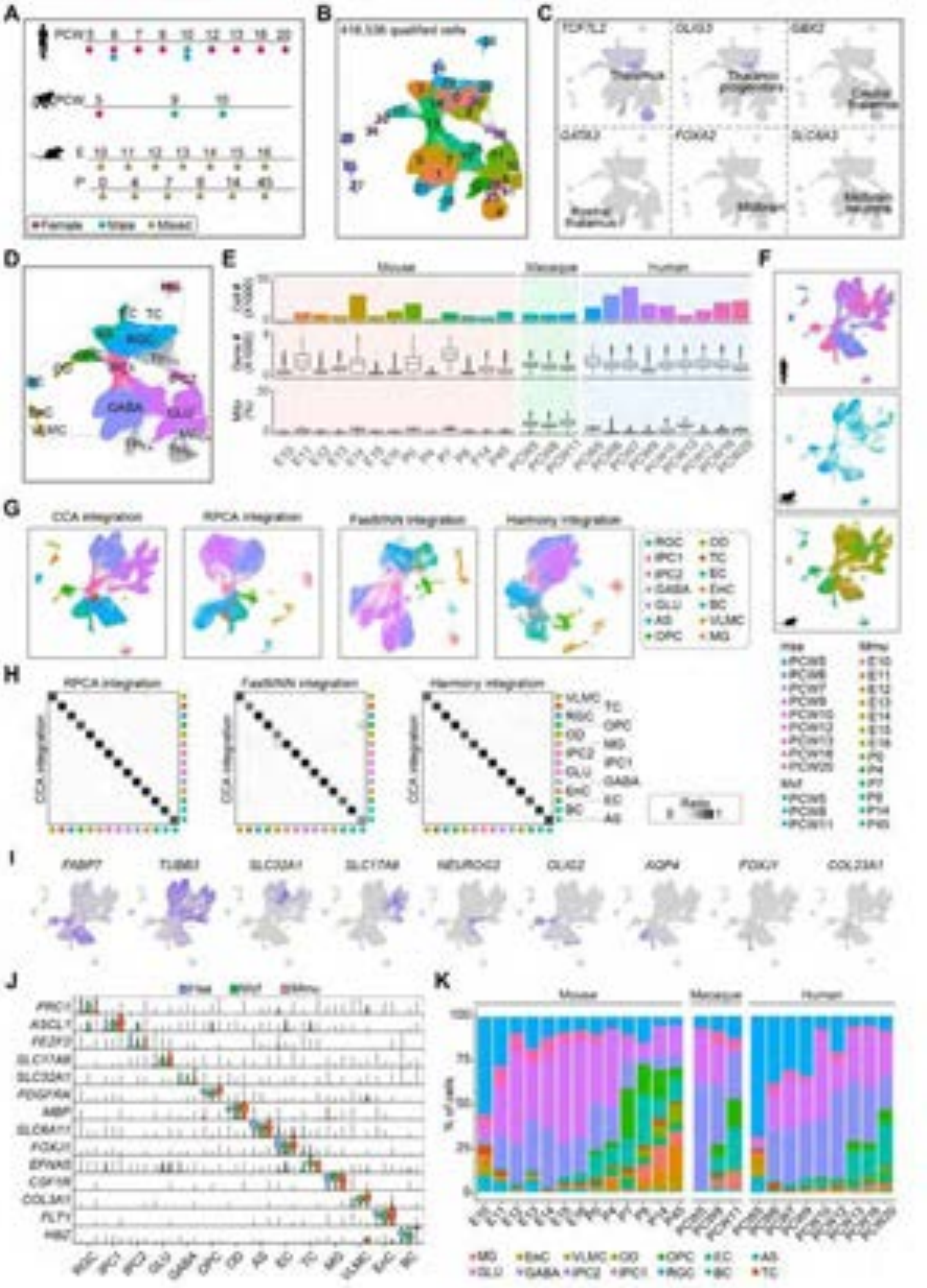


Figure S1: An overview of sc/snRNA-seq data from the hypothalamus of humans, macaques and mice, related to Figure 1.

(**A**) Diagram illustrating the developmental stages and genders of samples collected across different species. PCW, week post conception, E, embryonic; P, postnatal.

(**B**) UMAP visualization of a total of 416,536 qualified cells collected from the human, macaque and mouse developing hypothalamus.

(**C**) Feature plots illustrating the expression pattern of genes specific to thalamus and midbrain. Thalamic progenitors are characterized by *OLIG3* expression, thalamic neurons specifically express *TCF7L2*, *GBX2* and *GATA3*, while midbrain cells display enriched expression of *FOXA2* and *SLC6A3* expression.

(**D**) Cell annotation in UMAP with hypothalamic cells highlighted within dashed circles. Contaminated cells from adjacent brain regions (e.g. thalamus and midbrain) are teased out for subsequent analyses. RGC, radial glial cells; IPC, intermediate progenitor cells; AS, astrocytes; OPC, oligodendrocyte precursor cells; OD, oligodendrocytes; EC, ependymal cells; GABA, GABAergic neurons; GLU, glutamatergic neurons; TC, tanycytes; BC, blood cells; EnC, endothelial cells; MG, microglia; VLMC, vascular and leptomeningeal cells; $\mathrm{TH_{Ex}}$, thalamic cells for exclusion; $\mathrm{MB_{Ex}}$, midbrain cells for exclusion.

(**E**) Bar and box charts showing the cell count, gene number per cell, UMI number per cell and mitochondrial gene count per cell for each species across various developmental time points.

(**F**) UMAP pots showing the distribution of hypothalamic cells from humans (Hsa, 195,166 cells), macaques (Mcf, 25,847 cells) and mice (Mmu, 130,729 cells), color-coded by different developmental time points.

(**G**) UMAP visualization of integrated datasets using CCA, RPCA, FastMMN and Hormony algorithms.

(**H**) Heatmaps showing the consistency of cell clustering results across different integration methods.

(**I**) Feature plots illustrating the expression pattern of *FABP7*, *TUBB3*, *SLC32A1*, *SLC17A6*, *NEUROG2*, *OLIG2*, *AQP4*, *FOXJ1* and *COL23A1*. for each cluster.

(**J**) Violin plots showing the expression of conserved cell markers within each cell class across different species.

(**K**) Histogram showing the relative proportion of cells in each cell class at different developmental time points.

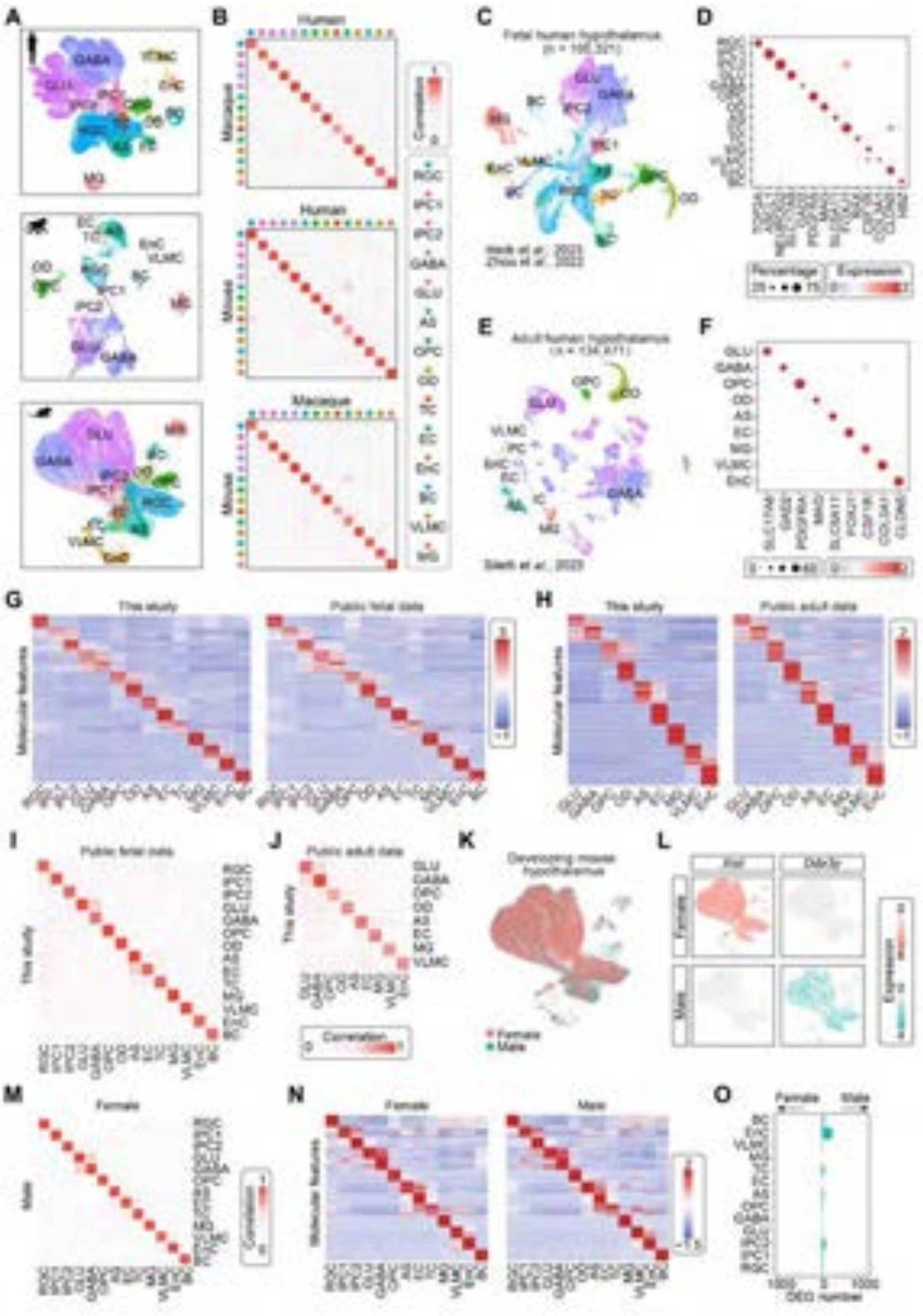


Figure S2: Conserved cell classes across different species and datasets, related to Figure 1.

(**A**) Independent clustering of human, macaque and mouse datasets.
(**B**) Correlation analysis of the major cell types across different species.

(**C**) UMAP visualization of the clustering analysis of public fetal human hypothalamus.
(**D**) Dot plot showing the conserved marker gene expression for each fetal hypothalamic cell class.
(**E**) UMAP visualization of the clustering analysis of public adult human hypothalamus.
(**F**) Dot plot illustrating conserved marker gene expression for each adult hypothalamic cell class.
(**G and H**) Heatmaps showing common molecular features across our dataset in comparison to public fetal (G) and adult (H) hypothalamic datasets.
(**I and J**) Correlation analysis of the major cell types between our dataset and public fetal (**I**) or adult (J) hypothalamic datasets.
(**K**) UMAP plots showing female and male hypothalamic cells from prenatal and perinatal mice.
(**L**) Feature plots illustrating the expression patterns of *Xist* and *Ddx3y* in the developing hypothalamus of both female and male mice.
(**M**) Correlation analysis of the major cell types between female and male mice.
(**N**) Heatmaps showing the common molecular features across female and male mice.
(**O**) Stacked barplot showing the number of differentially expressed genes across major hypothalamic cell types between female and male mice.

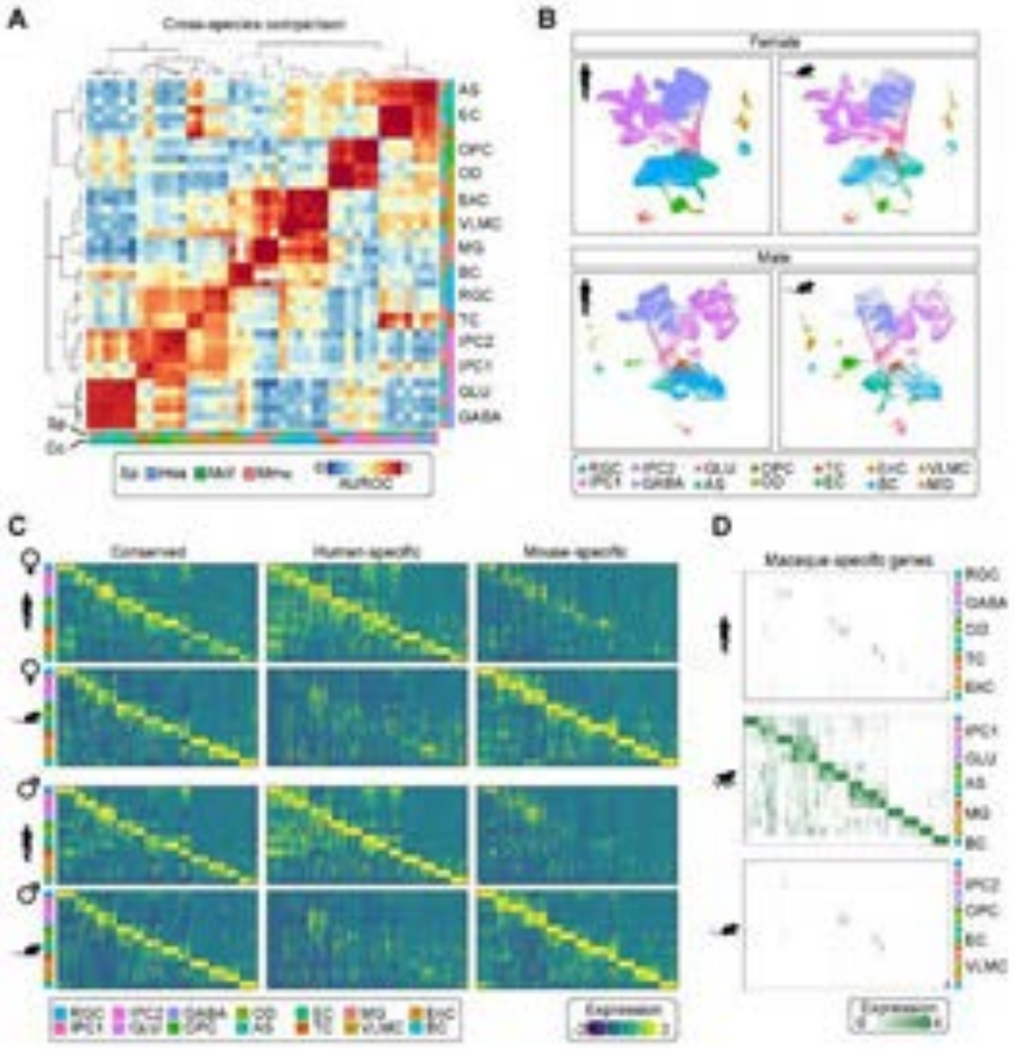


Figure S3: Conserved cell classes across different species and sexes, related to Figure 1.
(**A**) Heatmap showing the area under the receiver operator characteristic curve (AUROC) scores of cell signatures across species. We used Metaneighbor program to compare gene expression profiles. Sp, species; Cc, cell classes.
(**B**) UMAP plots showing female (up) and male (bottom) hypothalamic cells from both humans and mice.
(**C**) Heatmaps showing the conserved and species-specific molecular features in each cell type from female and male humans and mice.
(**D**) Heatmap showing macaque-specific molecular features in each cell type.

