Title: Spatially-resolved molecular sex differences at single cell resolution in the adult human hypothalamus

Authors

Bernard Mulvey[1,2], Yi Wang[3], Heena R. Divecha[1], Svitlana V. Bach[1,2], Kelsey D. Montgomery[1], Sophia Cinquemani[1], Atharv Chandra[1], Yufeng Du[1], Ryan A. Miller[1], Joel E. Kleinman[1,2], Stephanie C. Page[1,2], Thomas M. Hyde[1,2,4], Keri Martinowich[1,2,5,6], Stephanie C. Hicks[3,7,8,9], Kasper D. Hansen[3,7,8,10+], Kristen R. Maynard[1,2,5+]

Affiliations

1. Lieber Institute for Brain Development, Johns Hopkins Medical Campus, Baltimore, MD, 21205, USA
2. Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, Baltimore, MD, 21205, USA
3. Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, 21205, USA
4. Department of Neurology, Johns Hopkins School of Medicine, Baltimore, MD, 21205, USA
5. The Solomon H. Snyder Department of Neuroscience, Johns Hopkins School of Medicine, Baltimore, MD, 21205, USA
6. Johns Hopkins Kavli Neuroscience Discovery Institute, Baltimore, MD, USA
7. Department of Biomedical Engineering, Johns Hopkins School of Medicine, Baltimore, MD, 21218, USA
8. Center for Computational Biology, Johns Hopkins University, Baltimore, MD, 21205, USA
9. Malone Center for Engineering in Healthcare, Johns Hopkins University, Baltimore, MD, 21218, USA
10. Department of Genetic Medicine, Johns Hopkins School of Medicine, MD 21218, USA

+co-corresponding authors

Correspondence

Kristen R. Maynard
kristen.maynard@libd.org

Kasper D. Hansen
khanse10@jhu.edu

Abstract

The hypothalamus contains multiple regions, including the ventromedial hypothalamus (VMH) and arcuate (ARC), which are responsible for sex-differentiated functions such as endocrine signaling, metabolism, and reproductive behaviors. While molecular, anatomic, and sex-differentiated features of rodent hypothalamus are well-established, much less is known about these regions in humans. Here we provide a spatially-resolved single cell atlas of sex-differentially expressed (sex-DE) genes in human VMH and ARC. We identify neuronal populations governing hypothalamus-specific functions, define their spatial distributions, and show increased retinoid pathway gene expression compared to rodents. Within VMH and ARC, we find correlated autosomal expression differences localized to *ESR1/TAC3*-expressing and *CRHR2*-expressing neurons, and extensive sex-DE of genes linked to sex-biased disorders including autism, depression, and schizophrenia. Our molecular mapping of disease associations to hypothalamic cell types with established roles in sex-divergent physiology and behavior provides insights into mechanistic bases of sex bias in neurodevelopmental and neuropsychiatric disorders.

Introduction

The hypothalamus (HYP) is a specialized brain structure that modulates behavioral and physiological drives essential to survival–ranging from social behavior[1] to metabolism[2] (**Fig. 1A**). This functional breadth is mediated by different HYP regions ('nuclei') with distinct circuitry, neurotransmitters, and neuropeptides. The tuberal portion of HYP contains the ventromedial hypothalamus (VMH) and arcuate (ARC), two nuclei that in model species show sex-differentiated development[2] and sex divergence at the levels of behavior[3–5], physiology[6–8], cell types[9,10], and gene expression[8,11].

The rodent VMH has molecularly distinct anatomic subdivisions[3,12] linked to regulation of metabolism, appetite, and parenting behavior. VMH modulates both adrenal[13] and peripheral nervous catecholamine release[14,15], and contains a neuronal population expressing corticotropin-releasing hormone receptor 2 (*Crhr2*)[16], suggesting critical roles in regulating the autonomic nervous system and stress-related behaviors[17,18]. VMH roles in metabolism[19] include glucose-mediated modulation of brain-derived neurotrophic factor (BDNF) expression[7], potentially in a sex-differential fashion[20]. VMH is necessary for several sex-differentiated mouse behaviors, including mating, maternal aggression, and male fighting[1,21–23]. Both of these functions are partially mediated by ventral lateral subdivision (VMHvl) neurons coexpressing estrogen receptors (*Esr1*) and progesterone receptors (*Pgr*)[3,21,22] and sexually differentiated during developmental androgen exposure[10].

The ARC is adjacent to the VMH and serves two primary functions, the first of which is metabolic regulation. ARC neurons detect circulating nutrients and nutritional status hormones (e.g., leptin)—in part via primary cilia[24]—and release neuropeptides including ghrelin, growth hormone releasing hormone (GHRH)[25], neuropeptide Y (NPY), and agouti-related peptide (AgRP)[26] throughout the HYP and pituitary to modulate metabolism and behavior. Second, *Kiss1* (kisspeptin)-*Esr1* coexpressing ARC neurons regulate cyclic release of gonadotropin releasing hormone (GnRH) from HYP, resulting in pulsatile pituitary release of luteinizing hormone (LH) and follicle-stimulating hormone (FSH)[9,27,28]. *Kiss1*- and metabolism-regulating populations of ARC bidirectionally influence one another to alter metabolism/growth according to reproductive status in a sex-differentiated manner[6,29].

While the ARC and VMH have been extensively studied in rodents, little is known about their function in the human brain despite their fundamental importance to survival and evolution and potential roles in neuropsychiatric disorders. At the molecular scale, *in vitro* studies of human ARC-like neurons have identified enrichment in genomic regulators associated with major depressive disorder (MDD) or bipolar disorder[30]. Furthermore, diagnostic features of several neuropsychiatric disorders implicate VMH/ARC functions: examples include eating disorders, which by definition involve appetitive/metabolic (dys)regulation; major depressive disorder, which can involve increased or decreased appetite; and anxiety/posttraumatic stress disorders, which may stem from deficits in fear extinction (in part a VMH-regulated process[31]). Notably, all of these disorders have a sex-differential incidence[32–36]. Likewise, sex differences in molecular and behavioral phenotypes relevant to mental health are documented in many rodent HYP studies[1,3,4,21,22,31,37]. In sum, observations ranging from animal models to clinical phenotypes suggest that the HYP may mediate sex differences in molecular risk for—and presenting features of—neuropsychiatric/neurodevelopmental disorders, highlighting the importance of investigating molecular sex differences in human VMH and ARC.

Molecular profiling of the human HYP, especially individual subregions, has been limited. Previous analyses of bulk RNA-sequencing (RNA-seq) data from adult human tissue by UKBEC[38] and GTEX[39,40] revealed substantial sex differences in HYP, though these could not be attributed to specific regions[41,42]. To date, single-cell[43] and spatial transcriptomic[44] atlases from adult human HYP tissue have included only male samples and were underpowered for sex-differential analyses, respectively. To our knowledge, transcriptional sex differences within molecularly-defined HYP cell types have been the focus of only one report, with limited information on tissue anatomy and sample handling[45]. Thus, significant questions remain regarding the regional enrichment and cell type-specificity of molecular sex differences in the human VMH and ARC.

Here, we utilized the 10x Genomics Visium and Xenium platforms to profile spatial gene expression in human postmortem VMH and ARC from 8 psychiatrically healthy adult donors evenly split by sex. A preliminary analysis of transcriptome-wide Visium data was used to prioritize genes for follow up at single cell resolution with a custom gene panel using Xenium. Using these complementary platforms, we identify marker genes and molecularly distinct cell types for VMH and ARC and highlight conserved and divergent features relative to rodents. Sex-differential expression (sex-DE) analyses in VMH and ARC revealed that the transcriptome-wide effects of sex are well-correlated between these two regions. Further, we identify potential transcription factors (TFs) whose putative regulatory targets are enriched for sex-DE, including gonadal hormone- and retinoid receptor-binding TFs. Across cell types residing in VMH or ARC, we identify 122 genes that are both sex-DE and linked to neuropsychiatric disorders with sex differences in prevalence or clinical features. Finally, we provide evidence that VMH and ARC cell types are enriched for sex-DE genes associated with disorders that show sex differences in prevalence rates, including autism, depression, and schizophrenia. This state-of-the-art multisample, multimodal spatial atlas of molecular features and sex differences in the adult human VMH and ARC is made available via interactive web resources at research.libd.org/spatial_HYP.

Methods

3.1 | Postmortem human tissue samples

Postmortem human brain tissue from 8 adult neurotypical donors of European ancestry was obtained at the time of autopsy following informed consent from legal next of kin, Office of the Chief Medical Examiner of the State of Maryland, under the Maryland Department of Health IRB protocol #12-24, from the Western Michigan University Homer Stryker MD School of Medicine, Department of Pathology under the WCG protocol #20111080, and from the NIMH Human Brain Collection Core, protocol #90-M-0142. Using a standardized strategy, all donors were subjected to clinical characterization and diagnosis. Macro- and microscopic neuropathological examinations were performed, and subjects with evidence of neuritic pathology or other neuropathological abnormalities were excluded. Details regarding tissue acquisition, processing, dissection, clinical characterization, diagnosis, neuropathological examination, RNA extraction and quality control (QC) measures have been previously published[47]. Demographics for the 8 neurotypical control donors included in the study are listed in **Table S1**. All subjects had BMIs between normal range and class I obesity (22.3-31.6) and no history of neuroendocrine disorders. Female subjects were premenopausal and not pregnant at time of death. Information on donor use of oral contraceptives was not systematically collected. Briefly, fresh frozen coronal brain slabs with clearly visible thalamus, hypothalamus, putamen, globus pallidus (internal and external segments), optic tract, anterior commissure, fornix, and posterior amygdala (MNI: -7.04) were selected for VMH and ARC dissection. Using a hand-held dental drill, tissue blocks of approximately 10 X 20 mm were dissected, encompassing the hypothalamus, optic tract, anterior commissure and fornix. White matter (WM) tracts were used as landmarks to identify the approximate location of VMH and ARC for microdissection. Tissue blocks were stored in sealed cryogenic bags at -80°C until cryosectioning.

3.2 | Tissue processing and quality control

At the time of cryosectioning, tissue blocks were acclimated to the cryostat (Leica CM3050s) at -14°C and mounted onto a chuck with OCT (TissueTek Sakura). Approximately 50 µm of tissue was trimmed from the block to achieve a flat surface, and several 10 µm sections were mounted onto pre-chilled glass slides for anatomical validation quality control (QC) experiments using single molecule fluorescent in situ hybridization (smFISH) with RNAScope technology (Advanced Cell Diagnostics [ACD]). RNAscope was performed as

described below to verify the presence of neuroanatomical landmarks with probes specific for the VMH (Hs-NR5A1-C1, catalog no. 553151), paraventricular nucleus/supraoptic nucleus ([PVN/SON]; Hs-OXT-C2, catalog no. 538811), ARC (Hs-NPY-C3, catalog no. 416671), and WM tracts (Hs-MBP-C4, catalog no. 411051). Successive rounds of RNAscope were completed until the anatomical plane containing VMH and ARC was reached. This location approximately corresponds to the area between the optic tract (as the ventro-lateral landmark) and the fornix (as the dorsal landmark). Blocks were then scored with a razor blade to isolate the VMH and ARC nuclei in ~6.5 X 6.5 mm squares for the Visium assay. Adjacent 10 µm tissue sections were mounted onto pre-chilled Visium Spatial Gene Expression slides (part number 2000233, 10x Genomics) (**Fig. S1**) and onto Superfrost plus glass slides, which were banked for additional RNAscope validation experiments. In some cases, tissue sections from the same block were collected onto multiple Visium arrays to ensure inclusion of target structures. In these cases, distance along the anterior-posterior axis between experiments did not exceed 100µm. Following Visium analysis, tissue blocks were re-mounted for cryosectioning, approximately 50 µm of tissue was trimmed off to achieve a flat surface, and several 10 µm sections were mounted onto pre-chilled glass slides and banked for validation studies. Blocks were then re-scored with a razor blade to isolate the VMH and ARC nuclei in ~12 X 6.5 mm strips and mounted onto Xenium slides (Part Number 1000460, 10x Genomics). 3-4 tissue strips were fitted onto each Xenium array.