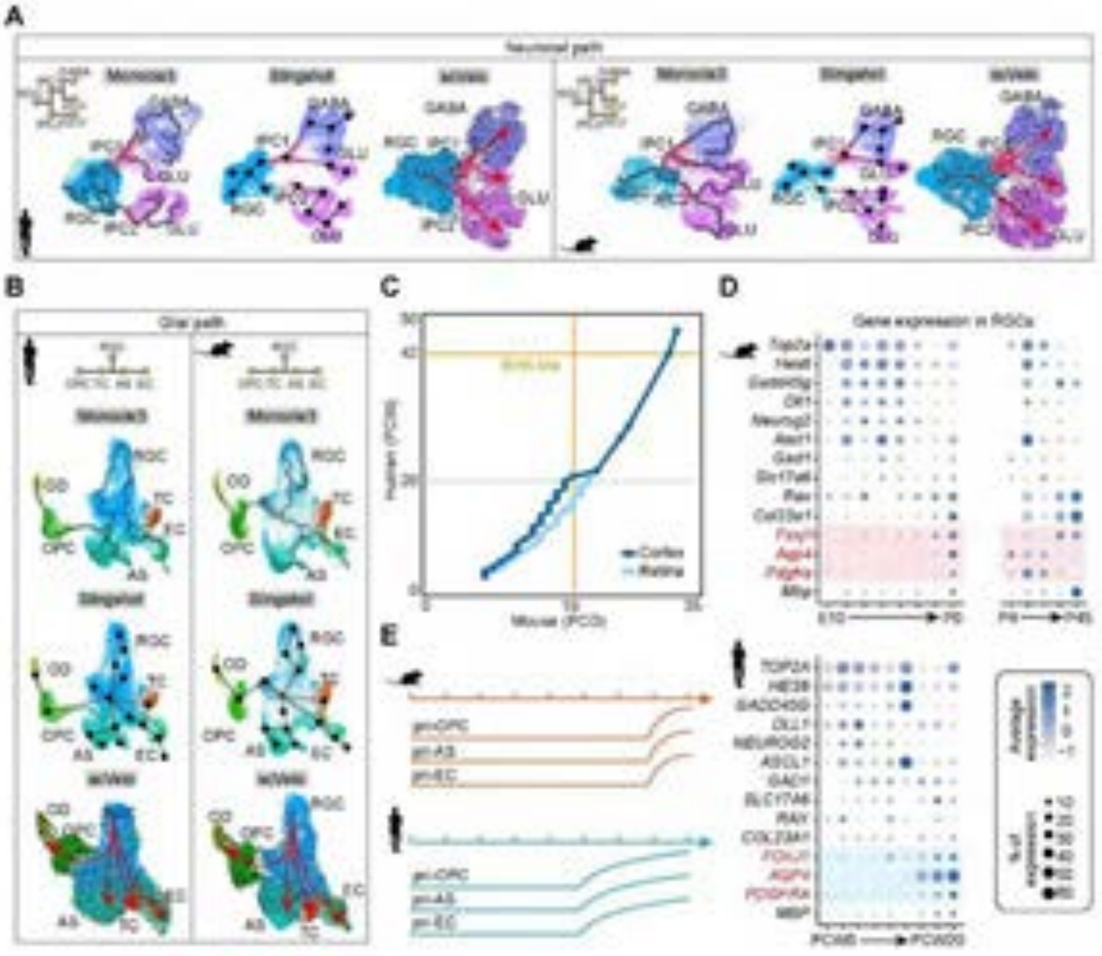


Figure S4: Conserved neurogenic lineage progression and asynchronous gliogenic pace in humans and mice, related to Figure 1.

(**A and B**) UMAP visualization of the developmental trajectories for neurogenic (A) and gliogenic lineages (B) in the human and mouse developing hypothalamus, inferred by Monocle 3, Slingshot, and scVelo. Notably, RGCs differentiate into both IPC1 and IPC2, with IPC1 further bifurcating into GABAergic and glutamatergic neurons.

(**C**) A line chart depicting the stage correspondences of cortical and retinal development between mice and humans[1]. This visualization effectively demonstrates the aligned developmental timeframes in both species. Mid-fetal stage at PCW20 in human best matched mouse perinatal (P0) stage.

(**D**) Dot plots showing the developmental expression dynamics of cell type-specific markers in mouse and human RGCs. Transcriptional priming, denoted by mRNA expression of lineage-specific marker genes such as *ASCL1* and *NEUROG2* for neuronal differentiation, *PDGFRA* for OPC differentiation, *AQP4* for astrocyte differentiation and *FOXJ1* for ependymocyte differentiation, occurs in RGCs before embarking on the path of differentiation. The asynchronous expression of *PDGFRA*, *AQP4* and *FOXJ1* between mice and humans are highlighted in colors, suggesting asynchronous timing of gliogenesis in different species.

(**E**) A schematic diagram illustrating the temporal occurrence of glial cells in the mouse and human hypothalamus, with human gliogenesis commencing relatively earlier.

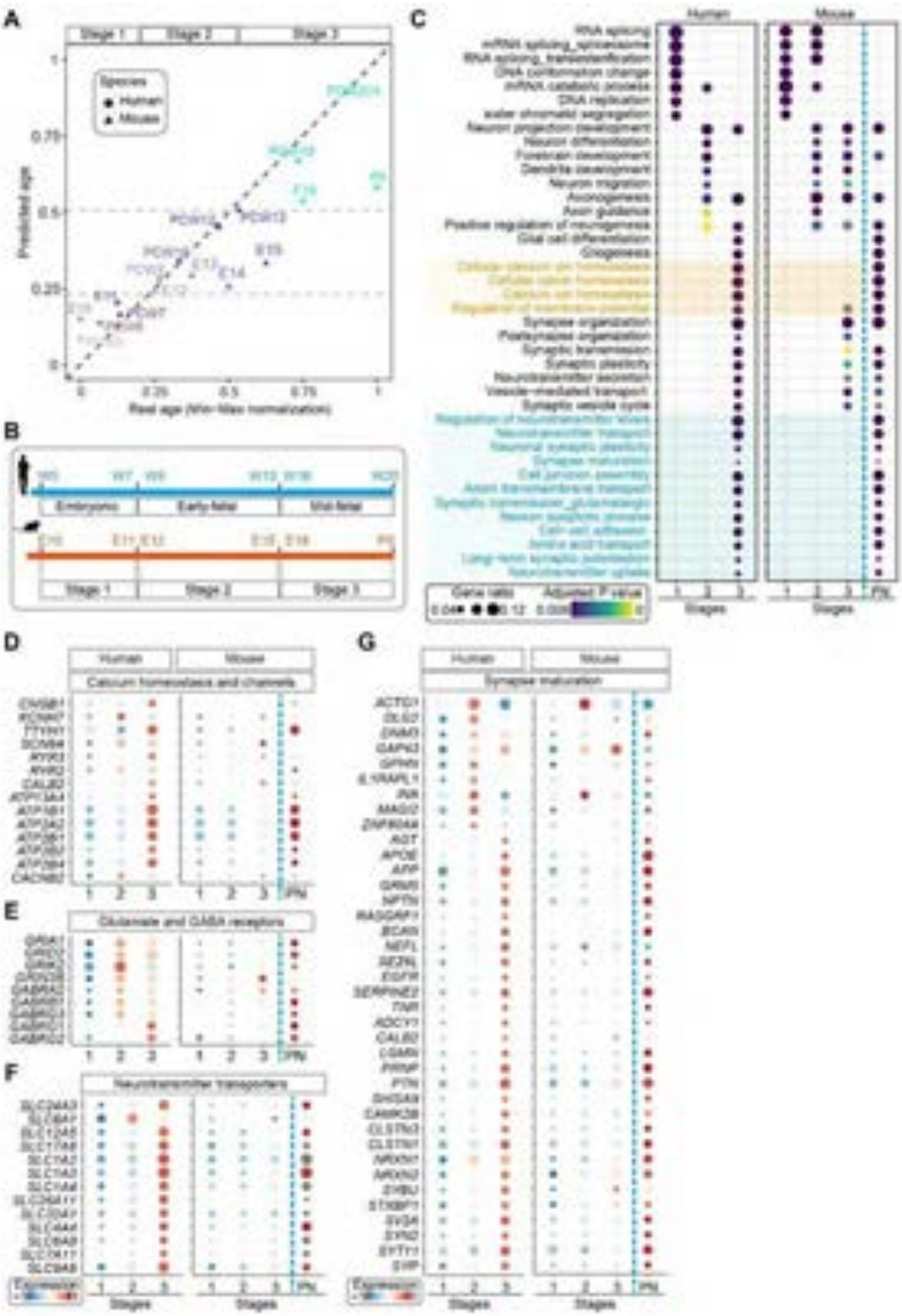


Figure S5: Asynchrony in neuronal maturation between humans and mice, related to Figure 1.

(**A**) Scatter plot depicting the real age (x axis) and the age predicted by *TranscriptomeAge* (y axis) of human and mouse datasets. Real ages were normalized by Min-Max normalization and subdivided into three stages: stage 1 (embryonic stage), stage 2 (early-fetal stage) and stage 3 (mid-fetal stage). We applied a lasso regression model to predict the transcriptome age of each sample.

(**B**) Schematic showing the corresponding developmental stages between human and mouse hypothalamus.

(**C**) Dot plots showing the developmental stage-specific biological processes in both humans and mice. The human-specific enrichment of calcium homeostasis and synaptic maturation is highlighted in colors, suggesting the asynchronous neuronal maturation in the human and mouse developing hypothalamus. PN, postnatal stage.

(**D-G**) Dot plots showing the expression dynamics of calcium channels (D), glutamate and GABA receptors (E), neurotransmitter transporters (F) and genes involved in synaptic maturation (G) in mouse and human hypothalamic neurons across different developmental stages.

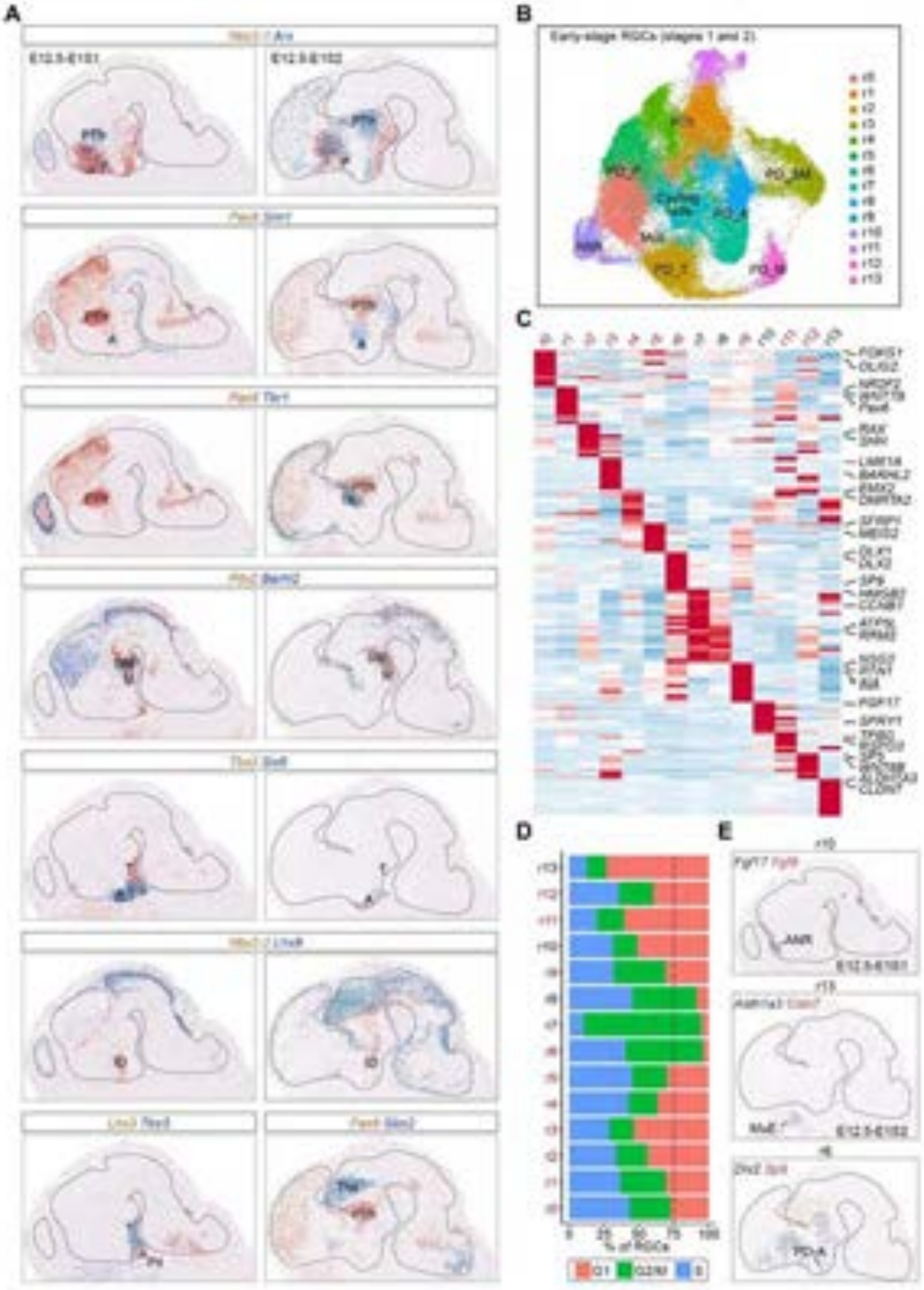


Figure S6: Identification, annotation and validation of hypothalamic progenitor domains using spatial transcriptomic data, related to Figure 2.

(**A**) Spatial visualization of *Nkx2-1*, *Arx*, *Pax6*, *Sim1*, *Tbr1*, *Barhl2*, *Pitx2*, *Tbx3*, *Six6*, *Nkx2-2*, *Lhx9*, *Lhx3* and *Gbx2* expression in E12.5 sagittal brain sections using the mouse organogenesis spatiotemporal transcriptomic atlas (MOSTA)[2]. PTh, prethalamus; P, preoptic zone; A, anterior zone; T, tuberal zone; M, mammillary zone; SM, supramammillary zone; TE, thalamic eminence; ID, extrahypothalamic diagonal; Tha, thalamus; Pit, pituitary.

(**B**) Subclustering and spatial annotation of early neural progenitor cells at stages 1 and 2. Cycling cells (nP7 and nP8) without spatial identities, anterior neural ridge (ANR) cells (nP10) and mucosal epithelial (MuE) cells (nP13) were removed for subsequent analysis.

(**C**) Heatmap showing the marker genes enriched in distinct neural progenitor domains.

(**D**) Stacked bar chart showing the proportion of cells across various cell cycle phases in each cluster of neural progenitors.

(**E**) Spatial visualization of marker gene expression for nP10 (marked by *Fgf17*), nP13 (denoted by *Aldh1a3*) and nP6 (featured by *Dlx2* and *Sp9*) in E12.5 sagittal mouse brain sections. Despite that neural progenitors in nP6 cluster are predominantly proliferative, they display a specific spatial distribution within the A zone.

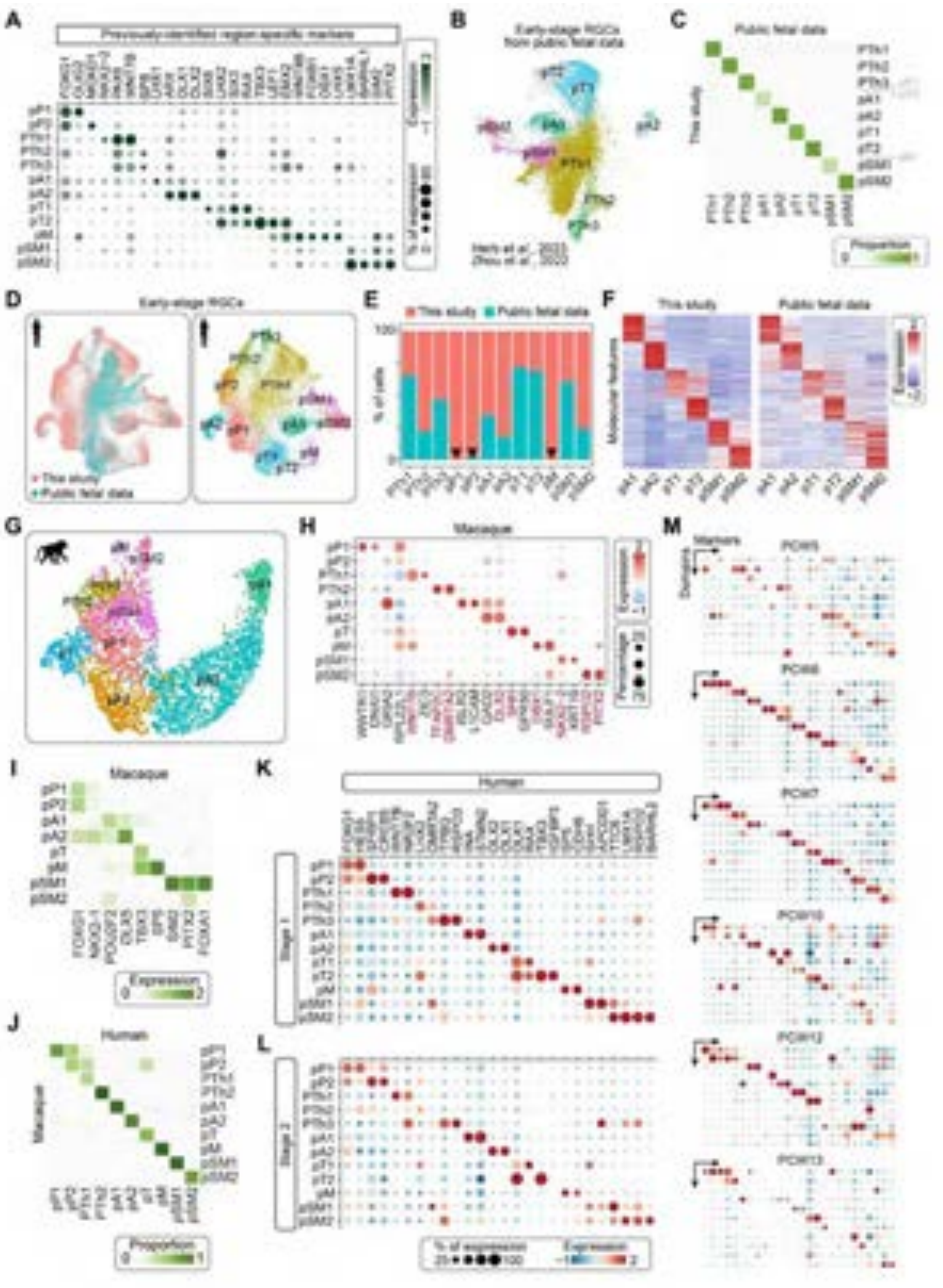


Figure S7: Molecular features of distinct hypothalamic neural progenitor domains, related to Figure 2.

(**A**) Dot plot showing the expression of putative region-specific marker genes (e.g. *PAX6*, *NKX2-1*, *TBX3*, *SIX6* and *BARHL2*) in the 9 hypothalamic and 3 prethalamic progenitor domains.

(**B**) UMAP visualization of neuronal progenitor domains identified from public fetal datasets.

(**C**) Correspondence between our data and public fetal data by label transfer. Three subtypes of neural progenitors identified in our study are absent in the public data.
(**D)** UMAP plots showing the integration and clustering of early RGCs from both our study and public fetal data.
(**E**) Stacked bar plots illustrating the ratio of cells across our dataset and public datasets in various neural progenitor subtypes, with arrowheads marking the RGC subtypes absent in the public data.
(**F**) Heatmaps showing the molecular features of distinct progenitor domains shared between our study and public fetal data.
(**G**) UMAP visualization of macaque progenitor domains.
(**H**) Dot plots showing marker gene expression across distinct macaque progenitor domains.
(**I**) Conserved TFs expressed in macaque progenitor domains.
(**J**) Inter-species correspondence of RGC subtypes between humans and macaques by label transfer.
(**K-M**) Dot plots showing robust and consistent patterns of signature gene expression in distinct neural progenitor domains at different developmental stages.

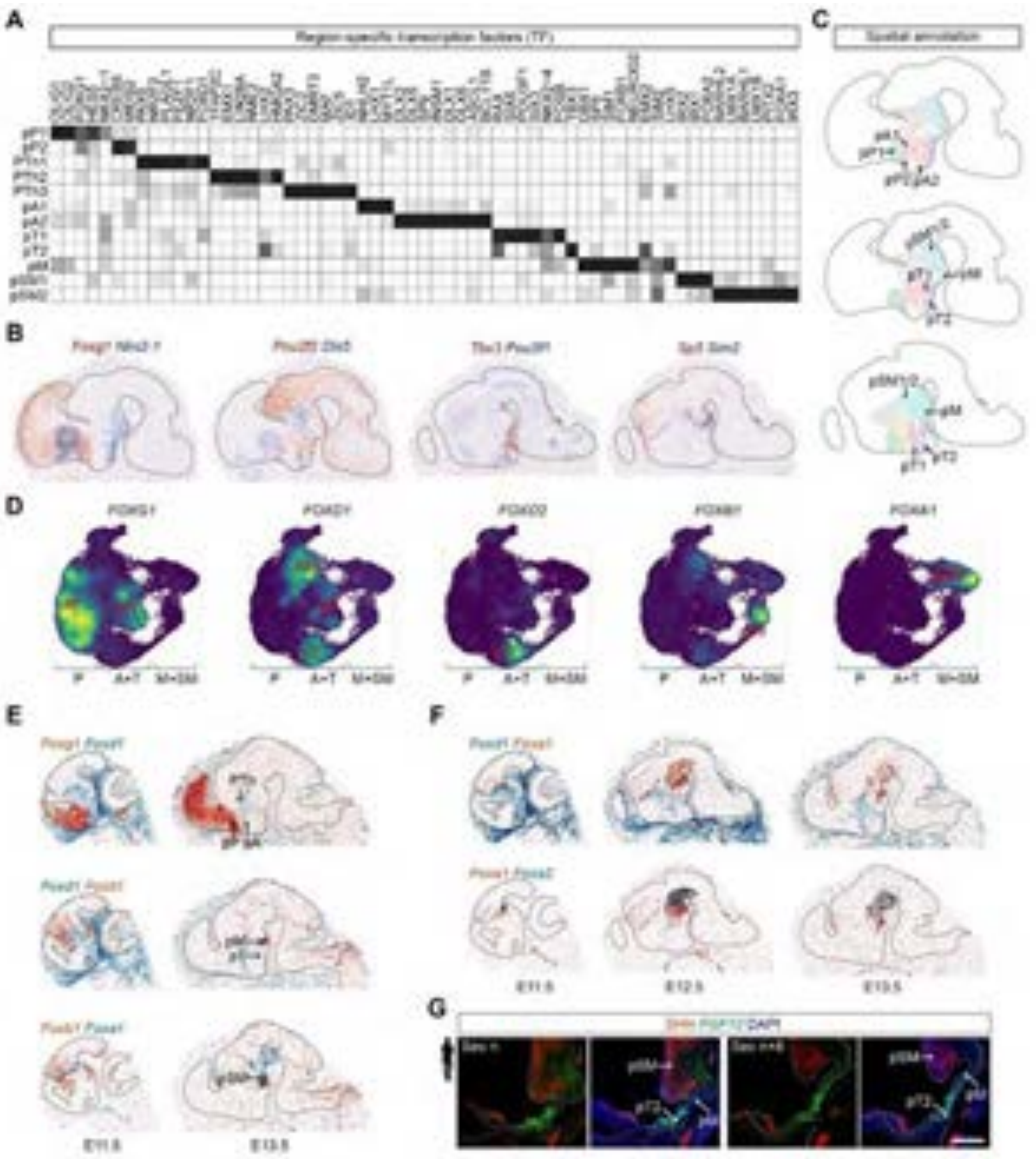


Figure S8: Segmentation of hypothalamic progenitor domains by FOX gene family, related to Figure 2.

(A) Heatmap showing the specific expression of distinct TF sets in various neural progenitor domains.

(**B**) Spatial visualization of conserved region-specific marker genes (*Foxg1*, *Nkx2-1*, *Pou2f2*, *Dlx5*, *Tbx3*, *Pou3f1*, *Sp5* and *Sim2*), along with schematic diagrams indicating the spatial annotation of each hypothalamic progenitor domain.

(**C**) Schematic diagram showing the spatial distribution of hypothalamic neural progenitor domains.

(**D**) Feature plots showing the expression of various members of the *FOX* gene family within specific progenitor domains.

(**E**) Spatial transcriptomic analysis showing the expression of *Foxg1*, *Foxd1*, *Foxb1* and *Foxa1* in E11.5 and E13.5 mouse brain sections.

(**F**) Spatial visualization of *Foxd1*, *Foxa1* and *Foxa2* expression patterns in E11.5, E12.5 and E13.5 mouse brain sections.

(**G**) Sample images showing the spatial expression of *SHH* and *FGF10* mRNA in a sagittal view of PCW6 human fetal brain sections. *FGF10* was restricted in pT2 domain and mutually exclusive with *SHH* in the hypothalamic floor plate. Scale bar, 500 μm.

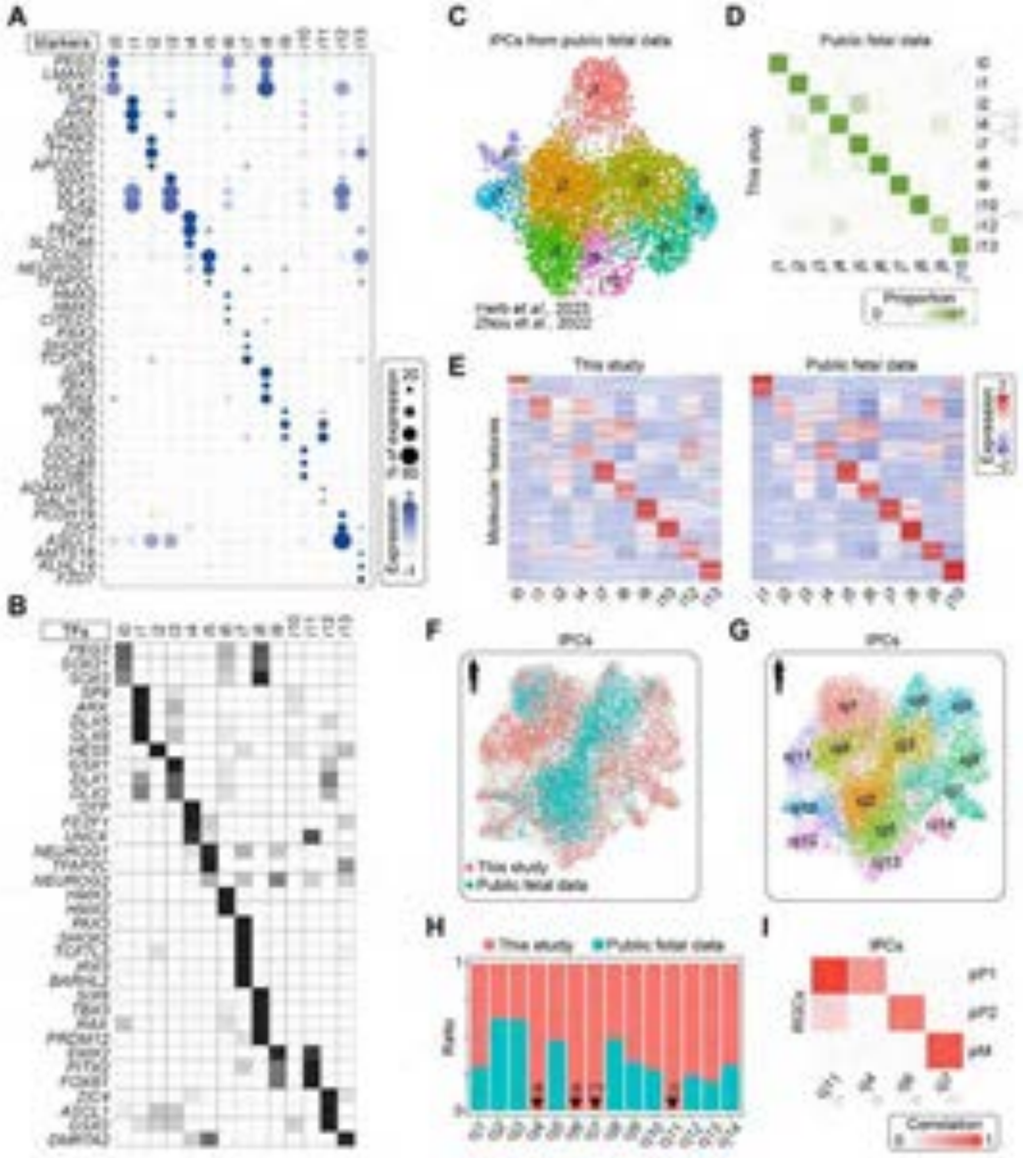


Figure S9: Identification of IPC subtypes, related to Figure 3.
(**A**) Dot plot showing the molecular signatures in each IPC subtype.
(**B**) Heatmap demonstrating the subtype-specific TFs among IPCs.
(**C**) UMAP visualization of IPC subtypes identified from public fetal data.
(**D**) Correspondence between our data and public fetal data by label transfer. Four IPC subtypes identified in our study are absent in the public data.
(**E**) Heatmaps showing the similar molecular features of distinct IPC subtypes in our study and public fetal data.
(**F and G**) UMAP plots showing the integration (F) and clustering (G) of IPCs from both our study and public fetal data.
(**H**) Stacked bar plots illustrating the ratio of cells across our dataset and public datasets in various IPC subtypes, with arrowheads indicating subtypes that were absent in the public data.
(**I**) Correlation analysis of RGC and IPC subtypes that were missing in the public data.

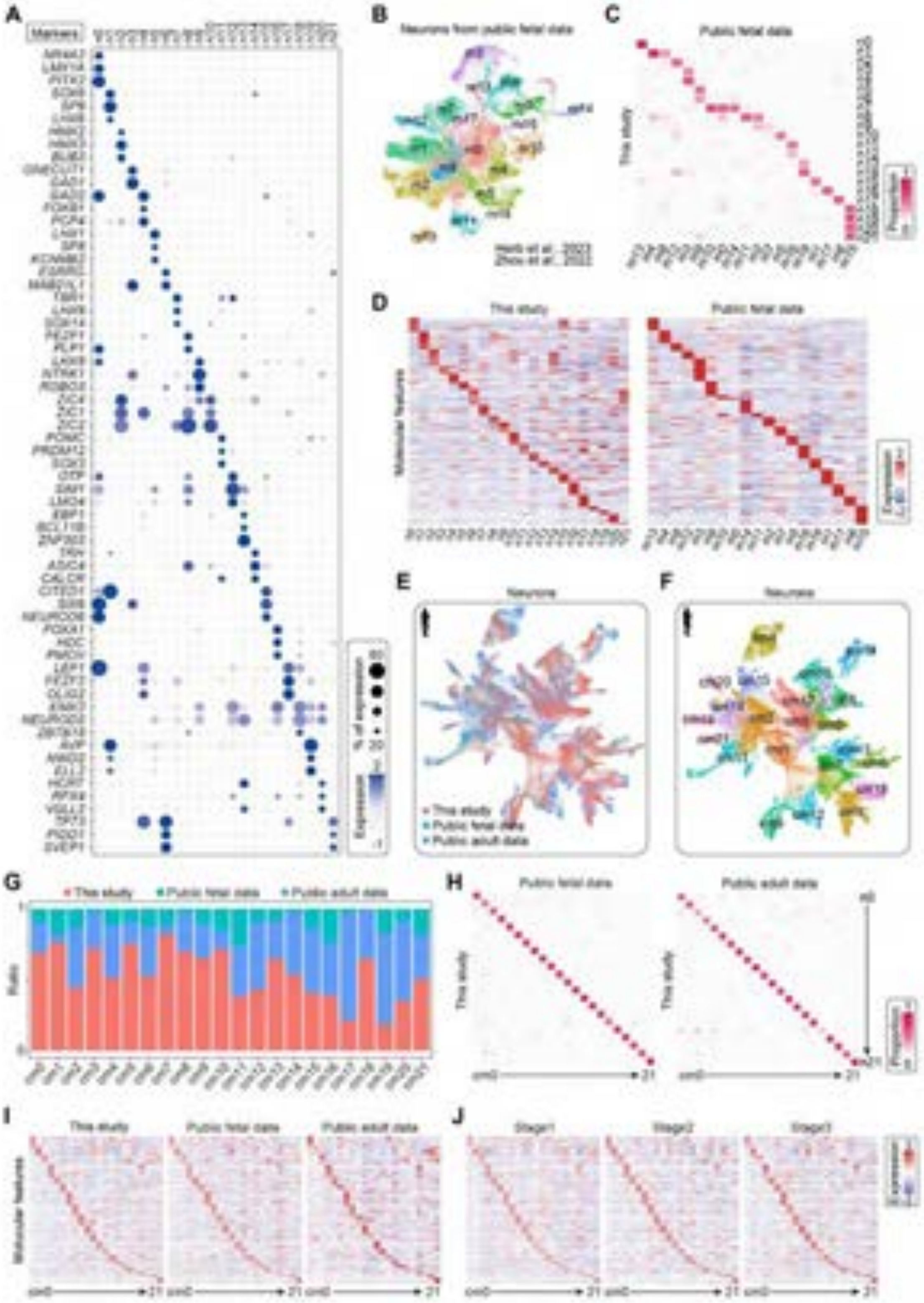


Figure S10: Identification of neuronal clusters for lineage inference, related to Figure 3.
(**A**) Dot plot illustrating the signature genes in each neuronal subtype.
(**B**) Independent clustering of neurons from public fetal data.