3.3 | Visium data generation

Visium Spatial Gene Expression slides were processed according to the manufacturer's protocol (protocol number CG000239, Rev G, 10X Genomics) as previously described[48,49]. The tissue permeabilization step was performed for 18 minutes based on tissue optimization experiments in other brain regions[49,50]. For each Visium slide, H&E staining was performed and the slide was coverslipped. Tissue sectioning and mounting to slide, sequencing library preparation, sequencing, H&E staining, and microscopy were performed across two sites (listed for each step per sample in **Table S1**). Briefly, microscopy images were acquired either at 40x on a Leica CS2 slide scanner equipped with a 20x/0.75NA objective and a 2x doubler, or on an Olympus VS200 slide scanner equipped with a 10x/0.4NA objective. Following removal of the coverslip, tissue was permeabilized, cDNA synthesized, and sequencing libraries generated. Libraries were loaded at 300 pM on a NovaSeq6000 System (Illumina) at the sites listed in **Table S1**. Sequencing was performed according to manufacturer's instructions, targeting ~60K reads per spot.

3.4 | Xenium Data Generation

Xenium slides were fixed and permeabilized according to 10X Genomics' *Xenium In Situ for Fresh Frozen Tissues – Fixation & Permeabilization Protocol* (CG000581, Rev C). Briefly, slides were warmed for 1 minute prior to submersion in 3.7% formaldehyde for 30 minutes. Slides were then permeabilized with 1% sodium dodecyl sulfate solution (SDS) for 2 minutes and rinsed with PBS. This was followed by a 1-hour incubation in 70% methanol on ice. Following PBS washes, slides were placed in their respective cassettes and incubated with 0.05% PBS-T. For probe hybridization, ligation, and amplification, the 10X Genomics protocol: Xenium In Situ Gene Expression was followed (CG000582, Rev E). The probe hybridization solution was then prepared. We used a combination of the off-the shelf "human brain" probe panel designed by 10X Genomics (Xenium Human Brain Gene Expression Panel, Part No.1000599) and our custom probe panel (10X Genomics; Xenium Custom Gene Expression Panel, Part No. 1000561). Details of the custom probe panel design are described in section 3.8.1. Probe mix was applied to each slide, a cassette lid was applied to each slide cassette and the slides were incubated at 50°C overnight.

The next day, the slides were washed twice in 1X PBS-T followed by an incubated wash in post hybridization wash buffer (Xenium Post Hybridization Wash Buffer, Part No. 2000395, 10X Genomics). Three

washes of 1X PBS-T were completed before addition of ligation mix to each slide. Slides were then incubated at 37°C for 2 hours. After ligation, the slides were washed three times in 1X PBS-T, and then incubated in an amplification master mix at 30°C for 2 hours. After amplification, slides were washed three times in TE buffer followed by three washes in 1X PBS and then incubated in diluted reducing reagent B for 10 minutes at room temperature. Slides were then washed in 70% ethanol followed by two washes of 100% ethanol. Autofluorescence solution was then added to each slide and incubated in the dark for 10 minutes at room temperature. Slides were then rinsed three times with 100% ethanol before drying on a 37°C preheated thermocycler for 5 minutes. Additional washes were performed before adding Xenium Nuclei Staining Buffer, which incubated for 1 minute in the dark at room temperature. Slides were then washed three times in 1X PBS-T and stored at 4°C prior to being loaded on the Xenium analyzer at the Johns Hopkins Single Cell & Transcriptomics Core (SCTC). The solutions needed to run the Xenium analyzer were prepared according to the manufacturer's instructions. Region selection was guided by our team with instrument operation performed by the SCTC Core.

Upon completion of the Xenium analyzer run, post-xenium H&E was completed following 10X Genomics Xenium *In Situ* Gene Expression - Post-Xenium Analyzer H&E Staining protocol (CG000613, Rev A) with the following modifications: reduced staining time in hematoxylin, from 20 minutes to 17 minutes, and reduced eosin staining time, from 5 minutes to 2 minutes. Briefly, slides were quenched in a 10mM sodium hydrosulfite solution for 10 minutes, slides were then washed four times in milli-Q water before being stained with Mayer’s Hematoxylin for 17 minutes (Sodium hydrosulfite, Cat No. 157953-5G, Sigma Aldrich; Hematoxylin Solution, Mayer's, Cat No. MHS16, Sigma Aldrich). After nuclear staining was completed, slides were rinsed three times in milli-Q water before submerging into bluing buffer (Bluing Solution, Cat No. CS702, Dako). Slides were washed once in milli-Q water, once in 70% ethanol, and once in 95% ethanol before being introduced to alcoholic eosin (Eosin Y Solution, Alcoholic, Cat No. 3801615, Leica). After cytoplasmic staining was completed, the slides were rinsed twice in 95% ethanol and twice in 100% ethanol before being cleared in two changes of xylene (Xylene, Reagent Grade, Cat. No. 214736, Millipore Sigma). Slides were then coverslipped using Vectamount mounting media and allowed to dry overnight before imaging (VectaMount Permanent Mounting Medium, Cat No. H-5000, Vector Laboratories Inc.). Imaging was completed at 20X and 40X using a Leica Aperio CS2 digital pathology slide scanner. While H&E images were not utilized in the current study, they have been made available with other supplemental data through a Globus endpoint (see *Data Availability*).

3.5 | Single molecule fluorescence *in situ* hybridization (smFISH)

3.5.1 | Sample preparation and imaging

smFISH assays were performed with RNAscope technology utilizing the RNAscope™ Multiplex Fluorescent Reagent Kit v2 (catalog no. 323100; Advanced Cell Diagnostics [ACD]) in reference to the manufacturer’s protocol (document no. UM323100, rev B; ACD) with the following RNAscope probe sets: Anatomical quality control: Hs-NR5A1 (assigned Opal dye 570), Hs-OXT (assigned Opal dye 620), Hs-NPY (assigned Opal dye 690), Hs-MBP (assigned Opal dye 520) (catalog no. 553151, 538811-C2, 416671-C3, 411051-C4, respectively; ACD); Experiment 1: Hs-NR5A1 (assigned Opal dye 570), Hs-LAMP5 (assigned Opal dye 620), Hs-GAD1-01 *or* Hs-GAD2-ver2 (assigned Opal dye 690), Hs-SLC17A6 (assigned Opal dye 520) (catalog no. 553151, 487691-C2, 573061-C3 or 415691-C3, 415671-C4, respectively; ACD); Experiment 2: Hs-KISS1 (assigned Opal dye 520), Hs-ESR1 (assigned Opal dye 570), Hs-TAC3 (assigned Opal dye 690) (catalog no. 507981, 310301-C2, 507301-C3, respectively; ACD). Specifically, 10-µm tissue sections were fixed in 10% neutral buffered formalin for 30 minutes at room temperature (RT). After 2 PBS washes, the samples were sequentially dehydrated in 50%, 70%, 100%, and 100% ethanol solutions each for 5 minutes at

RT, treated with hydrogen peroxide for 10 minutes at RT, washed 2 times again with PBS, and treated with protease IV for 30 minutes at RT. After another 2 PBS washes, sections were incubated in a solution of the four target probes for 2 hours at 40°C, washed 2 times in wash buffer (catalog no. 310091; ACD), and stored in SSC for up to 48 hours at 4°C. Afterward, the sections underwent amplification with AMP1 through AMP3 for 30, 30, and 15 minutes, respectively, at 40°C with 2 washes with wash buffer performed before each new reagent was added. In sequence, the sections underwent 4 cycles of 15-minute incubation with HRP-Cx, 30-minute incubation with a fluorescent Opal dye solution, and 15-minute incubation with HRP blocker, all at 40°C and with 2 washes with wash buffer in between. Each cycle used a different fluorescent Opal dye (520, 570, 620, or 690; catalog no. FP1487001KT, FP1488001KT, FP1495001KT, and FP1497001KT, respectively; Akoya Biosciences) diluted 1:500 in TSA buffer. Finally, the tissue sections were treated with DAPI for 20 seconds at RT, decanted, and then coverslipped with Fluoromount G mounting medium (catalog no. 0100-01; Southern Biotechnology). Coverslipped slides were stored in the dark for at least 24 hours before imaging. Slides were imaged as z-stacks (at least 10 per image) using a Nikon AXR confocal microscope system powered by the NIS-Elements imaging software with a Nikon APO lambda D 20x / 0.80 objective and/or Nikon APO lambda S 40x / 1.25 objective. For anatomical quality control, slides were imaged in a single plane with a Nikon Plan APO lambda D 2x / 0.1 objective. All confocal images at the resolutions presented/analyzed are available with other supplemental data files through a Globus endpoint (see *Data Availability*).

3.5.2 | Puncta & Colocalization Quantification

For Experiment 2 HALO (Indica Labs) was used to segment and quantify fluorescent signals for each probe in single cells. Nikon *.nd2* files were imported into HALO and viewed at the manufacturer's recommended magnification (40x) for punctate probe signal quantification. The FISH-IF module was used to quantify RNA transcripts (copy counts) within each detected object (i.e. a nucleus with radially defined "membrane" at 3 µm to estimate a cell) in accordance with the manufacturer's guidelines: HALO 3.3 FISH-IF Step-by-Step guide (Indica labs, v2.1.4 July 2021). For each image, fluorescent detection parameters were optimized for DAPI and probe signal to ensure image faithfulness in accordance with the HALO 3.6 User Guide (Indica labs, February 2023). To ensure accurate counting of RNA transcript copies assigned to each cell for each of the 3 RNAScope probe channels, the median probe signal intensity value of positive cells per signal was input into HALO as the representative intensity value parameter for that channel's copy intensity parameter consistent with previously published work[51].

HALO *.csv* outputs were imported into R. As previously described[52], *k*-means clustering was performed on HALO variables Signal Area and Copy Count with a $k = 3$ to cluster "high," "medium," and "low/no" expressing cells) for each of the 3 RNAScope probe channels. High (signal area: *KISS1*–41-85um^2; *ESR1*–17-62um^2; *TAC3*–56-127um^2; and copy count: KISS1–27-82; *ESR1*–6-33; *TAC3*–12-68) and medium (signal area: *KISS1*–13-48um^2; *ESR1*–4-19um^2; *TAC3*–20-61um^2; and copy count: *KISS1*–2-29; *ESR1*–1-12; *TAC3*–3-32) clusters were deemed positive. All cells were plotted as circles with x and y coordinates as determined by HALO and colored by positivity for each channel. The ARC region of interest (ROI) was selected by drawing a polygon around the area expressing *TAC3* and *KISS1* for each donor. A data frame was then produced containing all information for each segmented cell in the ROI, including cell phenotypes (channel(s) meeting the positivity definition above). Phenotype occurrences were then tabulated to quantify cell phenotype counts and proportions. All HALO settings files are provided on GitHub.

3.6 | Visium H&E image segmentation and raw data processing

Multi-sample slide images acquired on the Leica CS2 were first processed using VistoSeg software[53], which divides the whole-slide Visium images into individual capture area images using the *splitSlide* function. The Olympus images were of single capture areas and did not require such processing. Images for each

capture area were aligned in Loupe Browser (10x Genomics) with the gene expression spots captured. Slides were then processed using the Spaceranger v. 2.1.0 software from 10X Genomics, yielding feature counts for each spatial location ("spot") per sample using the *.json* output from Loupe browser, the sample image, and FASTQ sequencing files.

3.7 Visium Data Analysis

All analyses were performed using R 4.4.1 unless otherwise noted. Analysis was performed within the *SpatialExperiment* R framework[54], coercing outside datasets in other formats to this framework where necessary using *zellkonverter*[55]. Spatial expression and clustering plots were generated using *SpatialLIBD*[56] or *escheR*[57]. *ggplot2* was used to generate heatmaps and scatter/volcano plots unless noted. Extended color palettes were generated using the package *Polychrome*[58].