(**C**) Correspondence of neuronal subtypes between our study and public fetal data by label transfer.
(**D**) Heatmaps showing the similar molecular features of neuronal subtypes between our study and public fetal data.
(**E and F**) UMAP plots showing the integration (E) and clustering (F) of neurons from our study, public fetal and adult data.
(**G**) Stacked bar plots illustrating the ratio of cells across our dataset and public datasets in various neuronal subtypes.
(H) Correspondence of neuronal subtypes between our study and public datasets by label transfer following data integration.
(**I and J**) Heatmaps showing the molecular features shared across different datasets (I) and developmental stages (J).

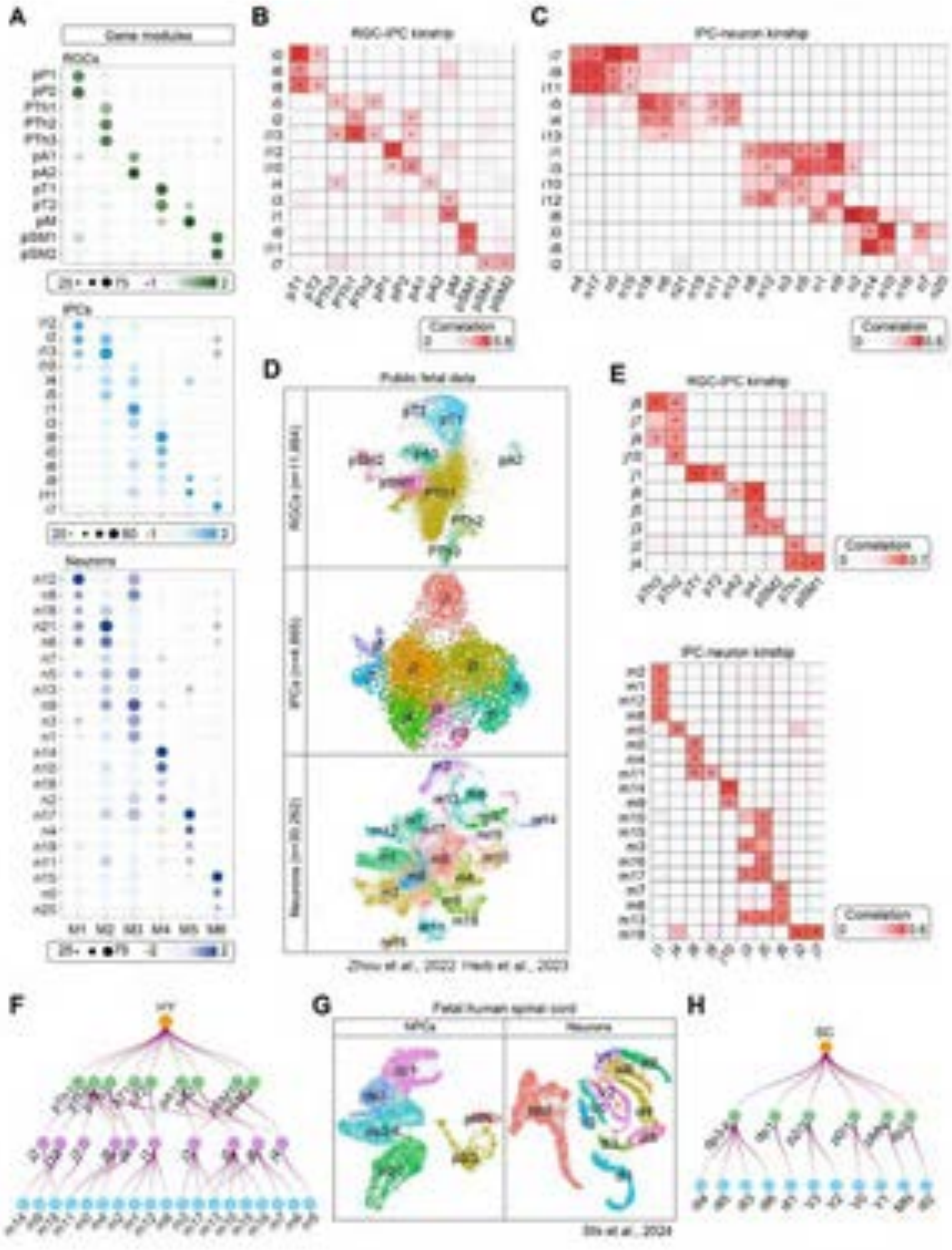


Figure S11: Inference of neuronal lineages, related to Figure 3.
(**A**) Modular transcriptomic analysis revealing that the region-specific gene modules in RGCs are transmitted to their progeny cells including both IPCs and neurons at the cell subtype level. Signature gene sets for each progenitor domain were identified by cosine similarity algorithm, with COSG score > 0.1.
(**B and C**) Heatmaps showing the pairwise correlation of RGC and IPC subtypes (B), as well as IPC and neuron subtypes (C). The cell subtype pairs with strong correlation coefficients (*P* value < 0.05) were marked with a "+" sign.

(**D**) Independent clustering of RGCs, IPCs, and neurons from public fetal data.
(**E**) Pairwise correlation of RGC, IPC and neuronal subtypes identified from public fetal data.
(**F**) Computational reconstruction of hypothalamic neuronal lineage tree using public data.
(**G**) Clustering analysis of RGCs and neurons public fetal human spinal cord data.
(**H**) Inference of neuronal lineage tree for the developing human spinal cord.

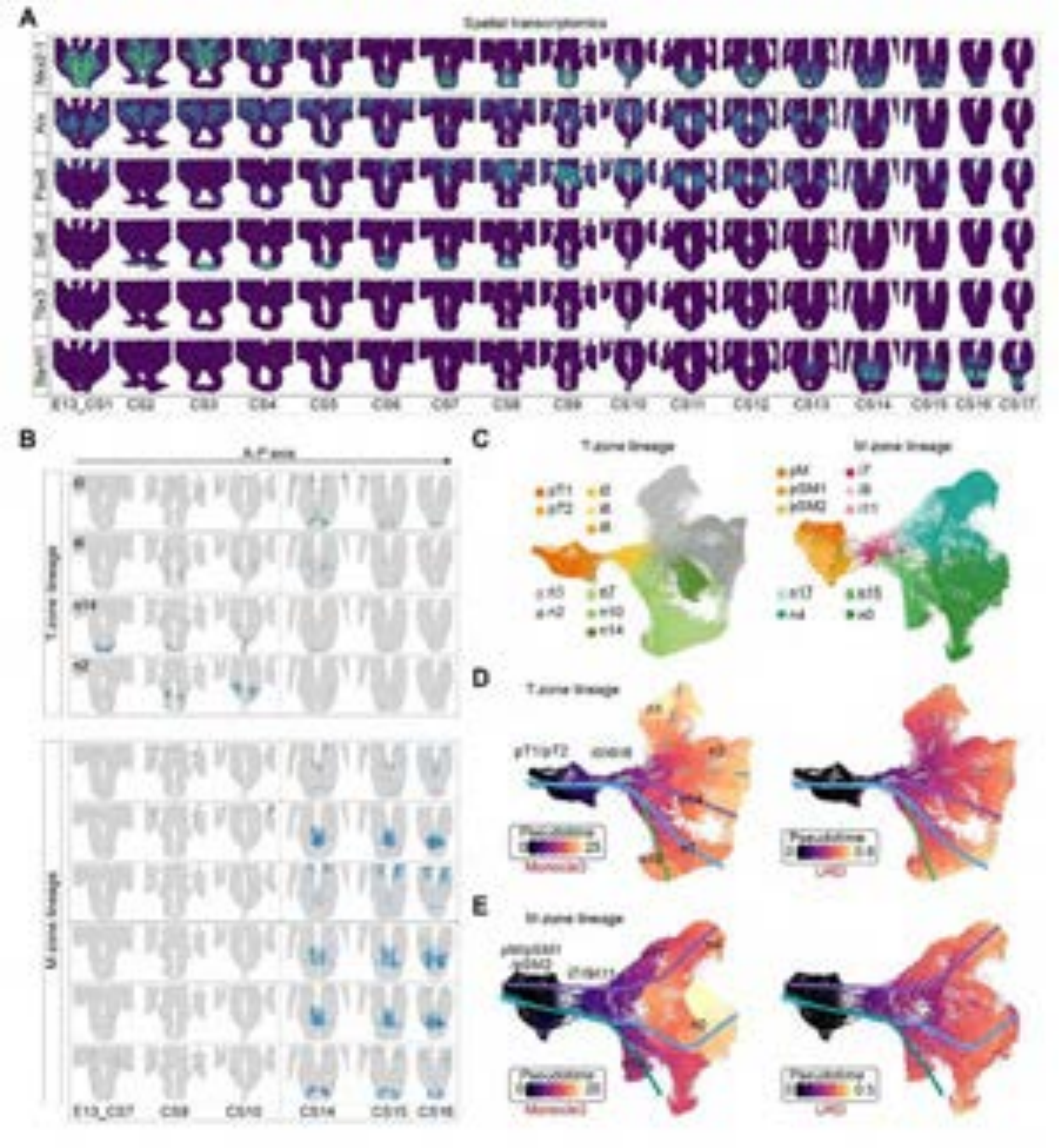


Figure S12: Spatial transcriptomic and trajectory analyses of hypothalamic sublineages, related to Figure 3.

(**A**) Serial coronal images showing the spatial expression of *Nkx2-1*, *Arx*, *Pax6*, *Six6*, *Tbx3* and *Barhl1* in the mouse hypothalamus at E13.5. These coronal brain sections (n=17), ordered along the anterior-posterior (AP) axis, were subjected to Stereo-seq for spatial transcriptomic analysis.

(**B**) Spatial visualization of signature gene modules for each progenitor and neuronal subtype within T- and M-zone lineages in E13.5 mouse hypothalamus.

(**C**) Subclustering analysis of cells in the inferred T-zone (left) and M-zone (right) lineages.

(**D and E**) Pseudotemporal analysis of T-zone (D) and M-zone (E) lineages using monocle 3 and URD approaches. Solid and dashed lines show the differentiation paths of neural progenitors in T- or M-zones.

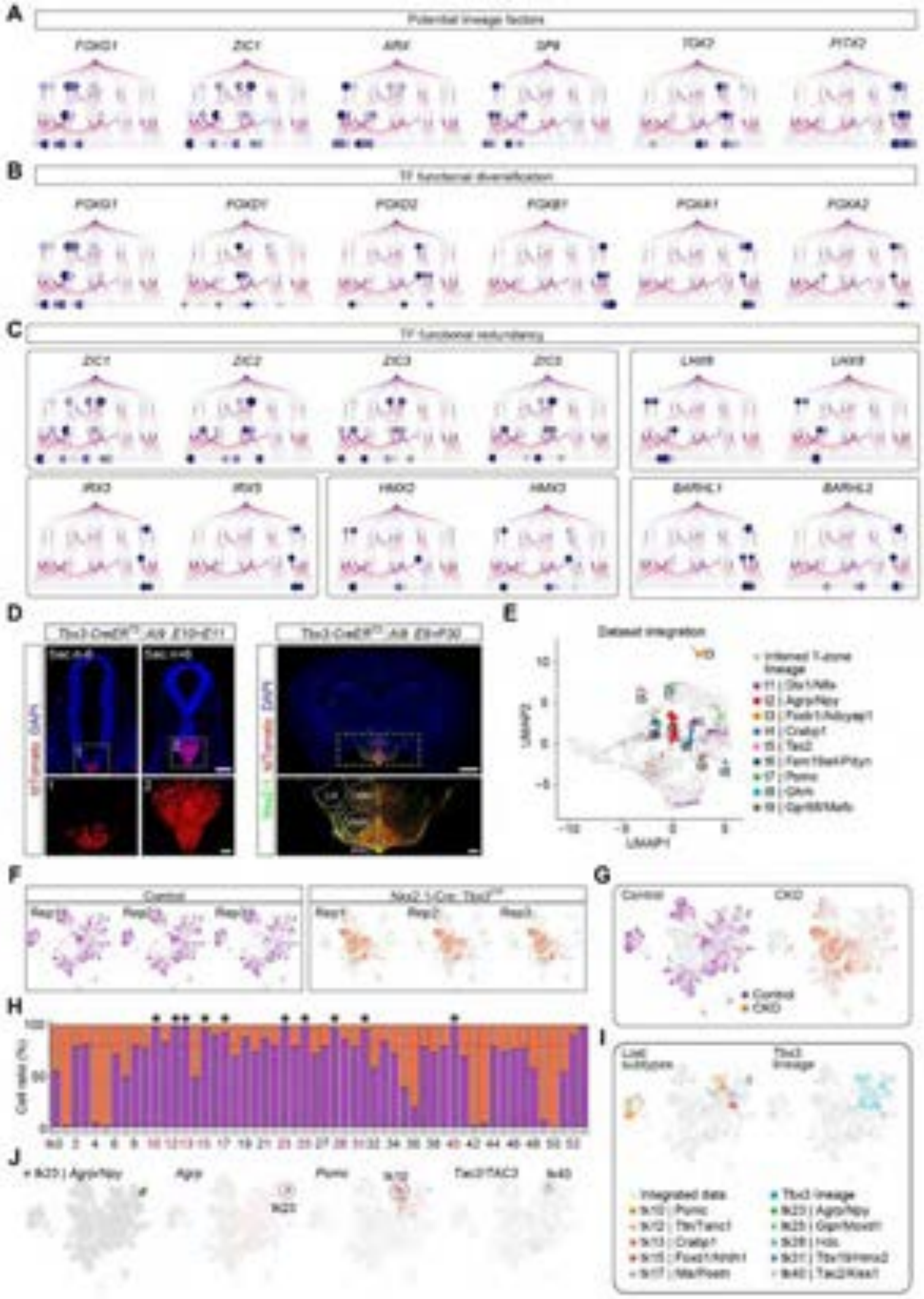


Figure S13: Inference and functional validation of potential lineage factors, related to Figure 3.

(**A**) Crown diagrams showing the candidate lineage factors specifying distinct hypothalamic sublineages.

(**B and C**) Crown diagrams demonstrating the functional diversification (B) and redundancy (C) of potential lineage-specific TFs in individual multigene families.

(**D**) Representative images showing genetic labeling of pT2 progenitor domain by tdTomato and spatial distribution of progeny neurons (co-stained with Nkx2.1) using *Tbx3-CreER*^{T2}*::Ai9* mice, which were induced with tamoxifen at E10 and sacrificed at E11 (left) or P30 (right). LH, lateral hypothalamus; DMH, dorsomedial nucleus; VMH, ventromedial nucleus; TU, tuberal nucleus; ARC, arcuate nucleus. Scale bars, 200 and 50 µm (left panels); 1000 and 200 µm (right panels).

(**E**) Mapping of all single-cell data from Tbx3-derived lineage onto UMAP plot that represents T-zone lineage. Experimentally traced neuronal subtypes are color-coded.

(**F**) Sample replicates of snRNA-seq data from both control and Tbx3 conditional knockout (CKO) mice at P14. We bred *Nkx2.1-Cre::Tbx3*$^{F/F}$ mice to delete *Tbx3* gene in the hypothalamus.

(**G**) UMAP plots showing cell composition and distribution in control and CKO datasets. Control cells are labeled in purple, while CKO cells are in orange.

(**H**) Stacked bar chart showing the relative ratios of control and CKO cells in each neuronal subtype. The red dashed line indicates when the proportion of CKO cells falls below 20%. Noteworthy, the neuronal subtypes with a significant reduction in CKO groups are marked with asterisks and red text. We calculated the statistical significances of these difference using three softwares: scCODA, Cacoa, and speckle.

(**I**) UMAP plots highlighting the lost neuronal subtypes in CKO groups (left) and Tbx3-derived (pT2) neuronal sublineage (right) in colors. Single-cell data of inferred pT2 sublineage was randomly under-sampled and reversely mapped onto the UMAP plot integrating control and CKO neurons. Lost subtypes of neurons in CKO mice are labeled in distinct colors.

(**J**) UMAP plots showcasing three neuronal subtypes (AgRP, POMC and KNDy) lost in CKO groups.

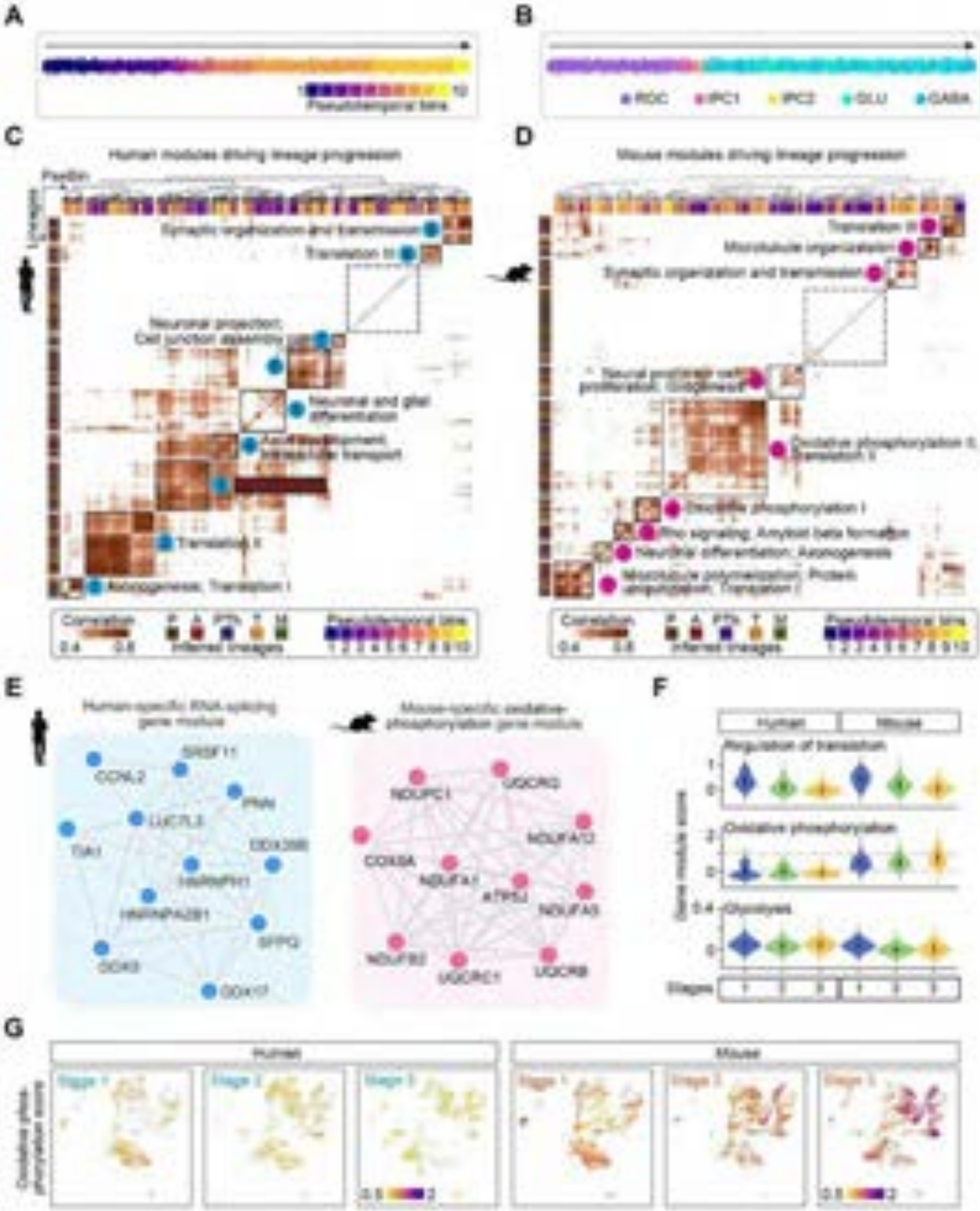


Figure S14: Conservative and divergent gene modules in human and mouse neuronal lineages, related to Figure 3.

(**A and B**) Beeswarm plots showing the subdivision of hypothalamic neuronal lineage into 10 pseudo-temporal bins (A) and 5 cell groups (B).

(**C and D**) Main heatmaps depicting pairwise similarities between all NMF transcriptional programs in human (C) and mouse (D) neuronal lineages, ordered by hierarchical clustering. Nine meta-programs are indicated by squares and numbers, and annotated by GO enrichment analysis.

(**E**) Network graphs showing the protein interaction networks for the human-specific RNA-splicing gene module (left) and mouse-specific oxidative-phosphorylation gene module (right).

(**F**) Violin plots showing the gene module scores of protein translation, oxidative phosphorylation and glycolysis across species at various developmental stages.

(**G**) Feature plots showing the divergent oxidative phosphorylation score between humans and mice at different developmental stages.

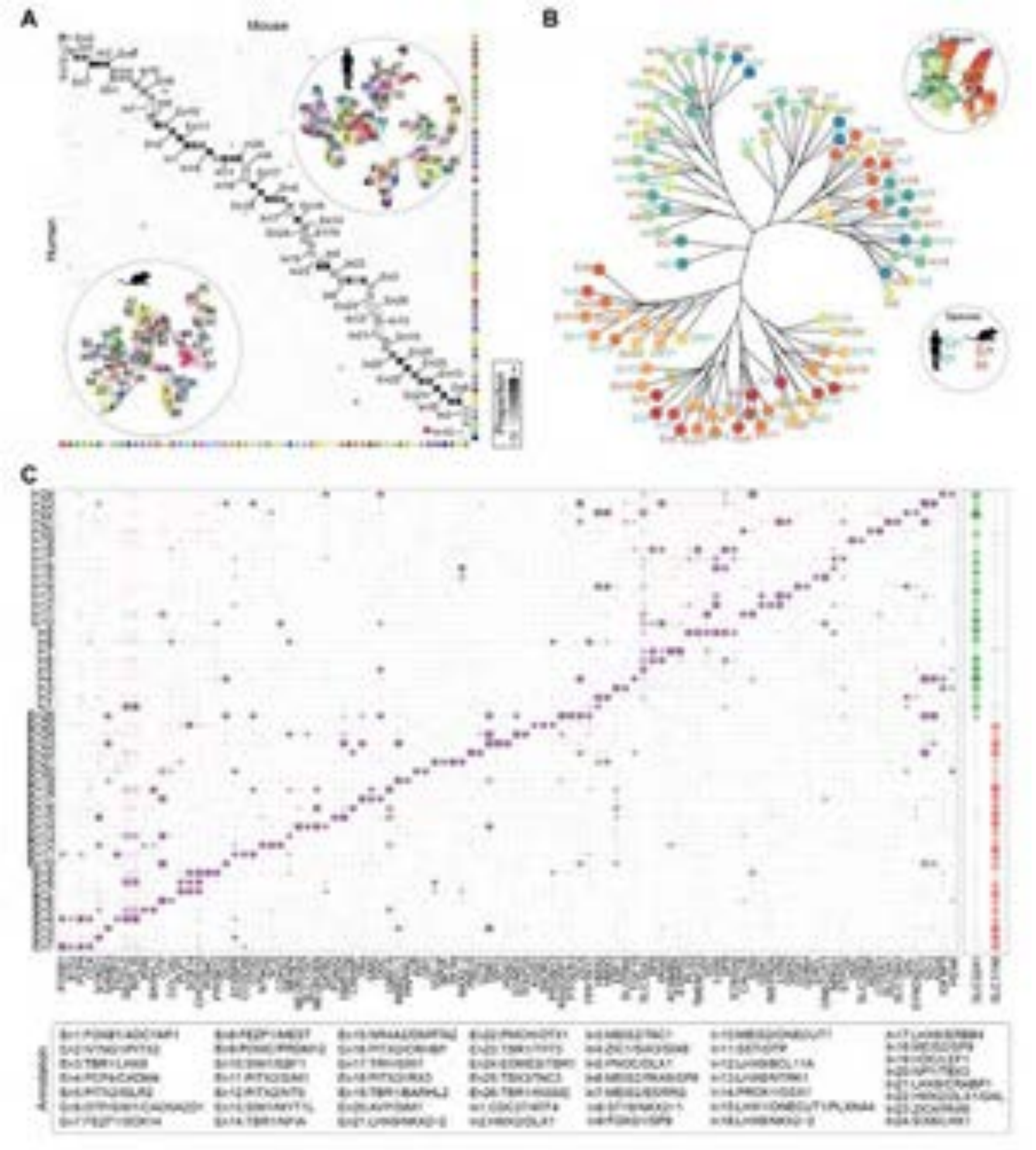


Figure S15: Molecular annotation of distinct neuronal subtypes, related to Figure 4.

(**A**) UMAP plots and heatmap illustrating the independent clustering of human and mouse neurons and their inter-species conservation, respectively.

(**B**) Dendrogram tree illustrating the hierarchical clustering of diverse neuronal subtypes across species. Glutamatergic and GABAergic neurons are distinguished by warm (red, orange, and yellow) and cool (green and blue) colors, respectively. Neurons from humans and mice are labeled in blue and red, respectively.

(**C**) Heatmap plot depicting the specific marker genes for each neuronal subtype. Excitatory (En) and inhibitory (In) neurons are distinguished by expression of *SLC17A6* and *SLC32A1*, respectively. Neurons are annotated by neurotransmitters, neuropeptides and specific TFs.

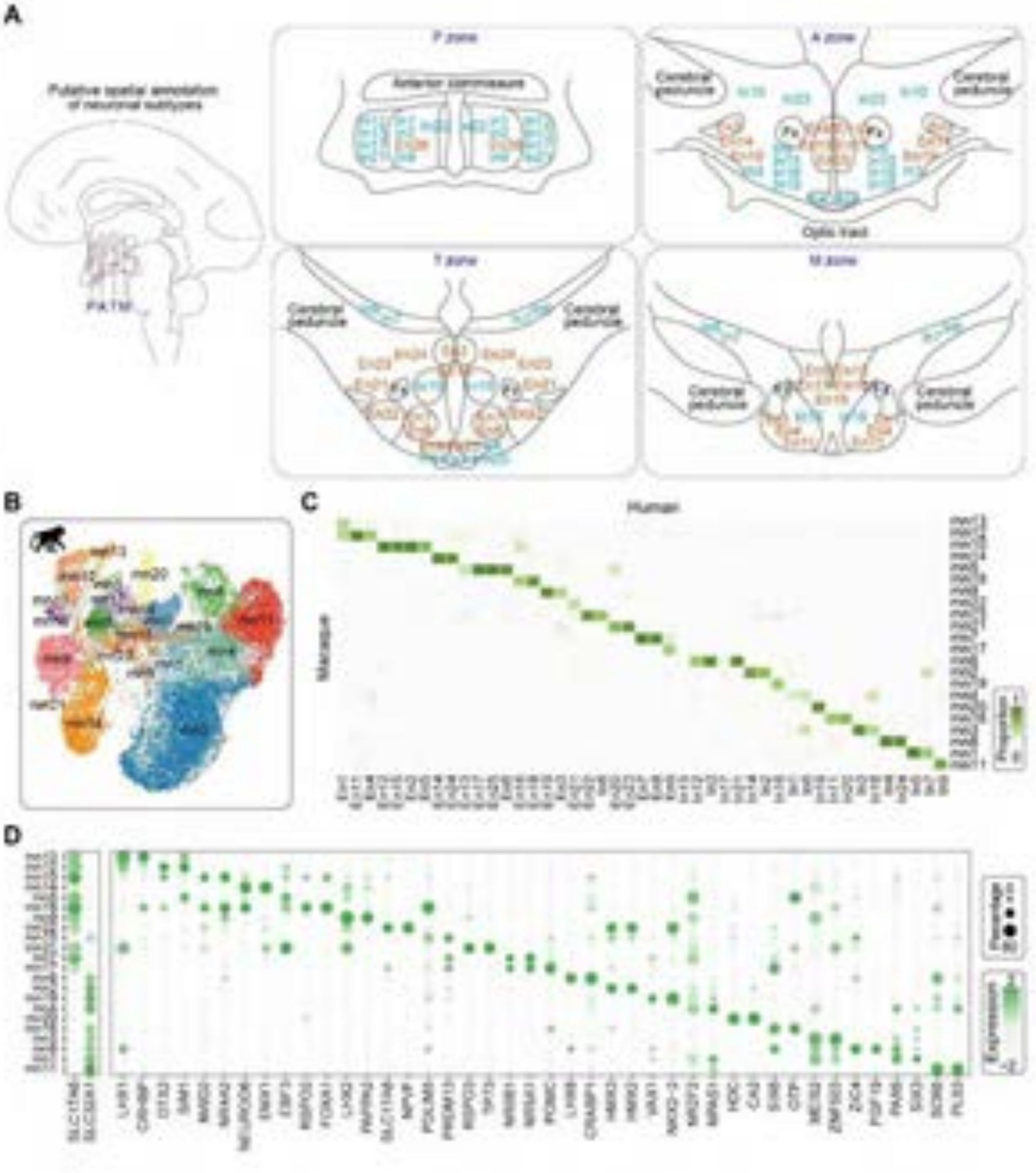


Figure S16: Spatial annotation of distinct neuronal subtypes conserved between humans and macaques, related to Figure 4.

(**A**) Schematic diagrams depicting the potential spatial distribution of distinct neuronal subtypes in P, A, T and M zones along the AP axis. Spatial position of most, if not all, neuronal subtypes in diverse hypothalamic nuclei was predicted and annotated by the expression of known spatial codes, neuropeptides or other canonical markers[3-7]. For example, TRH (En17) and AVP (En20) endocrine neurons were distributed in paraventricular nucleus, while POMC (En9), AgRP (In20) and KNDy (En25) were located in arcuate nucleus.

(**B**) Independent clustering of macaque neurons.

(**C**) Identification of neuronal-subtype homology between humans and macaques.

(**D**) Dot plot showing the molecular signatures in each macaque neuronal subtype.

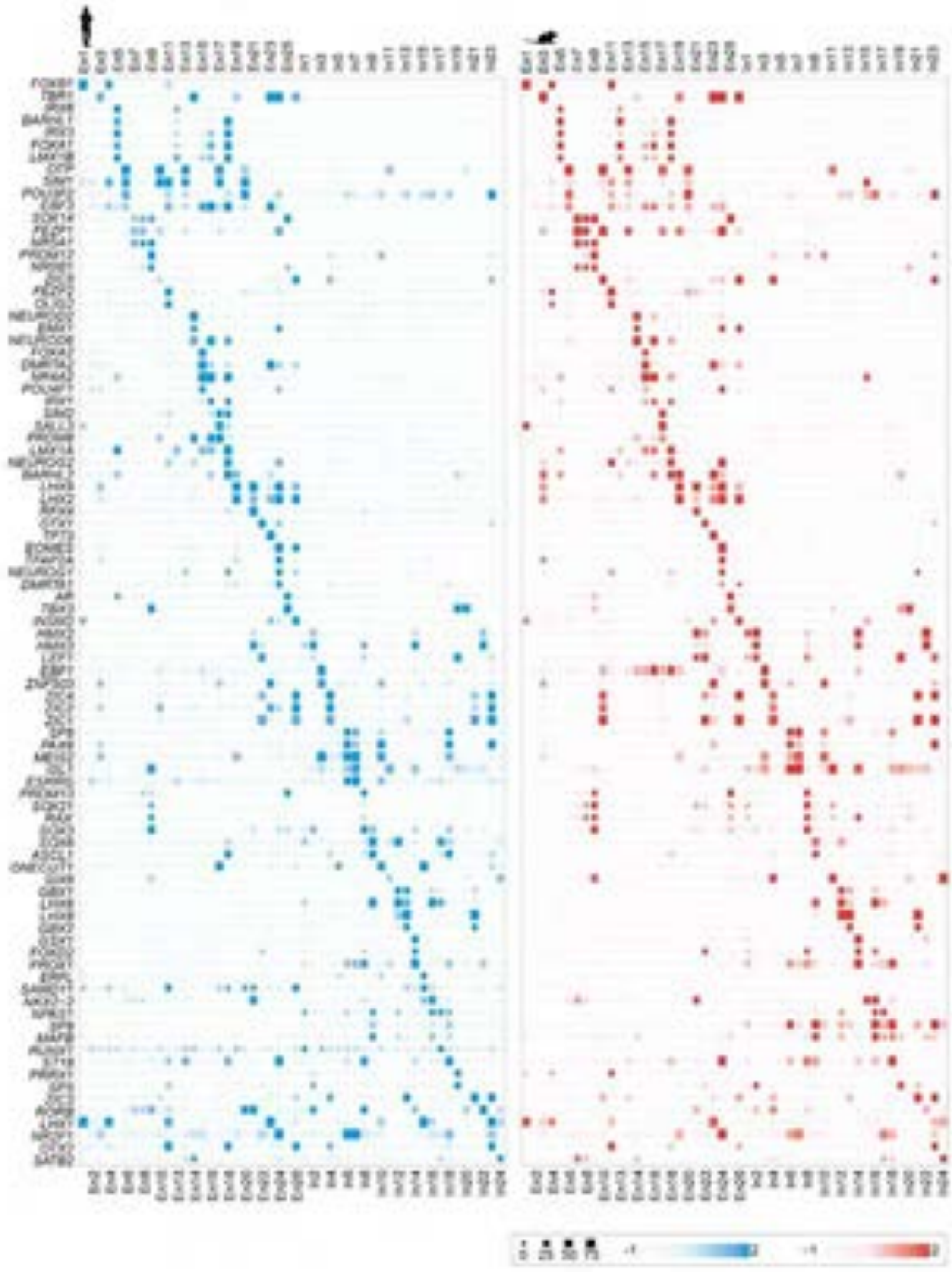


Figure S17: Neuronal subtype-specific TFs conserved between humans and mice, related to Figure 4.
Heatmaps showing the specific expression of 89 conservative TFs in each neuronal subtypes within both the human (left) and mouse (right) developing hypothalamus.