3.7.1 | QC and Normalization

Fraction of mitochondrial reads and number of unique genes identified were calculated for each spot. Plots of these values were manually examined for spatial patterning in each sample to ensure selection of filter thresholds that would not remove putative biological information (for example, WM tracts would be expected to have low gene diversity in a contiguous area, a matter of biological rather than technical variance). Using *SpotSweeper*[59], we first detected and removed high-end outliers relative to surrounding spots in terms of mitochondrial read percentage, low-end outliers for number of UMIs, or low-end outliers for number of unique genes. Subsequently, we filtered remaining spots for high mitochondrial percentage, low UMI count, and low unique gene count. We noted during exploratory analyses that the highest mitochondrial read fractions were largely in spots later identified as VMH—containing the only excitatory neurons of note in the samples—and spots along its lateral edge. We thus set a lenient threshold of 50% for mitochondrial read content in the final analysis. We plotted histograms of UMI and gene count distributions across all remaining spots and identified a clear drop-off in spot frequencies within the histograms at approximately ≤210 UMIs and ≤126 unique genes, suggesting the boundary between technical artifact and biologically low content. Indeed, such spots were almost exclusively localized to sample edges, and these thresholds were therefore used to remove low-content spots. For the retained spots, library sizes (number of reads) per spot were calculated and TMM normalization[60] across all spots in all samples was performed to get log-normalized ('expression') values per gene per spot.

Of 4,992 Visium spots possible in a sample, tissue occupied 4,621±341 (mean ± standard deviation) spots. Spots with high mitochondrial read proportion (≥50%) were removed to preserve a potentially biological pattern of high mitochondrial counts adjacent to VMH (**Fig. S2C**). Subsequent filtering for low unique molecular identifiers (UMIs) and/or gene diversity (≤210 UMIs, ≤126 unique genes) resulted in 4,508±363 spots for downstream analysis. The filtered data included 6,059±5057 UMIs/spot (mean±SD) spanning 2,536±1470 unique genes/spot, covering 45,074 spots and 30361 unique genes altogether. Additional pre- and post-filtering metrics are included in **Table S1**.

3.7.2 | Feature Selection

We used the filtered data to identify highly-variable genes (HVGs, spatially-naive) and spatially variable genes (SVGs) using the packages *scran*[61] and *nnSVG*[62], respectively. For HVGs, we used sample as the blocking factor and collected the top 10%ile and 20%ile of resultant genes using *scran::getTopHVGs(..., prop=0.1 or 0.2)*[61] (**Table S2**).

In the nnSVG approach, each sample is analyzed independently beginning with gene filtering. Here, we filtered to genes with at least 3 counts in at least 0.5% of spots in a sample. We subsequently recalculated the

logcounts for use within nnSVG to account for dropped genes. We extracted and concatenated the nominally significant SVGs from each sample (**Table S2**) and determined the experiment-wide SVGs by averaging their ranks across samples where the gene achieved nominal significance. We then filtered the union of samplewise SVGs to those that achieved nominal significance in three or more samples. These features were sorted in ascending order of mean rank (among samples where the feature achieved nominal significance for spatial variation) to collect the same number of features as contained in 10%ile and 20%ile HVG sets.

3.7.3 | Selecting the Number of Clusters

In a preliminary subset of samples, we used each of the four feature sets described above for performing dimensionality reduction concurrent with batch correction using *Harmony* with sample ID as the grouping variable[63]. The default output of 50 reduced dimensions were appended to the *SpatialExperiment* object for downstream use separately for each feature set.

To determine an expected number of spatial domains, we turned to the Allen Brain Atlas for adult human[64] and tabulated the number of unique anatomic areas denoted along the entire anterior-posterior span of the VMH, yielding 14 potential anatomically-defined domains. We assumed WM would comprise a single, homogeneous domain and added this to get a grand total of 15 potential anatomic areas: dorsal, central, and ventral VMH; ARC; periventricular nucleus, tuberal part; lateral hypothalamic area, tuberal part; posterior hypothalamic nucleus; dorsomedial nucleus; tuberomamillary nucleus; lateral tuberal nuclei; paraventricular nucleus; magnocellular nucleus of the lateral hypothalamic area; supraoptic nucleus; accessory secretory cells of lateral hypothalamus; and generic WM.

In this same preliminary subset of data, we also performed spatially-agnostic, weighted nearest-neighbors clustering using *buildSNNgraph* from *scran*. We utilized *Harmony* batch-corrected principal components calculated on the 10%ile or 20%ile highly variable gene sets, and constructed two SNN graphs for each feature set, using k=5 or k=10 nearest neighbors, for a total of four initial neighbor graph spaces. From each of these four neighbor graphs, we divided the tree into 4, 5…47 clusters using *igraph*::*cut_at* and aggregated these assignments. We determined the optimal, spatially-agnostic clustering resolutions for maximal within-cluster concordance relative to between-cluster concordance using the *fasthplus* method[49,65]. The values returned as *seg_psi_est*, *seg2_psi1_est*, and *seg2_psi2_est* represent the three estimated clustering values most optimal for a given graph. These 12 estimates included two values rounding to 9, three in the range 14.6-15.6, three ranging 17.0-19.8, and a maximum value of 30.8 leading us to select the recurrent minimum (9), the value consistent with anatomic expectations (15), the upper end of this range (20), and the highest value (31) as the number of clusters to generate in the final data.

We then utilized *BayesSpace*[66] with parameters of k=15 and 20,000 iterations for spatially-aware clustering of the *Harmony* reduced dimensions for each of the four variable-feature sets. The resulting domain assignments from each feature set were comparable when plotted onto the spatial data and plots of the top principal component; however, UMAP plots labeled by domain assignment revealed that the majority of domains separated into multiple sectors of UMAP space, except when using the “10 %ile” nnSVG set. We therefore used this feature set and its derived spatial domains for all subsequent analysis.

3.7.4 | Dimensionality Reduction and Clustering

For the analyses of the full dataset presented here, we only utilized an equivalently-derived “10 %ile” nnSVG feature set for dimensionality reduction by principal components analysis (PCA), UMAP, and sample-level batch correction using *Harmony* (**Data S1**) (see *Data Availability*) with initialization of the lambda parameter—which alters the strength of correction—from the data itself, rather than requiring a user-defined initial value (an approach which we term “lambda null” or “lambda NA” throughout the methods and in certain Supplemental files). We used the Harmony batch correction as input dimensions for clustering in *BayesSpace* at resolutions of k=15, 20, and 31. All results presented reflect *Harmony BayesSpace* clusters computed with

k=15 and 60,000 iterations. We produced additional clustering assignments with the nnSVG “10th %ile” feature set using reduced dimensions from MNN batch correction[67] or *Harmony* with its default lambda value of 1 as inputs to 60,000-iteration BayesSpace runs at *k*=15, 20, and 31 (see *Data Availability*). Marker, sex-DE, and GSEA results are provided for clusters defined after using the development version of *Harmony* (“lambda null”-enabled) and BayesSpace *k*=15 (individual clusters), *k*=15 (VMH/ARC clusters joined into single clusters), *k*=20 and *k*=31 (see *Data Availability* and *Supplemental Materials*). Since spots can contain multiple cells, we emphasize that we subsequently call these clusters “domains" to emphasize that they do *not* necessarily correspond to single cell types or subtypes as would be identified in single-cell or -nucleus approaches.

3.7.5 | Marker Gene Identification

Marker genes were comprehensively determined using several complementary statistical approaches. We primarily utilized the *registration_wrapper* function in *spatialLIBD*[56] to calculate *t*-test statistics comparing log counts in each domain vs all others (this method being most analogous to the logFC comparisons assessed using Cohen’s *d* in *scran*). We performed select analyses using the *scoreMarkers* function from *scran*[61] to identify marker genes for select pairs of domains (e.g., to compare VMH subdomains) using multiple statistics: a Cohen’s *d* statistic for mean log fold-change (logFC) between compared groups; an AUC statistic, indexing the probability that a gene in a random spot drawn from domain/group 1 is greater than in a random spot drawn from domain/group 2; and the logFCs calculated when only considering spots from each domain/group where a given gene was detected. We focus on *SpatialLIBD t*-test statistics in the main text for global marker identification, and on *scran* output statistics for comparisons between VMH or ARC subdomains. One-to-one comparisons were performed for all combinations of putative VMH clusters or putative ARC clusters identified in clusterings at *k*=15, 20, and 31, the results of which are provided in the Supplement. Putative VMH/ARC subdomains from the additional were identified based on marker genes and referred to throughout the supplement as (*VMHorARC)(k).(digit)*, where each subdomain is assigned a sequential number arbitrarily (e.g., “VMH20.1” and “VMH20.2” are two subdomains of VMH identified from *k*=20 BayesSpace analysis). To discern genes differentiating among VMH or ARC clusters from Visium, we additionally performed ‘one-vs-all-other’ analyses, wherein each cluster from a domain was separately compared to other clusters *only* from that domain by subsetting the data only to spots with VMH or ARC cluster labels before performing marker analysis using *spatialLIBD* as above. For *k*=15 or 20, this amounted to a single pairwise comparison for each domain, as VMH and ARC each spanned two clusters at both of these clustering resolutions.

3.7.6 | Spatially Variable Gene Identification within VMH and ARC

To further explore whether we could identify spatial patterns of gene expression within the VMH and ARC specifically, we subset the data to those spots assigned a VMH or ARC domain as above and re-ran nnSVG on individual samples, only considering spots within ARC or within VMH. One sample was excluded from this analysis for VMH due to an impractically small number of VMH spots (sample v5459A_M, 330 spots) for SVG analysis. One sample was likewise excluded from ARC analysis for the same reason (sample v8667A_F, 207 spots). The threshold for gene inclusion for SVG analysis in a given sample was ≥1 count in at least 5% of spots within the domain being analyzed. The ranks of nominally significant genes per sample were tabulated and aggregated to their means across those samples where the gene achieved nominal significance, as was done for the feature selection steps described above. As the collective structure of VMH and ARC domains was similar at all clustering resolutions, we only performed these SVG analyses on the domains as defined by the *k*=15 BayesSpace results.

3.7.7 | Pseudobulk Sex-Differential Expression Analyses

We set out to test each Visium cluster/domain for sex differences in expression. To achieve this, spots with a given cluster/domain level were pseudobulk aggregated within-sample using *dreamlet* 1.3.1. We then subsetted to clusters found in at least 5 samples (at least 2 per sex). Pseudobulk counts were preliminarily normalized to total counts using TMM normalization. We then programmatically set a low-expression filter by detecting the smaller sex group (in terms of sample number) among the retained samples for a given cluster, detecting the second minimum in a histogram of gene-wise mean log counts. This value was used to perform low-expression filtering such that retained genes which adhered to an approximately normal distribution. TMM factors were then recalculated for the remaining genes, followed by DE analysis using *voomLmFit*, with donor as a blocking factor (which is internally passed to *DuplicateCorrelation*), *adaptiveSpan* set to true to identify the number genes to use for smoothing the mean-variance trend, *sample.weights* set to true to perform an analysis analogous to the *voomLmFit* predecessor *voomWithQualityWeights*, and fitting a model of *~0+Sex* to test for sex effects. This was run for all batch correction and clustering resolutions described above, with results available in the Supplement.

3.7.8 | Gene Set Enrichment Testing of Markers and Sex-DEGs in GSEA-MSigDB and for Transcription Factor Target Genes

We performed comprehensive, threshold-free analyses of genes marking spatial domains to confirm expected functional roles and identify potentially unexpected ones. Using the *t*-statistics of all genes tested in marker analysis from *spatialLIBD*[56], we applied Gene Set Enrichment Analysis[68,69] as implemented in the package *fGSEA*[70] to analyze marker enrichment in gene sets from MSigDB[68,71,72]. We included “hallmark” sets, along with sets representing chromosomal regions; curated sets from publications, including pathway databases like KEGG; encompassing predicted microRNA (miRNA) / transcription factor (TF) targets; defined in ontologies; and reported human cell type markers. (These correspond to gene sets “h.all”, “c1.all”, “c2.all”, “c3.all”, “c5.all”, and “c8.all”, respectively; all versioned “2022.1”.) We also included the MSigDB “oncogenic signature”/”c6.all” dataset as a null comparator and to screen for transcriptomic changes suggestive of HPA tumors, which are relatively common and whose presence would run counter to our present goal of profiling healthy neurotypical hypothalamus.