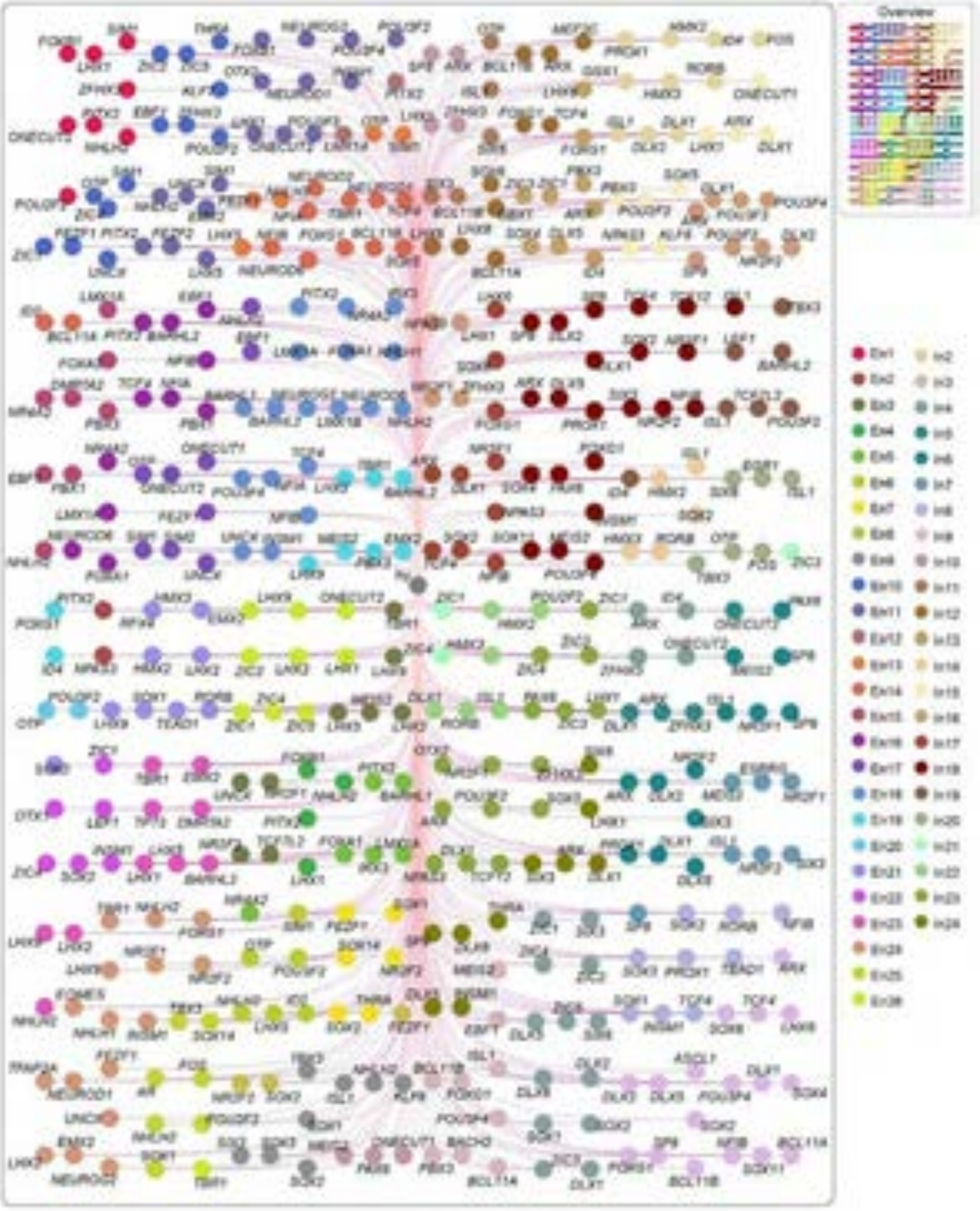


Figure S18: Potential neuronal subtype-specific regulons predicted by CESNI, related to Figure 4.
An overview of the cell type-specific gene regulatory networks inferred by CESNI (STAR Methods) in the developing hypothalamus. The color-coded nodes represent potential regulons in different neuronal subtypes.

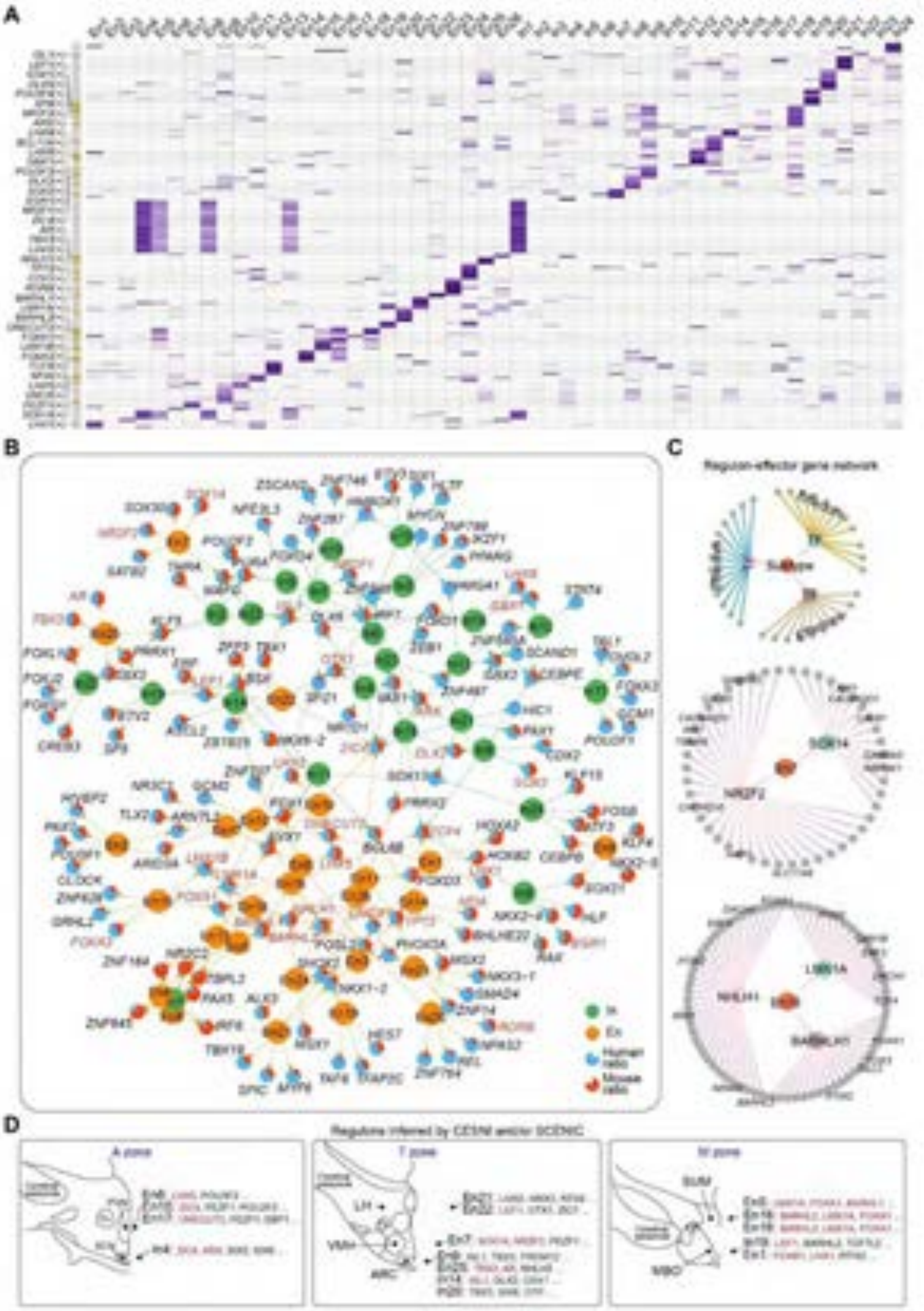


Figure S19: Neuron subtype-specific regulons predicted by SCENIC, related to Figure 4.
(**A**) A heatmap showing the activity of 174 potential regulons for distinct neuronal subtypes in the hypothalamus, which were inferred by SCENIC.

(**B**) Network analysis of the top 6 potential regulons for each neuronal subtype. Excitatory and inhibitory neurons are indicated by orange and green nodes, respectively. Pie charts for each regulon reflect its relative enrichment in mouse (red) and human (blue) neurons.

(**C**) Sample network diagrams illustrating the potential master regulators and their effector genes in En7 (middle) and En18 (bottom).

(**D**) A schematic summarizing the potentially critical regulons for many neuronal subtypes distributed in paraventricular nucleus (PVN), suprachiasmatic nucleus (SCN), LH, VMH, ARC, supramammillary nucleus (SUM) and mammillary body (MBO). Fx, fornix. Regulons identified by both CESNI and SCENIC are labeled in red.

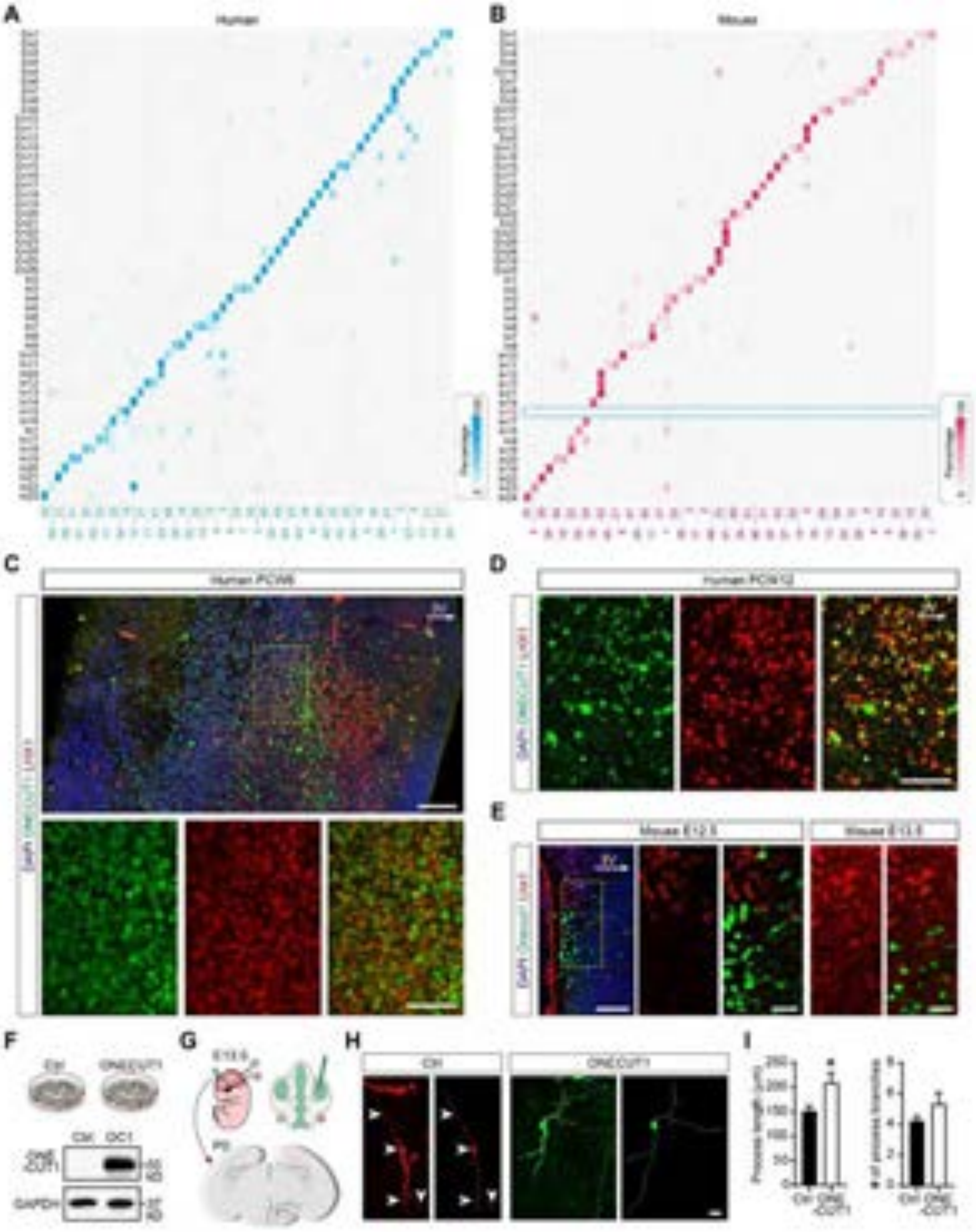


Figure S20: Identification of human-enriched neuronal subtype, related to Figure 4.

(**A and B**) Heatmaps showing human and mouse neuronal-subtype homologies, denoted by the proportion of cells overlapping between human and integrated neurons (A), or between mouse and integrated neurons (B). In15 subtype does not correspond to any mouse neuronal subtypes.

(**C**-**E**) Representative confocal images showing the species-specific prevalence of ONECUT1$^+$LHX1$^+$ In15 neurons in the human hypothalamus in comparison to mouse brains at different developmental time points (PCW6 and12 for humans; E12.5 and E13.5 for mice). 3V, third ventricle. Scale bars, 100 and 50 μm in (C) and (D); 50 and 20 μm in (E).

(**F**) Immunoblotting showing the successful expression of human ONECUT1 proteins in HEK293T cells transfected with the plasmid pCDH-EF1α-ONECUT1-HA-T2A-EGFP. OC1, ONECUT1.

(**G**) Schematic diagram illustrating the strategy for ectopic expression of human ONECUT1 in the embryonic mouse hypothalamus through *in utero* electroporation.

(**H**) Confocal images and reconstruction of neurons that were transfected with either tdTomato or ONECUT1-T2A-EGFP *in vivo*. Embryos at E13.5 were electroporated with plasmids *in utero* and collected after 5 days. Scale bar, 20 μm.

(**I**) Bar plots showing the process length and branch numbers in neurons (Control, n=20; ONECUT1, n=22). Data are shown as mean ± SEM ($*P < 0.05$ by two-tailed unpaired Student's *t* test).