We additionally aggregated several databases of tissue-*non*specific putative target genes of transcription factors (TFs) ascertained from model organism and human cell/tissue types. We collated TF-target sets from ChEA 2022[73], TRRUST[74], those mined from publication supplements by Rummagene[75], and those curated in the Enrichr[76] tool (Gene Expression Omnibus-mined results of TF perturbation and knockout experiments and TF protein-protein interactions). We subsequently filtered the results for each domain to those TFs expressed in that domain. Given that TFs are often lowly expressed at the mRNA level[77], we used lenient filters to define TFs as expressed: for marker gene enrichment, we required the at least 1% of spots to have log counts >0 for the TF in ≥5 (majority) samples and a mean [% of sample’s domain spots with nonzero counts of the TF] >5%. For sex-DE enrichment testing, we considered a TF expressed if ≥3 (smallest group size, when considering limited ARC or VMH spots in a Visium sample from each sex—see *Methods 3.7.6*) samples had at least 0.25% of spots in a domain with log counts >0 and a dataset mean of at least 1% of spots in the cluster with log counts >0. Our rationale for this difference in parameters was that our marker-based statistics were for broad domains, such that regulatory signatures defining the domain should be relatively abundant, whereas for sex differences, the expectation is not necessarily that all (below-detectable resolution) subdomains or cell types/subtypes in the domain—*i.e.*, a smaller number of spots than define the domain itself—drive the detected DE events.

Gene sets analyzed were limited to those containing 15-500 genes for MSigDb and TF-target analyses. We utilized signed, transcriptome-wide one-vs-all marker or male-vs-female DE *t*-statistics as our ranked list for the respective analyses. To retrieve enrichment statistics from every gene set for both male-upregulated

and female-upregulated expression patterns, we explicitly ran analyses for the top (male-upregulated / positive *t* statistic) and bottom (female-upregulated / negative *t* statistic) of the rank lists, as fGSEA testing on both directions simultaneously will otherwise return data for the effect direction with greater enrichment, even if both directions are significant. Additional sex-DE GSEA analyses were performed using the absolute value of DE *t*-statistics to identify gene sets enriched for sex-differential expression as a general phenomenon. To ensure the ability to resolve extremely small *p*-values, we set the fGSEA parameters *eps* to 0.0 to specify that p-values should be estimated for all enrichments, using 50,000 permutations for initial *p*-value estimations. To determine the most precise, independent enrichments for some of these analyses, we subsequently used the *collapsePathways* function in fGSEA. For marker gene analysis, we only examined "positive" enrichment, i.e., enrichments corresponding to genes more specific to a given spatial domain relative to others. For sex-DEGs, we used the same parameters and gene sets as above, but examined both directions of enrichment for each set of spatial domain's DEGs (as the test statistic's sign indicates male or female upregulation).

3.7.9 | GWAS-Derived Gene Set Enrichment Testing

To comprehensively examine whether neuropsychiatric disorder risk genes were enriched among markers or sex-DE genes, we collected genes reported in common-variant (genome-wide association study (GWAS)-based) and rare-variant studies across several modalities. GWAS-derived gene sets comprised those reported in GWAS studies, explicitly *excluding* genes assigned to loci on the sole basis of being the most proximal to a phenotype-associated variant, while retaining e.g. genes reported in the GWAS publications resulting from "post-GWAS" analyses such as MAGMA, quantitative trait locus (QTL) colocalization, etc. For schizophrenia and autism spectrum disorders, we also collected genes identified in, e.g., whole exome sequencing cohorts of patients and healthy sibling probands. We additionally collected transcriptome-wide association study (TWAS) analyses performed on GWAS summary statistics for the same traits (generally from the same GWAS studies); largely collected from webTWAS[78,79,80,81] and TWASatlas[82], as well as individual publications that performed TWAS. Finally, we gathered literature-mined, curated gene-disease pairs from DisGeNET[83]. For major depressive disorder, these included several genes of historical interest that have been explicitly demonstrated *not* to play a role in heritable depression risk[84]; we thus removed the genes covered by[84] from this set specifically. The sets from GWAS, TWAS, and DisGeNET and corresponding sources are provided in *Supplemental Materials*. For the analyses presented, we analyzed enrichments solely in the union of genes collected for each diagnosis across these three source types. As some of these gene sets fell outside the above bounds of 15-500 genes, the largest gene set size was elevated to 8200 to exceed the largest disease-gene list (BMI, union of all gene-associated modalities, >8100 genes).

3.7.10 | Spatial Registration

Given the paucity of publicly available, adult human hypothalamic single-nucleus RNA-sequencing, we opted to map our spatial domains to Allen Brain Atlas *in situ* hybridization (ISH) data from mouse[85] using *scCoco/cocoframer*. We provided the top 150 markers per Visium spatial domain based on one-vs-all logFC and set a minimum of 50 genes with adult mouse ISH data for a given domain's comparison to proceed. *scCoco* then was used to score the expression of query markers within each Allen Atlas Common Coordinate Framework (CCF) area of the hypothalamus. Scoring proceeds for one query marker set at a time (i.e., 50-150 genes) by ranking the ISH expression of each query gene in 10 voxels from every CCF hypothalamic area, then converting the median rank of each marker's expression to a value ranging 0-1 by taking 1 - (median rank / number of voxels queried). Finally, these converted values are averaged across all genes for each CCF region to get an enrichment score ranging from 0 (least overlap) to 1 (most overlap). Results using development *Harmony* with *BayesSpace* at *k*=15, 20, and 31 are provided in the Supplement.

We additionally utilized a recent mouse brain atlasing effort that included single-cell RNA-sequencing results from 45 sequencing libraries collected across over 25 donor mouse hypothalami from both sexes[86]. The

registration_wrapper command in spatialLIBD to identify marker genes of each cell “subclass” (a relatively coarse level of hierarchical clusters reported in the Yao data) detected in We then tested the correlation of marker *t* statistics from our spatial domains (k=15, 20, 31, and k=15 with collapsed VMH and ARC domains) with each mouse subclass’ *t*-scores for the same genes. We tested correlations using the top 25, 50, 75, 100, 250, 500, 750, and 1000 spatial domain marker genes (ranked by *t*-statistic), as well as using all genes with one-to-one mouse orthologs. We additionally repeated the analyses with all mouse VMH subclasses re-labeled as “VMH” and all ARC subclasses relabeled “ARC” to test domain correlation to collective VMH/ARC mouse types. We limited these analyses to a total of 47 subclasses—those represented by at least 15 cells per sample in 22 or more of the samples.

3.7.11 | Sex-DE comparison to adult mouse hypothalamus

To test whether the sex-DE patterns we identified in the human Visium data are similar to those in adult mouse hypothalamic gene expression, we returned to the mouse brain atlasing data used for spatial registration[86]. We first filtered the brain-wide-definition subclasses to those represented by ≥ 15 cells in ≥ 20 samples in the mouse hypothalamic single-cell datasets, then performed sex-differential expression analysis for each using *dreamlet*. To control for duplicate samples from the same donor mice, our differential expression model was ~donor_sex + (1|donor_id), using random effects to control for repeated donor sampling in the mouse study. These results and human gene orthologs were then retrieved using *orthogene*[87]. For comparisons to Visium, we performed sex DE analysis on each of 81 mouse cell subclasses, with the two mouse VMH subclasses re-labeled as ‘VMH’ and the four mouse ARC subclasses re-labeled as ‘ARC’. We then calculated the Spearman correlation coefficients for signed male-vs-female *t*-statistics for autosomal, non-mitochondrial genes with 1-1 orthologs between mouse and human, taking the 25, 50, 75, 100, 250, 500, 750, 1000, or all top-ranked human sex-DEGs (by |*t*-statistic|) with orthologs analyzed in both the human domain and mouse subclass/aggregate subclass. A similar analysis was performed between Xenium VMH and ARC neuron clusters and individual mouse VMH and ARC subclasses.

3.8 | Xenium Data Analysis

3.8.1 | Gene Selection for Xenium Custom Panel

Genes were selected for measurement on the Xenium platform in an initial subset of 7 Visium samples, all from donors in the final dataset. (One sample from the preliminary dataset was not included in the final data due to low read depth compared to all other samples; a second, adjacent tissue section sampled from the same donor was processed using Visium to collect higher-depth sequencing data and *is* included in the final dataset). This data subset was filtered using traditional QC approaches, removing spots with over 35% mitochondrial reads, under 600 UMIs, and/or under 450 unique genes detected. This data was integrated across samples with Harmony, clustered using BayesSpace with *k*=15, and analyzed for marker genes, VMH/ARC-local SVGs, and sex-DE, as described above, with the difference that *dreamlet* 1.0.0 was used for pseudobulk sex-DE analysis and a random effect of donor. Top marker genes of one or both VMH or ARC domains were visually inspected in sample-wise fashion for consistent patterns of specific or enhanced expression in that domain across samples. For SVGs, *nnSVG* results were aggregated across samples by taking the mean of ordinal gene ranks in samples nominally significant ($p<0.05$) for spatial variability in at least 3 samples. The top 100 mean-ranked genes each for ARC and VMH were visualized in all samples to determine whether there was a consistent spatial pattern of expression within the ARC or VMH across donors, and visually robust genes were selected for inclusion. For sex-DE genes, the preliminary analysis with *dreamlet* 1.0.0 returned a much higher (compared to the final dataset) number of significant DEGs at

FDR<0.05 due to algorithmic differences at that time, with approximately 150 in VMH and 400 in ARC. All genes meeting this significance threshold were plotted and visually examined in spotplots to determine whether the male/female pattern of expression difference was consistent. Further, visual inspection was used to rule out DE events that occurred due to genes predominantly detected at the edges and outside of VMH/ARC, which would indicate confounding due to differences in VMH/ARC boundaries relative to the neighboring domains primarily expressing the gene. Additional genes were selected on the basis of genes identified in model organism literature as marking VMH or ARC or their general relevance to sex differences in gene expression (e.g., androgen and estrogen receptors). We provided 10x Genomics with expression reference data for the panel design (balancing probing depth for each gene based on its expression) in the form of 1 male (v6197A_M) and 1 female (v6588B_F) Visium sample, with *k*=15 BayesSpace cluster labels from our preliminary analysis. A table of the final panel of genes and the rationale(s) for inclusion in the study is provided in the *Supplement*.

3.8.2 | Data Acquisition

Data was obtained from two slides (see *Methods*, section 3.4) each in two separate runs. The first two slides of Run 1 were acquired and post-processed by the Johns Hopkins SCTC using instrument firmware/xeniumranger version 1.5. The second two slides of Run 2 assayed at SCTC were acquired and post-processed with instrument firmware/xeniumranger 1.9. Due to changes in cell segmentation algorithms between these two Xenium versions, we uniformly re-postprocessed all data by using *xenium-resegment --with-nuclei=TRUE* in xeniumranger 1.7.

3.8.3 | Data QC and Preprocessing

Xenium data was initially assembled into an R *SpatialFeatureExperiment* (SFE)[88] object by importing xeniumranger outputs: cell boundaries, nucleus boundaries, and per-cell count data for transcript detections above quality score (QV) > 20. Cells and nuclei with segmentations containing holes as detected by the *nngeo* package were removed. Per-cell proportions of each type of negative control elements (negative control probes, negative control codewords, and unassigned/formerly BLANK codewords) were tabulated for each cell and added to the SFE. As the Xenium documentation does not address ‘Deprecated’ codewords or what they correspond to, this category of negative element was not considered further. Samples were then rotated or mirrored in space so that, for visualization, all tissue sections were oriented with the ARC in the bottom left-hand corner.

We then performed general quality control of the data as outlined in the documentation for the *Voyager* R package. Briefly, we first screened the data for nuclei that had either zero area or had an area ≥ 95% of their parenting cell’s segmentation (none were found). We then visualized the proportion of counts coming from each type of negative control element or of all elements considered jointly, confirming there were not marked spatial differences in negative element detection. We also examined the distribution of these proportions, finding that the vast majority of cells had 0 negative element reads and a small number of cells with ≥ 10% such reads. Following the *Voyager* approach, we used *scuttle::detectOutliers* to remove cells for which the proportion of reads contributed by any negative element (or their union) was > 3 median absolute deviations above the sample’s median. Cells deemed outliers in this manner also did not show any clear spatial pattern, though they were more visually dense in more cell-dense areas. Finally, we removed nuclei and corresponding cells wherein the nuclear area was under 400 pixels^2, resulting in an initial dataset of 870,870 cells/nuclei.