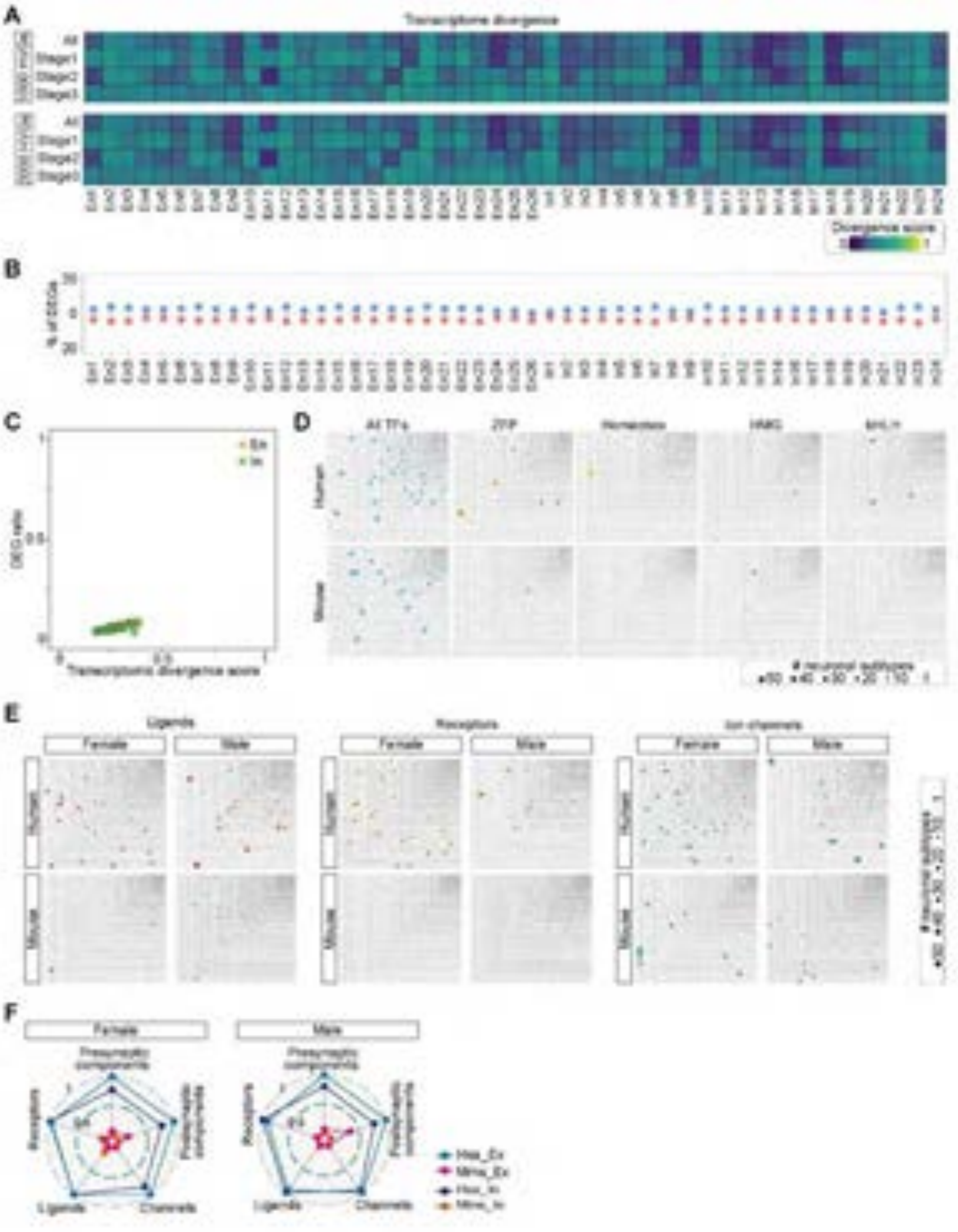


Figure S21: Conservation and diversity of hypothalamic neurons across species and sexes, related to Figure 4.

(**A**) Heatmap showing the transcriptomic divergence of distinct neuronal subtypes between humans and mice, analyzed with the top 1000 (top panel) and 2000 (bottom panel) highly variable genes (HVGs).

(**B**) Relative ratios of DEGs enriched in humans (blue) or mice (red) for each neuronal subtypes.

(**C**) Scatter plot showing the two-dimensional relationship between transcriptomic divergence score (x axis) and DEG ratio (y axis) for each excitatory (En) and inhibitory (In) neuronal subtype.

(**D**) Treemaps showing the frequency of human- or mouse-enriched TFs, including zinc finger proteins (ZFP), homeodomain family (Homeobx), high mobility group (HMG) and basic helix-loop-helix factors (bHLH). The size of each box represents the number of neural subtypes with species-specific gene enrichment.

(**E**) Treemap showing the frequency of human- or mouse-enriched ligands, receptors and ion channels in female and male samples.

(**F**) Normalized gene expression scores for ligands, receptors, channels, presynaptic and postsynaptic components in both human and mouse neurons across female and male samples. Ex, glutamatergic neurons; In, GABAergic neurons.

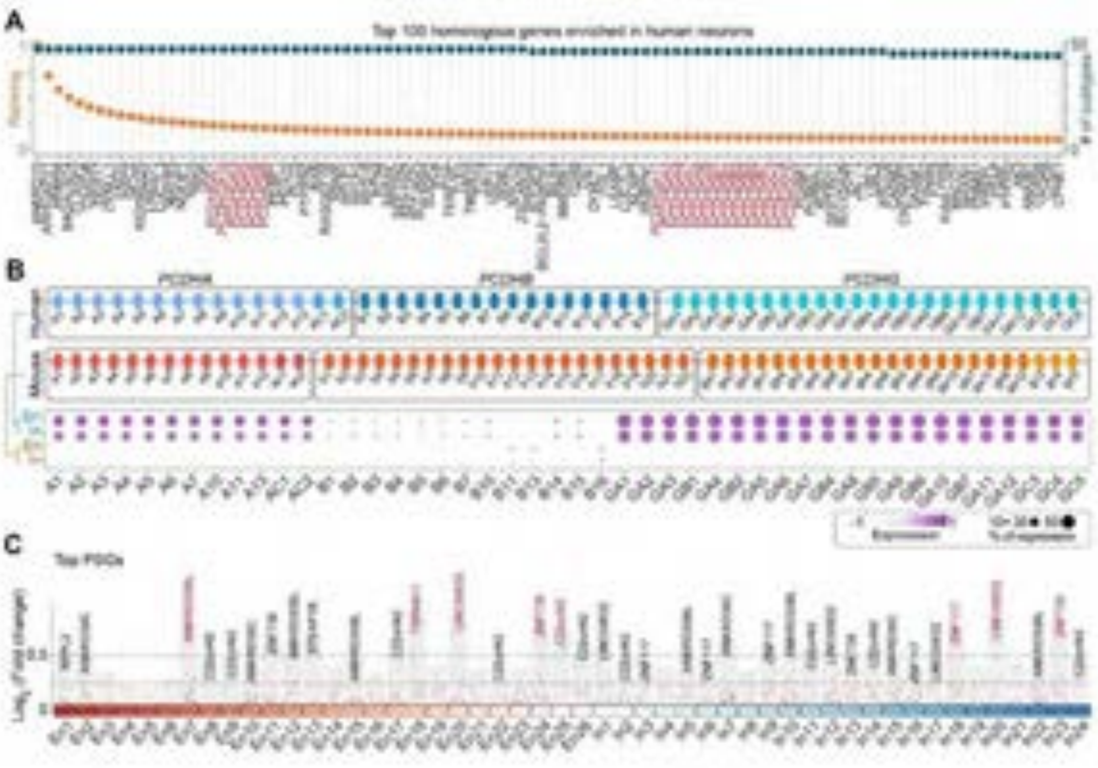


Figure S22: Genes enriched in human hypothalamic neurons, related to Figure 4.
(**A**) Ranked list of the top 100 orthologous genes that are upregulated in human neurons as compared with mouse neurons. Clustered protocadherin (cPCDH) genes are highlighted in red text.
(**B**) Dot plot showing the divergent expression level of cPCDH gene family in human and mouse neurons.
(**C**) Chart illustrating the subtype-specific primate-specific genes (PSGs) in human neurons. Genes with subtype specificity ($\log_2$FC > 0.5 and adjusted $P < 0.05$) are marked with red text.

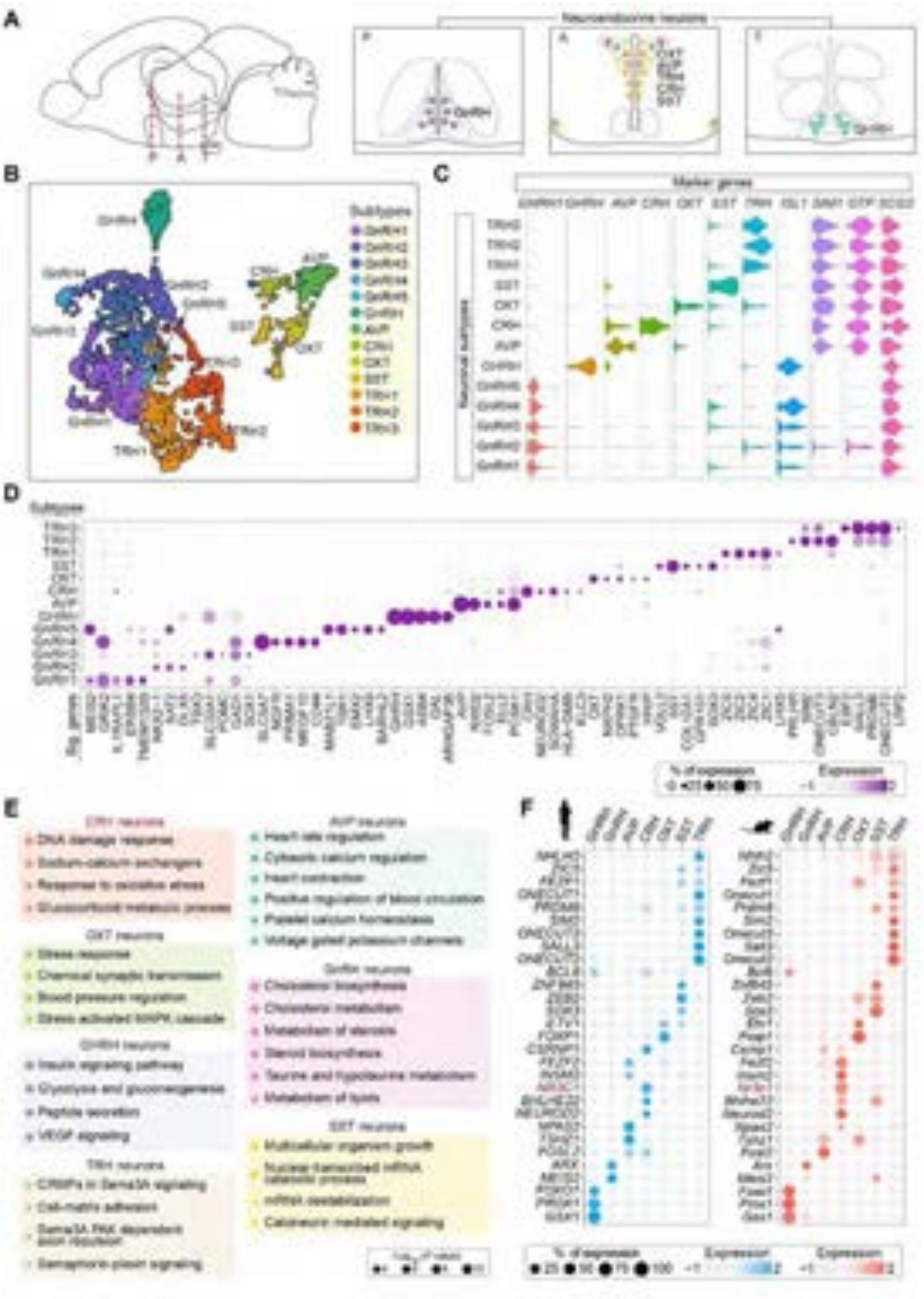


Figure S23: Molecular survey of hypothalamic neuroendocrine neurons during development, related to Figure 5.

(**A**) Schematic diagram illustrating the spatial distribution of neuroendocrine neurons, including GnRH, OXT, AVP, TRH, CRH, SST and GHRH subtypes, in the mouse hypothalamus.

(**B**) UMAP plot showing the subclusters of neuroendocrine neurons.

(**C**) Violin plots showing the expression pattern of *GNRH1*, *GHRH*, *AVP*, *CRH*, *OXT*, *SST*, *TRH*, *ISL1*, *SIM1*, *OTP* and *SCG2* in distinct subtypes of endocrine neurons.

(**D**) Dot plot showing the marker gene expression in various endocrine neuronal subtypes.
(**E**) GO functional enrichment analysis of molecular signatures for each neuroendocrine neuronal subtype.
(**F**) Subtype-specific TFs with species conservation. NR3C1 expression emerges early in prenatal CRH neurons.

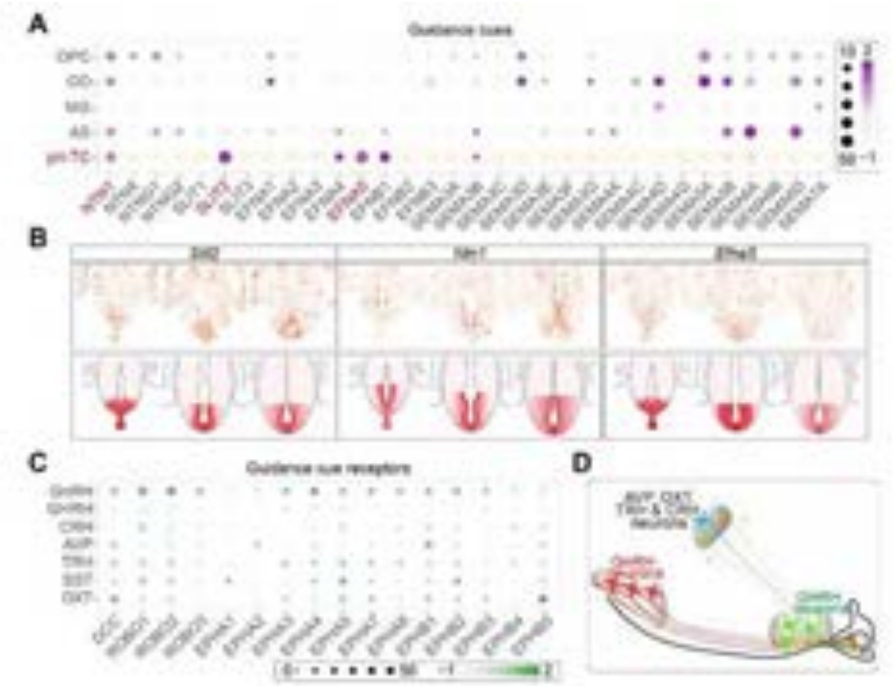


Figure S24: Expression patterns of guidance cues and their receptors in hypothalamic neuroendocrine neurons, related to Figure 5.

(**A**) Dot plot showing the relative gene expression of canonical axon guidance cues in OPC, OD, MG, AS and primitive tanycytes (pri-TC). *NTN1*, *SLIT2* and *EFNA5* are highly expressed in pri-TC.

(**B**) Spatial visualization of *Slit2*, *Ntn1* and *Efna5* expression in the E13.5 mouse hypothalamus with spatial transcriptomic data.

(**C**) Dot plot showing the expression of receptors for *NTN1*, *SLIT2* and *EFNA5* in neuroendocrine neurons.

(**D**) A schematic depicting the axonal projection of neuroendocrine neurons toward hypothalamic median eminence. margin to the tuberal hypothalamus.