Cell counts were normalized using both TMM normalization into log counts based on the number of total counts in a cell, and normalized without log transformation using *scuttle::normalizeCounts(log=FALSE)*. PCA and UMAP dimensionality reduction was performed to visually inspect for batch effects of donor and Xenium run; remarkably, the non-log-normalized data (the data used for clustering with Banksy, as advised by

its authors[89]) occupied a UMAP space shared across these potential batch covariates (**Fig. S3**), and thus did not require batch integration for further analysis.

3.8.4 | Xenium Clustering with *Banksy*

We subsequently set out to cluster the segmented Xenium cells using the *Banksy* package in R[89]. Per their documentation for multi-sample datasets, we placed an artificial buffer zone between the spatial coordinates of each tissue section, such that each section had unique x-coordinates and was "spaced" at least 1.5 tissue section widths apart from every other section. We then performed *non*-spatial clustering of the Xenium data, using non-log-normalized counts as input to Banksy for dimensionality reduction with parameters *kgeom*=6 (setting the physical 6 nearest cells to be defined as a cell's neighborhood) and the azimuthal Gabor filter setting and *lambda* both set as 0 for spatially-unaware dimensionality reduction, followed by Leiden clustering with a resolution parameter of 2. We elected to forego spatially-aware clustering in the Xenium data in order to examine *whether* genuinely distinct spatial clusters were present without biasing our analyses toward producing spatially arranged clusters *a priori*.

3.8.5 | Cell Type Annotation and VMH/ARC Domain Smoothing

We subsequently removed 8 sample/donor-specific clusters (one of which was a VMH cluster) from the SFE object to perform pseudobulk one-vs-all marker detection for each of the 33 clusters found across the dataset using *spatialLIBD* as described above for Visium. Based on marker genes, we were able to assign each cluster a putative identity, including 3 main VMH clusters and 5 ARC clusters. One ARC cluster, marked by *GAL* expression, was present both in the ARC and along the entire medial longitudinal axis of the VMH; as this cluster was not restricted to the traditional atlas boundaries of ARC (ventral to VMH and extending only along the ventromedial portion of VMH), we did not include this cluster in the smoothing procedure to define the ARC domain. For smoothing of the VMH, we excluded a fourth, rare cluster (7.3k cells total) tracing the approximate edges of VMH. We note that this cluster is annotated as "VMH boundary" in recognition of the spatial pattern, but was not marked by any canonical VMH genes (*SLC17A6*, *FEZF1*, *NR5A1*).

In order to define the spatial boundaries of VMH and ARC, we then performed a two-step spatial *k*-nearest neighbors (kNN) approach. In brief, this procedure first quantified the proportion of each cell's 10 (VMH) or 50 (ARC) spatially nearest neighbor cells (*including* those in sample-specific clusters) that were one of the three main VMH clusters or 4 main ARC clusters. A second round of kNN, obtaining the proportion of the 200 (VMH) or 500 (ARC) spatial neighbors labeled as VMH or ARC was performed, with intermediate VMH and ARC labels assigned based on thresholding the first kNN value for each region at 0.1, 0.2…1.0 (100 VMH/ARC combinations). We collected kNN values for the second step using each first-round threshold pair, and subsequently tested final definitions of ARC and VMH with thresholds of 0.1, 0.2…1.0 as applied to the second-round kNN smoothing (total of 10,000 combinations of VMH-ARC threshold pairs).

For each of these pairs, pseudobulk DE testing was done for marker genes between all cells passing the given final VMH or ARC threshold to quantify VMH/ARC marker gene enrichment in the respectively assigned domains. DE analysis was performed using *registration_wrapper* in *spatialLIBD* as we did for other gene marker analyses. First-second round combination pairs with robust enrichment of VMH markers in VMH and of ARC markers in ARC were visualized to determine whether the parameters also resulted in contiguous and directly adjacent VMH and ARC domains. For smoothing parameter combinations where VMH and ARC boundaries were adjacent, we further performed pseudobulk DE analysis to determine whether expected VMH and ARC markers were enriched on their respective sides of the boundary zone. To achieve this, we used the *sf* package to define the convex and concave hulls of smoothed VMH/ARC domains, then extended these boundaries by 5% of the median distance between hull vertices using *st_buffer* in *sf*. We performed a second round of VMH/ARC pseudobulk DE testing only comparing intra-VMH and intra-ARC cells within the overlap region.

Based on the magnitudes of enrichment for known marker genes, domain continuity, and domain adjacency, we selected first-round KNN thresholds of 0.2 (VMH) and 0.1 (ARC) followed by second-round KNN thresholds of 0.2 (VMH) and 0.5 (ARC). Cells that passed both ARC and VMH thresholds in the first or second step were deconflicted by assigning the higher KNN score (e.g., if the KNN round 1 proportion of ARC neighbors was 0.25 and of VMH neighbors was 0.3, then the cell would be labeled VMH going into the second round of KNN analysis). The labels of VMH or ARC were then assigned to all cells, regardless of cluster, to demarcate the boundaries of the VMH and ARC. Refined labels were determined by performing a one-vs-others pseudobulk enrichment approach similar to that described for Visium VMH/ARC clusters in section 3.7.5: that is, we tested pseudobulk expression on data subsetted only to ARC clusters or only to VMH clusters. This again was performed using tools from *spatialLIBD*.

Domain labels were then cleaned by removing cells outside of the main domain area that received the label; in the case of one donor, Br8741, both Xenium samples were instead processed with a second round kNN threshold of 0.15 for VMH due to holes in the domain assignments at higher thresholds. This donor contained the donor-specific VMH cluster, which was excluded from the smoothing procedure, resulting in fewer remaining VMH neurons in the sample and consequently in lower KNN proportions for VMH neuron type-neighbors.

For gene enrichment at the whole-domain level, cells from any of the 33 retained clusters falling within the boundaries of the VMH or ARC domain were given that domains label as their ‘cluster’, while any cells from a VMH or ARC cluster but found outside the two domains were removed from analysis. One-vs-all analysis was then performed with *spatialLIBD* using these modified labels such that VMH and ARC domains were treated as single clusters alongside the unmodified (non-VMH, non-ARC neuron) clusters; enrichment statistics for the VMH and ARC domains were then retrieved.

3.8.6 | Sex-Differential Expression Analyses

We set out to test each cell cluster for sex differences in expression local to the VMH or ARC. To achieve this, we first extracted all cells within either the VMH domain or the ARC domain as defined above, then performed pseudobulk aggregation within samples for each cell cluster using *dreamlet* 1.3.1. We then subsetted to clusters for which there were at least 6 samples with at least 10,000 total transcript detections for that cluster within the domain of interest. Pseudobulk counts were preliminarily normalized to total counts using TMM normalization. We then programmatically set a low-expression filter by detecting the smaller sex group (in terms of sample number) among the retained samples for a given cluster in each domain by detecting the second minimum in a histogram of gene-wise mean log counts to retain only those genes which adhered to an approximately normal distribution. Genes falling below this second minimum were discarded and TMM factors recalculated for the remaining genes. DE analysis was then performed with voomLmFit, using donor as a blocking factor (which is internally passed to *DuplicateCorrelation*), *adaptiveSpan* set to true to identify the number genes to use for smoothing the mean-variance trend, *sample.weights* set to true to perform an analysis analogous to the *voomLmFit* predecessor *voomWithQualityWeights*, and fitting a model of *~0+Sex+Xenium Run* to test for sex effects and control for potential variation between the two Xenium runs. The filtered pseudobulk object and full DE analysis output for each cluster within each domain were returned for visualization and downstream analyses.

In order to validate our sex-DE findings within VMH and ARC domains, we performed two additional high-confidence analyses. In the first, we utilized the *sf* package in R to define the convex and concave hulls for VMH and ARC, respectively, and trimmed these boundaries back by 5% of their respective domain areas to obtain “high-confidence” domains. Cells falling within these interior domain boundaries were then analyzed as above for sex-DE. In a second analysis, we sought to rule out cell segmentation effects that could have resulted in transcripts being incorrectly attributed to a given cell type (e.g., astrocyte endfoot processes could potentially end up within neuronal segmentation boundaries). We therefore also performed sex-DE only considering transcripts overlapping nucleus segmentations by assigning a given nucleus the same cluster label

as its parent cell segmentation. The nuclear count data was filtered to quality scores > 20 and totaled per gene per nucleus using the *MoleculeExperiment* package in R[90], followed by the same pseudobulk per-cluster-per-domain procedure as described above.

3.8.7 | Comparison of Xenium Clusters to Visium and to Mouse Hypothalamus

To assess the comparability of our cell type cluster labels to mouse hypothalamus, we performed *scCoco* analysis and spatial registration with *spatialLIBD* as described for Visium.

Results

Identification of human VMH and ARC using complementary spatially-resolved transcriptomic approaches

To better understand sex-differentiated gene expression at cellular and spatial resolution in the human VMH and ARC, we performed Visium spatial gene expression profiling followed by Xenium *in situ* sequencing (10x Genomics) on adjacent sections of postmortem human HYP tissue from male and female adult neurotypical donors (*N*=4 per sex, **Fig. 1A-C**). We used Visium to measure transcriptome-wide gene expression across the VMH and ARC, facilitating data-driven selection of genes for expression profiling at cellular resolution using Xenium (**Fig. 1C**). To locate the VMH and ARC, we performed single-molecule fluorescence *in situ* hybridization (smFISH) for *NR5A1* and *NPY,* marker genes for VMH and ARC, respectively. When the VMH and ARC were reached, tissue sections were first collected for Visium (10 total samples, 1-2 sections per donor) (**Table S1**). After quality control (**Fig. S2**, **Fig. S4**), the filtered Visium dataset included ~45K spots and ~30k genes.

To identify data-driven spatial domains representing VMH and ARC, we performed unsupervised spatial clustering. Specifically, we used 1,816 spatially variable genes (SVGs, **Table S2**)[62] as features for dimensionality reduction and batch correction[63] (**Data S1**), followed by unsupervised clustering at *k*=15, 20 or 31 (*Methods*). We present results using *k*=15 (**Fig. 1D**) as this value is consistent with the number of anatomical regions annotated in the Allen Brain Atlas in our samples. At *k*=15, we discerned 2 oval-shaped VMH clusters, which uniquely expressed the glutamate transporter *SLC17A6* and signature VMH marker genes *NR5A1* and *FEZF1* (**Fig. 1E**-**F**; **Fig. S1**, **Fig. S5**). Ventral to these clusters, we observed 2 ARC clusters enriched in *KISS1*—a regulator of gonadotropin releasing hormone (GnRH)—as well as neuropeptide-encoding genes *GHRH*, *AGRP*, *NPY*, and *POMC* (**Fig. 1G**, **Fig. S5**, **Table S3A-C**). We henceforth refer to these Visium (v) clusters as vVMH.1, vVMH.2, vARC.1 and vARC.2, and when considered together as 'vVMH' and 'vARC'.

We next performed gene set enrichment analysis (GSEA) for vVMH and vARC using MSigDB-indexed gene sets (*Methods*) (**Data S2**, **Fig. S6**). As expected, vARC was enriched for inhibitory neuron terms, vVMH for excitatory neuron terms, and both for terms related to neuropeptides. Consistent with previous literature[91], vVMH-enriched genes were overrepresented for kisspeptin signaling pathways. Further, "regulation of insulin secretion" genes were enriched among highly expressed genes in vVMH, consistent with its glucoregulatory function in rodents[92,93,94]. Interestingly, vARC-enriched genes were overrepresented for terms relating to primary cilia and their organizing body, the axoneme. We confirmed this by performing GSEA with a curated set of human primary cilia genes[95], and found ARC-specific enrichment for primary-cilium-localized genes (FDR<5×10^{-7}; **Table S4I**). Together, this provides evidence that we successfully identified ARC and VMH in our Visium samples.

Following an analysis of preliminary Visium data, a unique panel of 100 genes was designed for the Xenium platform (**Table S5**), which included vVMH and vARC marker genes, spatially variable genes (SVGs), sex-differential genes, and genes with spatial or sex patterns from rodent HYP literature (*Methods*). These 100 genes were supplemented by a commercially available "human brain" panel of 266 genes spanning markers for broad cell types largely derived from human cortex data. From all 8 donors, tissue sections adjacent to those profiled with Visium were collected for Xenium (*N*=7 male and *N*=6 female sections). After quality control (**Fig. S7**), data were clustered (**Fig. 1H**, **Fig. S3**), yielding 4 Xenium VMH and 5 Xenium ARC clusters, which we refer to as xVMH and xARC clusters. We then developed a spatial smoothing algorithm (*Methods*) to obtain anatomically contiguous domains corresponding to the xARC and xVMH (**Fig. 1I**). Consistent with Visium, Xenium domains for xVMH and xARC were respectively enriched for expected marker genes *NR5A1* (**Fig. 1J**) and *POMC* (**Fig. 1K**), with qualitatively strong anatomical agreement within donors and across platforms (**Fig. S8**).

Molecular characterization of VMH and ARC cell types and subdivisions

Because molecularly distinct subdomains have been previously described in rodent VMH[3,12], we investigated whether this was apparent in vVMH and vARC. At *k*=15, we were only able to identify 2 vVMH and 2 vARC clusters. For VMH, vVMH.1 and vVMH.2 shared top-ranked genes spanning several canonical VMH marker genes[96], such as *NR5A1*, *FEZF1*, *ADCYAP1*, and *KISS1R* (**Fig. S5**). Genes differentiating rodent ventromedial (vm) and vlVMH, such as *Foxp1* and *Esr1*, were not enriched in either vVMH cluster (**Table S3B**). For ARC, vARC.1 was dorsal to vARC.2 and shared canonical ARC marker genes, such as *POMC* (**Fig. 1G**) and *AGRP* (**Fig. S5**). Pairwise differential expression between vVMH clusters demonstrated that vVMH.1 was more enriched for canonical marker genes, while vVMH.2 was more enriched for glial markers (*e.g.*, *GFAP*, *MBP*; logFCs < 0.75; **Fig. S5**, **Table S6**). This was also the case for vARC, suggesting that vVMH and vARC clusters reflect local differences in glial content. In neither vVMH nor vARC did higher *k* values yield further molecularly distinct spatial clusters inside ARC or VMH domains (**Fig. S5**, **Fig. S9**, **Fig. S10**).

While discrete clustering approaches were not able to identify anatomical subdivisions for human vVMH or vARC, we next evaluated whether spatially-variable gene (SVG) expression might suggest potential cellular substructure in these domains. Considering each domain in isolation, we found 277 SVGs in vVMH and 1279 SVGs in vARC (**Table S7**). Top VMH SVGs of interest included ventral bias in somatostatin (*SST)* expression (**Fig. S11**) and medial/ventromedial bias of the progesterone receptor (*PGR*) (**Fig. S12**). Top ARC SVGs showed more marked patterns (**Fig. S13**-**Fig. S19**) and included *TAC3* and *GHRH* (**Fig. 2A**). Similarly, a top-200 SVG in all vARC samples was *DIRAS3* (absent laterally, **Fig. S19**), a gene expressed in developing *NPY*-expressing neurons of the human ARC[97], suggesting that some SVGs might represent distinct cell types. To further explore whether ARC SVGs might be indicative of different spatially-localized cell types, we examined 11 gene markers of 6 ARC clusters previously reported in developing human ARC, including *TAC3, GHRH,* and *DIRAS3*[97]. All 11 gene markers were significant SVGs in ≥6 tissue sections, with mean ranks skewed toward the top (Wilcoxon test, $p<1x10^{-7}$), and 8/11 were in the top 150 vARC SVGs overall. We thus concluded that there were likely distinct spatial distributions of ARC neuron types not resolvable as distinct clusters using Visium.

We therefore sought to evaluate VMH and ARC at cellular resolution using Xenium. Consistent with Visium results (**Fig. 2A)**, mutually exclusive *GHRH* and *TAC3* expression was reproducible within xARC (**Fig. 2B**). As introduced above, we identified and annotated 33 cell type clusters, including 5 xARC clusters and 3 xVMH clusters identifiable by marker genes and spatial distributions (**Fig. S3**, **Fig. S20A-B**, **Table S3G**). For xVMH, all three clusters shared enrichment for a number of VMH marker genes (*SLC17A6*, *NR5A1*, *FEZF1*, *NRGN*, *ANKRD34B*, *ADCYAP1*, *HS3ST4*; **Fig. S20B**). We therefore performed marker analysis restricted to these 3 clusters. xVMH.1 was highly enriched for *CRHR2*, while xVMH.3 was enriched for *LAMP5*. Interestingly, xVMH.2 was comparatively enriched for the oxytocin receptor (*OXTR),* but also for astrocyte (*GJA1*, *SOX9*) and microglia (*P2RY13*) markers (**Fig. S20C**, **Table S3I**); as such, we refer to xVMH.2 as "VMH-glia mixed." In addition to these 3 xVMH clusters, we noticed an *HCRT*-expressing cluster restricted to the VMH perimeter (**Fig. 2C**). Despite being neither exclusive to VMH nor enriched for canonical VMH markers, we included this "lateral border" population as a fourth xVMH cluster in downstream analyses based on its surprising spatial pattern. In contrast to xARC and our expectations of xVMH based on the literature in rodents, xVMH lacked clear spatial organization among its three main clusters (**Fig. S21**). For xARC (**Fig. 2D**, **Fig. S22**), we readily identified 3 out of 5 clusters based on known marker genes from the literature: xARC.1 as *AGRP*-expressing neurons, xARC.3 as a *TRH*-expressing subpopulation of *GHRH* neurons, and xARC.4 as *TAC3*- and *ESR1*-expressing neurons possibly representing *KISS1*-expressing KNDy neurons (**Fig. S20**). Although *KISS1* was not included in our Xenium panel, orthologues to secondary markers of mouse KNDy neurons[96,98] were also enriched in our *ESR1*-*TAC3* population (*ECEL1* and *NR4A2*, **Fig. S20**, **Table S3G**). smFISH further showed that ~14% of *ESR1*+/*TAC3*+ neurons in ARC also express *KISS1*, suggesting that KNDy neurons are included within xARC.4 (**Fig. 2E-F**, **Fig. S23**). For the two remaining clusters, xARC.5 was

comparatively enriched in oligodendrocyte markers (*OLIG1*, *OLIG2*, *PDGFRA*) along with *POMC*, and thus designated "*POMC*-Oligo Mixed," while xARC.2 was enriched for *SLC17A7* and glial markers and designated "*SLC17A7*-Glia Mixed" (**Fig. S20C**, **Table S3H**). Regarding relative enrichments of glial genes when comparing among xARC and xVMH clusters, we repeated our analysis using only nuclear-expressed transcripts and obtained similar results (**Table S3J-K**), supporting close interactions between glial processes and neurons in these clusters. In conclusion, the cellular resolution afforded by Xenium was able to resolve ARC, but not VMH, into spatially organized cell clusters that were not distinguishable using Visium.

Shared transcriptomic identities and divergent features of VMH and ARC across mouse and human

To examine how well vVMH and vARC domain marker genes corresponded to those of the adult mouse VMH and ARC, we compared our human data to a recently published single cell atlas of the adult mouse brain[86]. vVMH, vARC, and optic tract (OT) clusters shared gene marker signatures with mouse VMH, ARC, and white matter cell types, respectively (**Fig. 3A**, **Table S8E**). We obtained similar results when comparing marker gene expression against adult mouse *in situ* hybridization data from the Allen Brain Atlas[85] (**Table S8A-D**). In conjunction with our functional enrichment analyses (**Fig. S6**), these findings are consistent with signals, circuits, and cellular biology characterized in rodent ARC and VMH.

Given shared transcriptomic identities between human and mouse VMH and ARC, we next aimed to identify divergent features by examining previously unreported genes in the mouse literature that were enriched or selective to human vVMH and vARC (**Fig. 3**). In both vARC and xARC, we identified robust expression of *GABRE* (encoding a GABA-A receptor subunit, **Fig. 3B**) and selective expression of *CYP26A1* (**Fig. 3C**), a gene implicated in retinoid metabolism[99]. *CYP26A1* expression was strongest in xARC-*AGRP* cells from both sexes (**Fig. S24**). In vVMH and xVMH, we observed prominent enrichment of *ANKRD34B and LAMP5* (**Fig. 3D-E**). *LAMP5* is highly expressed in human cortical inhibitory neurons[100,101] and while previously described in rodent VMH[96], we were surprised by its robust expression within the densely excitatory VMH. Within xVMH, 94% of cells from the three primary xVMH neuron types coexpressed excitatory marker *SLC17A6* with *LAMP5* (42,767 of 45,332 *LAMP5*+ cells), while only 4% expressed *GAD1*/*GAD2* in the absence of *SLC17A6*. smFISH on adjacent tissue sections showed that *LAMP5* is coexpressed with *SLC17A6* and *NR5A1,* but absent from *GAD1*- and *GAD2*-expressing neurons in and surrounding VMH (**Fig. 3F**, **Fig. S25**). The expression of *LAMP5* in inhibitory neurons of the cortex and excitatory neurons of the VMH highlights the unique cellular identity of the VMH compared to other brain regions.

We next sought to identify transcriptional regulators of vVMH- and vARC-defining expression patterns by performing GSEA using aggregated databases of putative transcription factor (TF) target genes across mammalian tissues and cell types (**Table S4**, **Data S3**; see *Methods*). Gene targets of the glucocorticoid receptor (*NR3C1*) and of a putative retinoic acid receptor (*NR2F2*) were enriched in the vVMH[102]. Similarly, targets of the traditional retinoid receptors (retRs) *RXRA* and *RARA* were enriched in vVMH and vARC, respectively (**Fig. 3G**). Given limited discussion of retinoid pathways in rodent HYP, we compared the expression of *CYP26A1* and retRs between human vVMH/vARC and previously published VMH and ARC neuronal single cell 'subclasses' in adult mouse brain[85]. Notably, mouse VMH and ARC did not express *CYP26A1*. In addition, retRs showed greater expression in human vVMH and vARC relative to cell types from these regions in mouse brain, with particularly stronger gain of expression in human vVMH (**Fig. 3H**). This suggests that retinoid-mediated signaling pathways may have divergent roles in human and mouse HYP, consistent with prior observations of species differences in retR distribution between mouse and rat ARC[103].

Finally, several results suggested sex differences in vVMH and vARC, including androgen receptor (*AR*) target enrichment in VMH and *AR* interactors (*i.e.*, co-TFs)[104] target enrichment in ARC. Furthermore, genes upregulated or downregulated by perturbation of *EGR1,* a TF that sex-differentially modifies chromatin in the hippocampus[105], were respectively enriched in vARC or vVMH. These findings highlight previously

underappreciated signaling pathways in the human VMH and ARC and justify further examination of sex differences in these HYP regions.

Molecular sex differences in VMH and ARC domains

To assess transcriptome-wide sex-differential expression in vARC and vVMH, we performed pseudobulk DE analyses, which revealed 11 significant vVMH (**Fig. 4A**) and 26 significant vARC (**Fig. 4B**) sex-DEGs (FDR<0.05, **Table S9**). While all 11 significant sex-DEGs in the vVMH were allosomal, 13 of the 26 significant DEGs in vARC were autosomal. The distribution of p-values for both analyses showed an overrepresentation of p-values close to zero, suggesting a sizable number of 'true' differentially expressed genes in both domains, higher than the 11/26 we report here (**Fig. S26A-D**).

We observed a strong concordance between autosomal sex-differential expression in vARC and vVMH, with greater fold changes in vARC. Effects of sex were correlated (*t*-statistic Spearman rho=0.542; logFC Spearman rho=0.585; **Fig. S26A-B**), and the correlation increased when we restricted to genes nominally DE in either of the two domains (t-statistic rho=0.586; logFC rho= 0.683). Among nominal autosomal sex-DEGs in Visium, 211 were found in both domains, of which 210 had concordant effects. The one exception was histamine H1 receptor (*HRH1*), which showed greater vARC expression in females (logFC 0.77) and greater vVMH expression in males (logFC=0.49). Analysis of the p-value distribution of vVMH conditional on vARC (and vice versa) likewise suggested a substantial overlap among the sex-DEGs in the two domains (**Fig. S26E-F**).