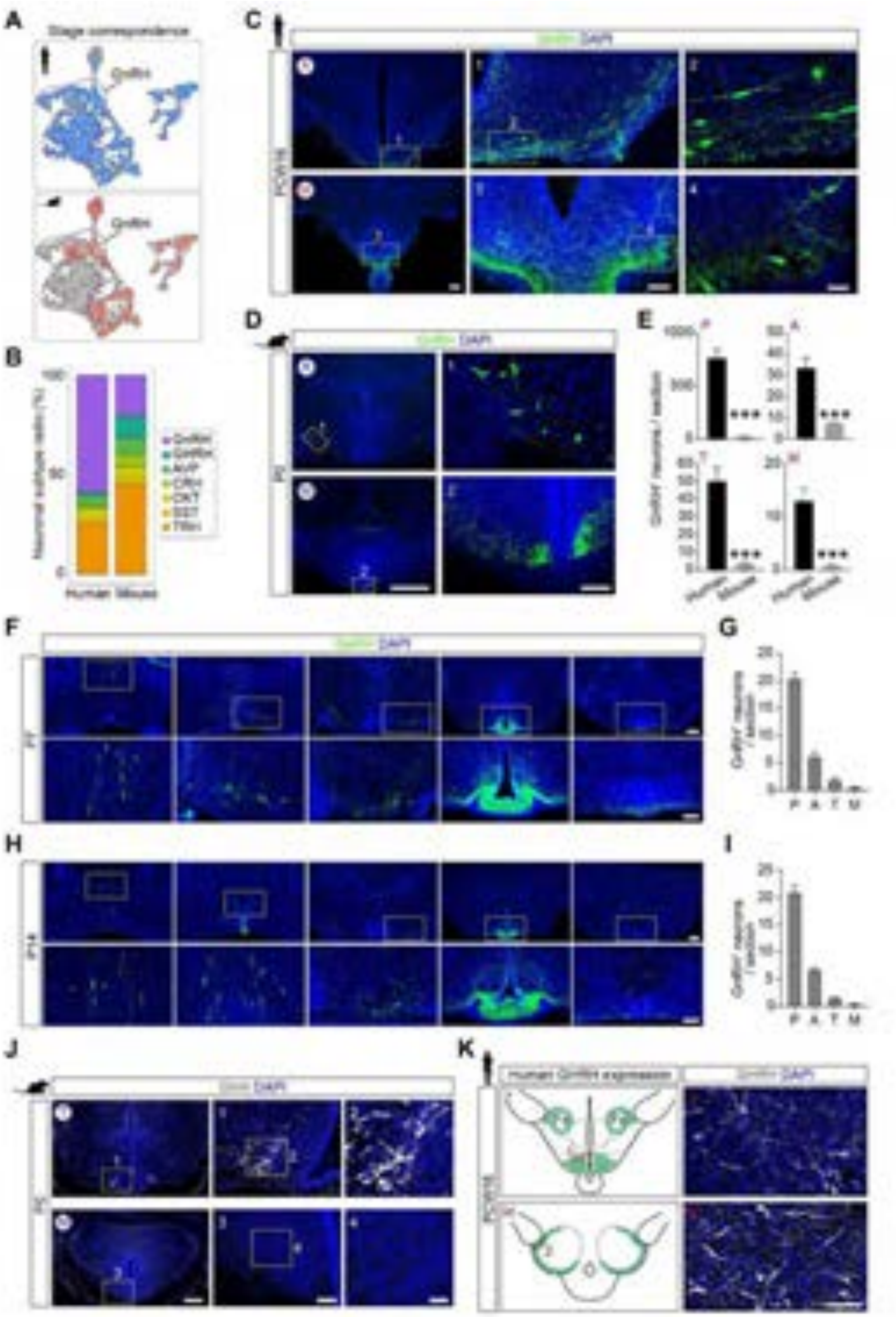


Figure S25: Interspecies differences of GnRH and GHRH neurons, related to Figure 5.

(**A**) UMAP plots showing the species difference of neuroendocrine neurons between humans and mice. Purple dashed lines outline the distribution of GnRH neurons.

(**B**) Stacked bar plot showing the relative proportion of each neuroendocrine neuronal subtype in humans and mice.

(**C and D**) Representative images showing the spatial dispersion of GnRH neurons in the A- and M-zones of human (PCW16, left) and mouse (P0, right) hypothalamus. Nascent GnRH neurons migrate along the ventrolateral margin to the tuberal hypothalamus. Scale bars, 500, 200 and 50 μm (C); 500 and 50 μm (D).

(**E**) Quantification of GnRH neuronal numbers in the human and mouse developing hypothalamus. Values represent mean ± SEM (n = 5 human and 15 mouse brain sections, $^{***}P < 0.001$ by two-tailed unpaired Student's *t* test).

(**F**-**I**) Representative confocal images and quantification of GnRH neurons in the P7 and P14 mouse hypothalamus along the rostrocaudal axis. Scale bars, 200 and 75 μm.

(**J and K**) Schematic diagrams and sample images demonstrating the widespread distribution of GHRH neurons in the T- and M-zones of the human hypothalamus compared to mice. Scale bars, 200, 50 and 20 μm (J), and 50 μm (K).

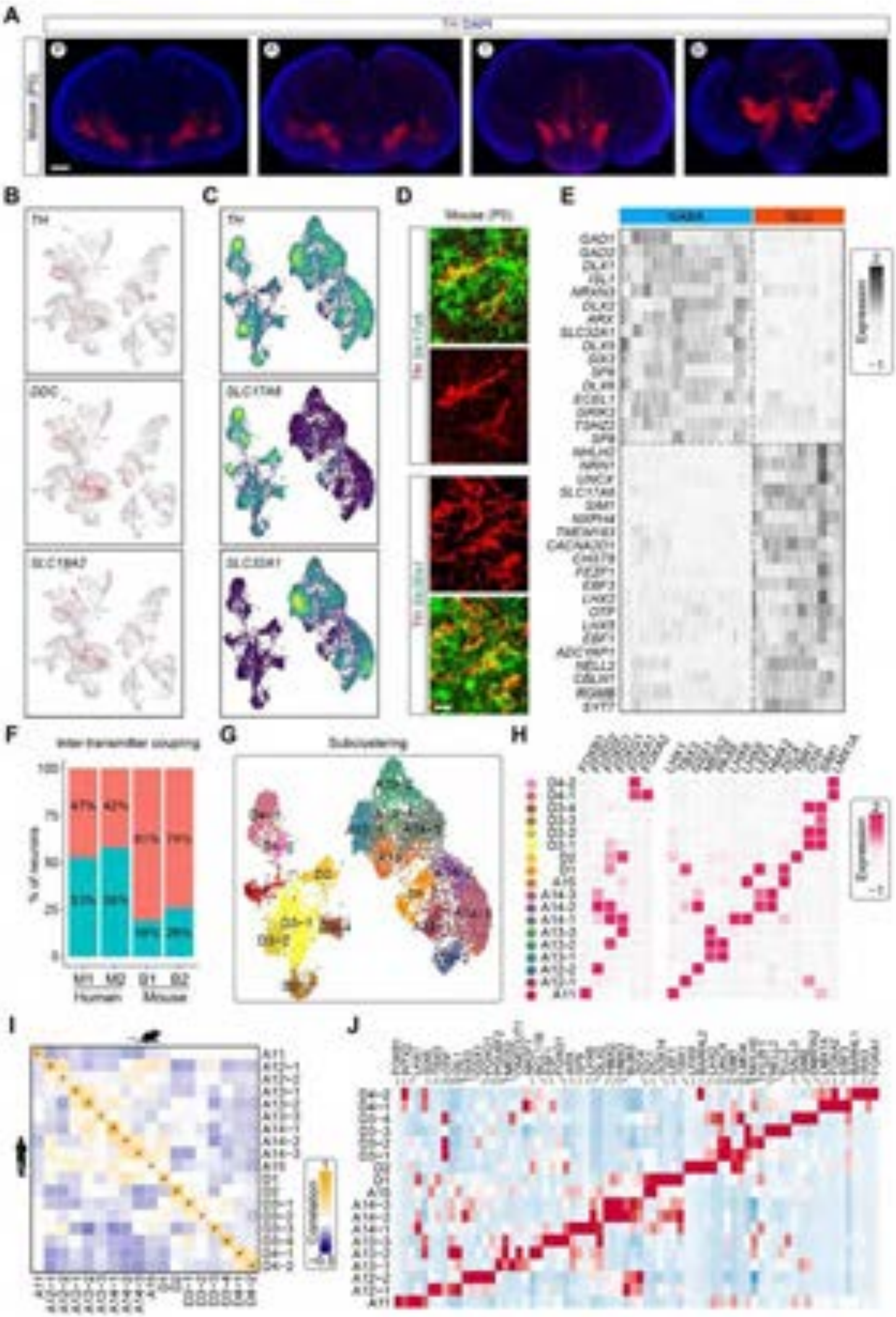


Figure S26: Molecular survey of hypothalamic dopamine neurons during development, related to Figure 5.

(**A**) Serial images showing the distribution pattern of TH-expressing neurons in the mouse hypothalamus at P0. Scale bars, 500 µm.

(**B**) Expression of *TH*, *DDC* and *SLC18A2* in the hypothalamic neurons.
(**C**) UMAP plots showing the overlap of *TH* expression with either *SLC17A6* or *SLC32A1*, suggesting the potential coupling between dopamine and glutamate or GABA transmitter.
(**D**) Co-labeling of TH with neurotransmitter transporters in the P0 mouse hypothalamus. *Slc17a6* and *Slc32a1* were detected by HCR-based *in situ* hybridization. Scale bars, 10 μm.
(**E**) Heatmap showing the divergent transcriptional features between GABAergic and glutamatergic TH-expressing neurons.
(**F**) Stacked bar plots showing the relative ratio of GABAergic and glutamatergic TH-expressing neurons in the human and mouse hypothalamus. Dopamine-glutamate coupling is more abundant in humans than in mice, regardless of data collection methods (M1 or M2) or batches (B1 or B2).
(**G**) Subclustering analysis of TH-expressing neurons in the hypothalamus.
(**H**) Heatmap showing the subtype-specific expression of distinct TFs among TH-expressing neurons.
(**I**) Pearson correlation analysis of diverse human and mouse dopamine neuronal subtypes.
(**J**) Heatmap showing the molecular signatures of each hypothalamic dopamine and dopamine-like neuron.

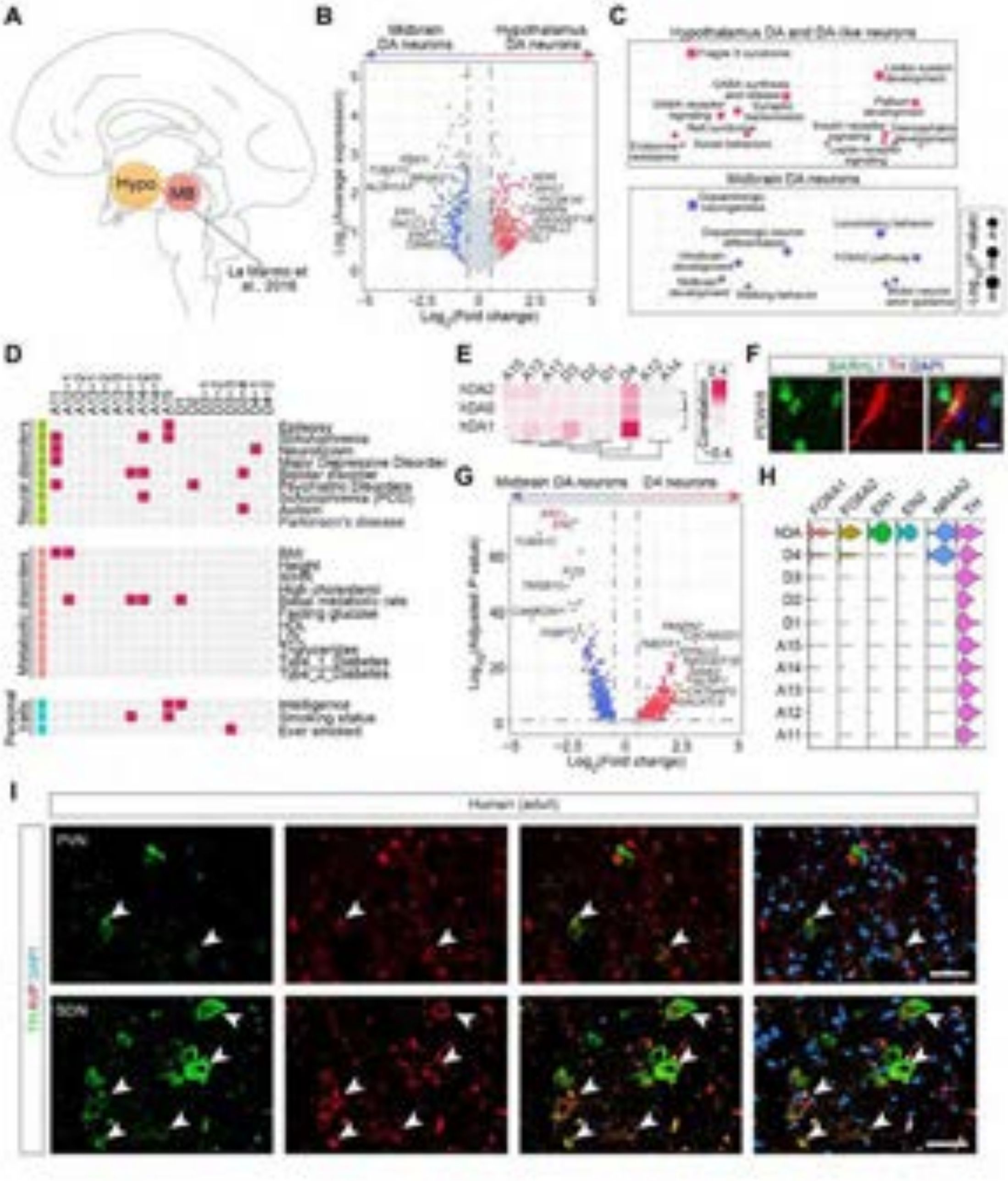


Figure S27: Hypothalamic dopamine neurons transcriptionally diverge from midbrain dopamine neurons, related to Figure 5.

(**A**) Anatomic location of hypothalamus and midbrain in the human brains.

(**B**) Volcano plot showing the genes that are differentially expressed in the hypothalamic and midbrain dopamine (DA) neurons[8]. Genes with significant differences ($\log_2FC > 0.5$ and adjusted P value < 0.05) are highlighted in colors.

(**C**) GO functional enrichment analysis of molecular signatures in both hypothalamic and midbrain DA neurons. Hypothalamic dopamine neurons show intricate links with metabolism, endocrine regulation, social behavior and forebrain development.

(**D**) GWAS disease and trait association of TH-expressing neurons in the hypothalamus.

(**E**) Pearson correlation between midbrain and hypothalamic DA neurons, suggesting a potential transcriptional similarity between the D4 subtype and midbrain DA neurons.

(**F**) Costaining of TH and BARHL1 in the human D4 neuronal subtype at PCW16. Scale bar, 10 μm.
(**G**) Volcano plot showing the genes differentially expressed in the D4 subtype and midbrain DA neurons. The marker genes of midbrain DA neurons, *EN1* and *EN2*, are denoted in red.
(**H**) Violin diagram illustrating the enriched expression of *EN1*, *EN2*, *FOXA1* and *FOXA2* in midbrain DA neurons, while *NR4A2* expression is shared between D4 subtype and midbrain DA neurons.
(I) Immunostaining showing the co-expression of TH and AVP in the adult human PVN and supraoptic nucleus (SON). Scale bars, 50 μm.

Supplemental references

1. Workman, A.D., Charvet, C.J., Clancy, B., Darlington, R.B., and Finlay, B.L. (2013). Modeling Transformations of Neurodevelopmental Sequences across Mammalian Species. The Journal of Neuroscience 33, 7368-7383. 10.1523/jneurosci.5746-12.2013.
2. Chen, A., Liao, S., Cheng, M., Ma, K., Wu, L., Lai, Y., Qiu, X., Yang, J., Xu, J., Hao, S., et al. (2022). Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. Cell 185, 1777-1792.e1721. 10.1016/j.cell.2022.04.003.
3. Kim, D.W., Washington, P.W., Wang, Z.Q., Lin, S.H., Sun, C., Ismail, B.T., Wang, H., Jiang, L., and Blackshaw, S. (2020). The cellular and molecular landscape of hypothalamic patterning and differentiation from embryonic to late postnatal development. Nature Communications 11. 10.1038/s41467-020-18231-z.
4. Briscoe, J., Pierani, A., Jessell, T.M., and Ericson, J. (2000). A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. Cell 101, 435-445. 10.1016/s0092-8674(00)80853-3.
5. Alvarez-Bolado, G., Paul, F.A., and Blaess, S. (2012). Sonic hedgehog lineage in the mouse hypothalamus: from progenitor domains to hypothalamic regions. Neural Development 7, 4. 10.1186/1749-8104-7-4.
6. Kim, D.W., Liu, K., Wang, Z.Q., Zhang, Y.S., Bathini, A., Brown, M.P., Lin, S.H., Washington, P.W., Sun, C., Lindtner, S., et al. (2021). Gene regulatory networks controlling differentiation, survival, and diversification of hypothalamic Lhx6-expressing GABAergic neurons. Communications Biology 4. 10.1038/s42003-020-01616-7.
7. Morales-Delgado, N., Castro-Robles, B., Ferrán, J.L., Martinez-de-la-Torre, M., Puelles, L., and Díaz, C. (2013). Regionalized differentiation of CRH, TRH, and GHRH peptidergic neurons in the mouse hypothalamus. Brain Structure and Function 219, 1083-1111. 10.1007/s00429-013-0554-2.
8. La Manno, G., Gyllborg, D., Codeluppi, S., Nishimura, K., Salto, C., Zeisel, A., Borm, L.E., Stott, S.R.W., Toledo, E.M., Villaescusa, J.C., et al. (2016). Molecular Diversity of Midbrain Development in Mouse, Human, and Stem Cells. Cell 167, 566-580.e519. 10.1016/j.cell.2016.09.027.