Based on a preliminary analysis of Visium data, we had prioritized 36 putative sex-DEGs for deeper characterization with Xenium. We also included the androgen receptor (*AR*) and estrogen receptors *ESR1* and *GPER1* (G-protein coupled estrogen receptor). To examine whether Xenium sex-DE patterns were consistent with Visium, we pseudobulked across all cells within xVMH or xARC or all spots within vVMH or vARC and examined domain correlations across platforms. Autosomal effects of sex (*t* statistics) were correlated across the two platforms (ARC, 277 genes, Spearman rho=0.63, $p<2.2\times10^{-16}$; VMH, 245 genes, Spearman rho=0.41, $p<5.5\times 10^{-11}$; **Fig. S26G**). The correlation increased when only considering the 36 genes earmarked for sex-DE follow-up on Xenium (ARC, 33 genes, rho=.80, $p<5.8\times10^{-7}$; VMH, 27 genes, rho=0.51, $p<0.01$; **Fig. S26H**). This suggests that sex effects at the domain level were consistent across platforms, regardless of DE significance.

To address whether sex-DE might be conserved in ARC and VMH between mouse and humans, we compared sex-DE signatures of Visium domains to adult mouse HYP single cells[86]. Leveraging the same brain-wide single cell atlas used to illustrate marker conservation of vVMH and vARC (**Fig. 3A**), we performed sex-DE analysis on 350K HYP cells from VMH/ARC mouse subclasses, aggregated to a single label per mouse domain (**Fig. S27A**, **Data S4**). Mouse VMH and ARC neurons did not achieve Spearman correlations >0.1 to vVMH or vARC when examining anywhere from 25 to several thousand top human domain sex-DEGs (**Table S10A**). This result suggests that either sex-DE patterns are weakly conserved between mouse and human in VMH and ARC, or that sex-DE in domains containing heterogenous cell types in the human does not recapitulate cell type-specific sex-DE patterns from the same domains in mouse.

Next, we examined the biological relevance of sex-DE genes in human VMH and ARC. Domain-level sex-DEGs on both platforms included genes with HYP-relevant functional roles and regulation. vARC showed greater female expression of *RAMP2* (logFC -1.49, FDR<0.05; **Table S9**), which directly interacts with calcitonin-family receptors and glucagon receptors[106], consistent with ARC's role in metabolic regulation. Female vVMH and xVMH both showed greater expression of *TAC1* (**Fig. 4C**), a neuropeptide transmitter, and male vVMH demonstrated elevated apelin receptor, *APLNR* (logFC 1.22, $p<0.01$). Interestingly, delivery of apelin into rodent VMH increases lordosis (reproductive behavior) in females[107], supporting a sex-differentiated role for this gene. In line with vVMH/vARC marker enrichment in retR-regulated genes, both vARC and xARC demonstrated female upregulation of a retinol dehydrogenase, *RDH10* (**Fig. 4D**) and male upregulation of the retR *RXRG* (**Table S9**).

Lastly, we examined domain-level expression of estrogen, progesterone, and androgen receptors using both platforms. As anticipated from rodent and human data[108,109], *ESR2* expression was absent in VMH and ARC (**Fig. S28A**). Visium expression of progesterone receptor (*PGR*) was moderate in both vVMH and vARC, consistent with its protein expression pattern in human HYP[110]. *PGR* was also nominally sex-DE (male-elevated) in vARC (logFC 0.94, $p<0.001$) and trended similarly in vVMH (logFC 1.23, p=0.053; **Table S9**). *PGR* (Visium only), *ESR1*, and *AR* were detected in a greater fraction of ARC spots/cells compared to that of VMH. Contrary to expectations from rodent literature showing enrichment of *ESR1* in the vlVMH[2–4,8,10,21,22], *ESR1* expression was sparse and spatially uniform in human xVMH. smFISH on Visium-adjacent tissue sections reproduced this finding (**Fig. S23A**; see *Data Availability* for full-tissue *ESR1* smFISH microscopy). In sum, we find that sex-DEGs of ARC and VMH are consistent with known regional functions and gene-regulatory pathways identified for these domains. The dense expression of *ESR1* in human ARC compared to VMH further suggests that DE observations are driven by hormonally-mediated (activational) signaling.

Single-cell analysis reveals especially marked sex effects in *ESR1*-expressing neurons of ARC

Next, to determine whether VMH and ARC showed cell type-specific sex differential expression, we examined all Xenium-measured genes for sex-DE within individual cell types in xVMH or xARC domains (**Fig. S29**, **Table S12A-B**). In xVMH, we found the most significant (FDR<0.05) sex-DEGs in excitatory *CRHR2*-expressing cells, involving 39 genes, with another six DEGs in the xVMH-Glia mixed cluster (**Fig. 5A**, **Fig. S30**). Among these, *TAC1* was female-upregulated in the xVMH-Glia mixed cluster with strong trends (FDR<0.1) towards female upregulation in *CRHR2* and *LAMP5* clusters (**Fig. 5B**). Interestingly, *Tac1* in rodents is male-upregulated in *Esr1*-expressing neurons of mouse VMH[111], further supporting a sex differential role for this gene. Male xVMH *CRHR2* neurons showed significantly greater expression of *MC4R*, an obesity-associated[112] melanocortin receptor also expressed in *Esr1*+ neurons of mouse VMH[8] (**Fig. 5B**). Lastly, *TMEM233* showed no appreciable DE in individual xVMH neuronal clusters (**Fig. 5B**) despite trending toward DE at the domain level (**Fig. 4A**), suggesting this particular observation is attributable to non-neuronal populations.

Sex-DE (FDR<0.05) in xARC was most extensive in *TAC3-ESR1* neurons (67 genes, **Fig. 5C**), and detected in all xARC populations except the *POMC*-Oligo mixed cluster (**Fig. S30**). *TAC3-ESR1* neurons showed specific and marked male upregulation of *CDH13*, a receptor for insulin-sensitizing peptide adiponectin[113] implicated in obesity and type 2 diabetes[114–116] (**Fig. 5D**). Highlighting retinoid pathway genes, *RXRG,* a male-upregulated DEG in vARC, was also significantly male-upregulated in three xARC clusters and a trend toward the same in *ESR1-TAC3* neurons. Consistent with domain level findings, retinoid metabolism genes were female-upregulated in xARC with *RDH10* showing DE in four populations and vARC marker *CYP26A1* showing DE in all five xARC clusters (**Fig. 5D**). No significant DEGs showed opposite directions of effect among or across xARC and xVMH cell types. However, in agreement with vARC and vVMH findings, *HRH1* was significantly upregulated in two xARC clusters from females while trending toward male upregulation ($p<0.07$) in xVMH *CRHR2* cells.

We returned to examining sex hormone receptor expression patterns, now at the finer resolution of individual cell clusters (**Fig. S28B**). The greatest proportion of cells with ≥1 count of *ESR1* included xARC *TAC3-ESR1*, xARC *AGRP*, and xVMH *CRHR2*, in addition to cells of microglia cluster 1 and oligodendrocyte progenitors (OPCs) residing in either xVMH or xARC. *AR* was generally found in a similar proportion of cells for each cluster-domain pair, with the notable exception of xARC *GAL-TRH-GHRH*, where it was detected in ~2-fold more cells than was *ESR1*.

Next, we continued our comparison to mouse HYP single cell sex-DE, examining correspondence between each xARC/xVMH neuron cluster and each mouse ARC/VMH cell subclass. As with Visium domain vs. mouse subclass comparisons, most cell populations showed greater human-mouse correlation of cell type marker gene ranks than of sex-DEG ranks. Interestingly, only the xVMH and xARC populations with glial

marker enrichment demonstrated comparable human-mouse correlations for marker and sex-DEG ranks (**Fig. S27B-C**, **Table S10B**). Overall, both marker and sex-DE correlations were low; the maximum Spearman correlation across all cell type marker and sex-DE comparisons was ~0.4.

To confirm that our findings were not confounded by ARC/VMH cell type intermixing at domain boundaries, we repeated the sex-DE analyses by reducing both the xVMH and xARC domain boundaries inward by 4%, which gave well-correlated t-statistics relative to the original domain boundaries for all 5 xARC clusters and the 3 main xVMH clusters (all Spearman rho > 0.99) (**Table S12F-H**). Similarly, we ruled out effects of cell segmentation by repeating the analyses using only nuclear transcript counts, which also yielded consistent results (**Table S12D-E** and **H**). Altogether, we replicated and refined several Visium findings at higher spatial resolution and identified specific xARC and xVMH cell types responsible for sex-DE signals. Moreover, we identified populations under the strongest transcriptional influence of sex in each domain and illustrate that molecular-functional HYP axes are influenced by sex in the human VMH and ARC.

Male upregulated genes in VMH and ARC are enriched for neurodevelopmental disorders

Given that the prevalence rates for many neurodevelopmental (NDD) and neuropsychiatric disorders are sex biased and that the HYP plays established functional roles in behaviors relevant for these disorders, we hypothesized that expression patterns defining VMH and ARC may be enriched in NDD and neuropsychiatric disorder gene sets. Toward this end, we collected gene sets from genome-wide association studies (GWAS), rare-variant studies (where applicable), transcriptome-wide association studies (TWAS)[78–82], and the database DisGeNET[83], which includes curated disease-gene links (**Table S4**). These gene sets were combined into a single gene set for each of 18 disorders. We then performed GSEA to identify whether disease-associated genes demonstrated marker-like expression in vARC or vVMH. While vARC did not show an enrichment of genes associated with neuropsychiatric disorders (**Table S11**), vVMH showed enrichment of genes associated with several sex-biased disorders, including autism spectrum disorders (ASDs), neurodevelopmental disorders (NDDs), and schizophrenia (SCZ) (**Fig. 6A**).

Next, we performed GSEA on sex-DEGs in vVMH or vARC to determine whether disease gene sets were enriched in male- and/or female-upregulated genes. Male-upregulated genes from both vVMH and vARC were overrepresented in genes associated with risk for ASDs and Tourette's syndrome (TS), which are disorders more prevalent in males (**Fig. 6B**, **Table S9E**). While our Visium analysis only identified three ASD-associated genes with nominal sex-DE in vARC, *TAC3-ESR1* cells in xARC showed sex-DE (FDR<0.1) for 11 such genes (*AR*, *CDH13*, *DCC*, *GABRA5*, *GABRQ*, *GLRA2*, *MYO16*, *MYT1L*, *NR4A2*, *PCLO*, *THBS1*; **Table S13**). *MYT1L*, a gene mutated in a syndromic form of ASD involving obesity[117], was significantly male-upregulated selectively in *ESR1-TAC3* neurons, providing previously unreported cell type context for this risk gene. These findings exemplify how gene panel prioritization using Visium improves our ability to detect sex-DEGs in anatomically-restricted cell types, while retaining transcriptome-wide information to link such genes to broader pathways.

We next turned to examining individual sex-DEGs on Xenium at FDR<0.1, focusing on 8 neuropsychiatric/neurodevelopmental disorders with prominent sex differences in prevalence and/or clinical features (ADHD, anorexia nervosa, ASD, MDD, Tourette's syndrome, SCZ, generalized anxiety disorder and panic disorder). In total, we obtained 122 disease-associated genes with sex-DE in ≥1 cluster within ARC and 35 in ≥1 cluster within VMH, with 29 genes in common between the two domains. These clusters include glial cells, interneurons, and other broad cell types that may interact differently with VMH- and ARC-specific neuronal populations (**Fig. S29**). Strikingly, there were 70 ARC/22 VMH sex-DEGs associated with MDD, 68 ARC/16 VMH genes associated with SCZ, and 31 ARC/12 VMH associated with ASD (**Table S13**). MDD-, SCZ-, and ASD-associated genes were also frequent among sex-DEGs of domain-specific clusters, with 25 genes in ARC neurons and 22 in VMH neurons for MDD, 31 ARC/16 VMH for SCZ, and 13 ARC/12 VMH for ASD. While our Xenium data is not a genome-wide assessment, it establishes that genes associated with

several sex-biased disorders are also sex-DE in individual cell types of the ARC and/or VMH, including their defining neuronal populations.

We then asked whether vVMH or vARC sex-DE corresponds to broader aspects of HYP biology or gonadal hormone signaling by performing similar GSEA analyses with pathway and ontology gene sets (**Data S5**). Male-biased vVMH genes were enriched for gene sets associated with aggression, which is consistent with behavioral functions of male rodent VMH in regulating aggressive behavior[21,23]. On the other hand, male-biased vARC genes were enriched for targets of *miR-30a*, a regulator of puberty onset downregulated by perinatal estrogen exposure in female rodents[118] (**Fig. 6C**). Progesterone-responsive genes were enriched among female-biased genes of vVMH (**Fig. 6C**), despite trends toward greater male *PGR* expression in both vVMH and vARC.

Given the above findings implicating specific gene regulatory pathways in sex-DE, we performed GSEA using TF-target gene sets and filtered results to TFs expressed in vVMH or vARC (see *Methods*) (**Fig. 6D**, **Data S6**). Sex hormone receptor targets were enriched for DE in both ARC and VMH. While *ESR1* target sets were enriched strongly in female vVMH and vARC, male vVMH was weakly enriched for one target set and for ESR1 protein interactors. Surprisingly, female vVMH showed stronger enrichment for inferred androgen receptor targets than did male vVMH. Interestingly, the above pathway GSEA findings indicated enrichment of vARC sex-DEGs for *AR* targets specifically upregulated via hormone-independent *AR* activator[119] and chromatin modifier[120] *PKN1* (**Data S5**).

Our TF-target analysis also suggested sex-DE regulation by multiple TF-coding genes associated with ASDs. Male vVMH was enriched for targets of two syndromic ASD genes, *MYT1L*[117] and *ASH1L*[121] (**Fig. 6D**). Moreover, *MYT1L* was itself nominally male-upregulated in xARC *ESR1-TAC3* neurons (**Fig. 6E**). On the other hand, female vVMH and vARC were enriched for targets of ZNF142, which is associated with a rare, female-predominant neurodevelopmental syndrome[122] and had divergent target patterns between sexes. Genes upregulated by disruption of *ZNF142* were enriched among female-biased vVMH and vARC genes, while those downregulated by disruption of *ZNF142* were enriched among male-biased vVMH DEGs. As these targets are ascertained from TF perturbation experiments, this would suggest that ZNF142 activity in VMH is *ordinarily* greater in males than in females. These results provide further evidence that molecular sex differences in the VMH and ARC are highly relevant for NDDs and ASDs.

Finally, given that the HYP plays a key role in energy metabolism, we also investigated gene sets relevant for body mass index (BMI) and obesity. GSEA did not reveal overrepresentation of BMI-associated genes in either vARC- or vVMH-enriched genes. However, GSEA using TF-target gene sets showed that both male and female vARC, as well as female vVMH, were enriched in target sets for *STAT3*, a TF necessary for leptin signal transduction. We also found female-specific enrichment of both vARC and vVMH for targets of *PPARA* and *PPARG*, which play key roles in lipid and sterol metabolism (**Fig. 6C**). Interestingly, both of these TFs appear to compete with estrogen receptor α for shared DNA binding sites[123–125]. In conclusion, these enrichments collectively illustrate that molecular sex differences in tuberal HYP are organized by gene pathways and regulatory hubs critical to HYP functions, gonadal hormone signaling, and disease risk.

Discussion

The human tuberal HYP contains regions of broad interest in behavioral and endocrinologic neuroscience, but its molecular anatomy has been relatively unexplored, especially in the context of sex differences. Here, we performed comprehensive spatial molecular profiling of two hypothalamic regions, VMH and ARC, in a carefully selected set of healthy male and female control donors enabling comparative transcriptomic analyses in the absence of critical brain or endocrinologic perturbations. Our findings highlight features of HYP that are distinct from more commonly studied cortical regions, including unexpected *LAMP5* expression in excitatory neurons of the VMH, complex/non-laminar cell type organization in the ARC, and distinctive neuropeptide expression in these domains and their surroundings. Moreover, we identified signatures of sex-regulated gene expression in VMH and ARC, which we localized to specific neuronal

populations. Finally, we highlight *ESR1*-expressing neurons in ARC and *CRHR2*-expressing neurons in VMH as hubs for molecular sex differences. These studies provide a roadmap for using transcriptome-wide spatial assays to prioritize genes of interest for finer characterization at cellular resolution using imaging-based spatial transcriptomics.

Given that ARC and VMH are small in size and lack clear anatomical boundaries, generating comprehensive molecular atlases for these regions in the human brain has been challenging to date. Using Visium, we were able to perform an *in silico* dissection of VMH and ARC based on gene expression. While we did not observe discrete spatial clusters within vARC or vVMH, we identified a preponderance of highly spatially variable genes (SVGs) in vARC, which overlapped with established marker genes of postmitotic human fetal ARC neurons[97]. We hypothesized that higher resolution assays, such as Xenium, would facilitate the identification of spatially restricted neuronal populations. Indeed, Xenium experiments recapitulated vARC SVG patterns and identified xARC clusters with distinct spatial localization. While there is clear evidence for substructure within the human ARC, we could not detect spatial subdivisions in human VMH. With Visium, VMH clusters were distinguished primarily by extent of glial gene expression, and in Xenium, VMH neuronal clusters were intermingled. We also did not find clear expression of rodent VMH subpopulation marker genes, including those previously reported to be spatially localized, such as *Esr1*-expressing neurons in the posterior ventrolateral rodent VMH.[3] Supporting our findings, a single-nucleus RNA-seq study of the developing human HYP through gestational week 20 (a stage at which nearly all neuronal populations have emerged) revealed only one postmitotic VMH cluster.[97] There are several possible interpretations for a lack of VMH substructure in our datasets. First, human VMH neuronal cell types may not appreciably exist in spatial niches, or such a phenomenon may only be discernible with very large, cell-resolution datasets and granular clustering. Second, human VMH cell types might be defined by different genes than rodent VMH cell types and for which spatial variability could not be detected within human vVMH. Third, it is possible that molecularly-defined VMH subdivisions exist in a specific anatomical plane across the anterior-posterior axis that was not captured in our study.

Nonetheless, marker genes of human vVMH and vARC strongly mapped to the corresponding domains in mice based on integrative studies with a previously published rodent dataset. Several marker genes, especially in ARC, point to conserved physiologic roles implicit in the expression of well-characterized neuropeptides such as *KISS1*, *GNRH*, *AGRP*, and *NPY*. Likewise, domain-enriched genes were enriched in functional terms consistent with rodent studies of VMH (e.g., regulation of blood glucose [92,93]) and ARC (e.g., primary cilia). Primary cilia are ubiquitous cell structures implicated in the role of ARC in detecting nutritional status. Mice deficient for primary cilium assembly show obesity[24]; likewise, mutations disrupting primary cilium localization of the melanocortin 4 receptor, which is expressed in ARC, are associated with obesity in humans[126]. Despite generally shared molecular identities between rodent and human ARC and VMH, we also identified human-specific marker genes, such as the retinoid pathway gene *CYP26A1*. Interestingly, the ARC of mouse and rat diverge in their expression of retinoid-synthesizing enzymes and retRs; moreover, their expression is regulated by photoperiod only in rats[103], further supporting our hypothesis that retinoid pathways are ‘evolutionarily plastic’ in the tuberal HYP. We identified additional convergent lines of evidence for retinoid activity in human VMH and ARC, including enrichment of VMH- or ARC-specifically expressed genes in regulatory targets of retRs and increased retR expression in these domains relative to those in mouse.

Given that the HYP is a sexually dimorphic brain region, we examined sex-DE in ARC and VMH. At the domain level, we found more significant genes and larger fold changes in ARC compared to VMH. Consistent with literature showing that rodent VMH populations are developmentally and sex-differentially established by androgens[10], our vARC and vVMH sex-DE signatures were enriched for androgen receptor (AR)-regulated genes. These included genes induced downstream of protein kinase PKN1 (**Data S5**), providing a possible mechanism for the establishment of sex-DE. Consistent with Visium analysis, we identified robust sex-DE, which we resolved to specific neuron subtypes. For example, we pinpointed *ESR1*-expressing neurons as the predominant site of sex-DE within ARC. We also identified sex-DEGs common to several or all domain neuron types, including the retinol dehydrogenase *RDH10*, which was DE in all xARC neuron clusters. Finally, sex-DE

was well-correlated between ARC and VMH, suggesting that sex similarly regulates expression in these two domains. Monosynaptic circuits connecting VMH-ARC have been traced in rodent HYP[127]; if such circuits exist in human HYP, sex-DE concordance may be partially driven by capture and subsequent quantification of mRNA from neurites of VMH neurons that project to the ARC.

To support translational research efforts, we conducted several cross-species analyses using available rodent datasets[85,86]. First, we compared sex hormone receptor expression and distribution between human and mouse HYP. Notably, *Esr1* expression in human VMH and ARC diverged from previous findings in rodents. In mice, *Esr1*-expressing neurons in the VMH, especially in the ventrolateral subdivision, are a well-studied hub of rodent sex differences[1,3,8]. By contrast, we observed uniformly sparse *ESR1* expression in human VMH, but dense, high expression in a ventromedial pocket of ARC. We validated this expression pattern using 3 complementary modalities: smFISH, Visium, and Xenium. *ESR1*-expressing cells in human ARC included a subpopulation of *TAC3*+/*KISS1*+ neurons likely representing KNDy neurons previously described in rodent ARC[5,128]. KNDy neurons play a key role in hormone feedback and the luteinizing hormone (LH)/follicle stimulating hormone (FSH) cycle[9,27,28], so it is not surprising that the xARC population containing these neurons would show the most sex-DEGs. At the same time, it is difficult to infer what proportion of our ARC *ESR1-TAC3* population comprised probable KNDy neurons as we were not able to include KISS1 on our gene panel. While smFISH indicated it was likely a minority of *ESR1* neurons, it is possible that gonadal hormones may have repressed key transcriptional markers of KNDy cells resulting in lower *KISS1* expression. This speculation is based on mouse studies showing that KNDy marker genes (*Kiss1, Tac2*) are repressed substantially by gonadal hormones[98,129,130], which we assume were circulating in our subjects. Finally, while human and rodent VMH/ARC show preserved molecular identities, sex-DE of human vVMH/vARC domains and xARC/xVMH cell types corresponds less strongly to that of broad VMH/ARC mouse cell types (**Fig. S27**). As we performed each human-mouse comparison considering cell/domain marker genes or sex-DEGs, we conclude that the comparatively weak correspondence of sex-DE patterns is not a product of technical differences (i.e., single-cell RNA-seq in mouse vs. spatial transcriptomic methods). While the human-mouse discordance in sex-DE may be influenced by cell type classification across species, divergence in sex-DE could also be explained by differences in the evolutionary niches that dictate suitable feeding, fear, and parenting behaviors.

Last, we explored relationships between HYP gene expression and genetic factors influencing BMI and neuropsychiatric disorders. Surprisingly, we did not find BMI-associated genes to be overrepresented in VMH or ARC, despite the role of HYP in appetite and metabolism. However, our method of gathering genes for this analysis yielded an improbable number of BMI-associated genes (over 8000, *i.e.*, a null expectation that >50% of query genes would be BMI-associated; results using subsets of the ascertained genes for each phenotype are in **Table S11** and **Table S9**). However, in ARC, we did note sex-DE of *CDH13*, a receptor for the insulin-sensitizing hormone adiponectin. In VMH, we observed sex-DE of the appetite-regulating and obesity-associated[112] melanocortin receptor *MC4R*. However, the most noteworthy finding, supported by multiple analyses, was the robust enrichment of male-upregulated genes of both vVMH and vARC for ASD and NDD risk gene sets. Moreover, male-upregulated genes were overrepresented in targets of TFs encoded by genes mutated in syndromic ASDs/NDDs, including *ZNF142*[122], *TCF20*[131], *ASH1L*[121], and *MYT1L*—a syndrome which includes obesity[117,132]. An interesting single-gene link to MDD, which occurs more often in females, also emerged: the sex-DEG *CDH13* is repressed by a human-specific lncRNA *FEDORA*, which has been shown to be sex-DE in MDD cortex and to exert sex-differential behavioral effects when expressed in mouse cortex[133–135]. Moreover, *MYT1L* was itself significantly male-upregulated in *TAC3-ESR1* KNDy neurons of the ARC. These findings highlight VMH and ARC as candidate HYP regions susceptible (perhaps sex-differentially) to common- and rare-variant mediated neurodevelopmental disorder risk.

We note several limitations in the present study. Due to low expression levels, we were not able to detect dopaminergic neuron marker genes in ARC, despite their known presence[136,137] and important function in lactation[138]. Additionally, there are methodological limitations to both spatial transcriptomic approaches employed. While Visium is transcriptome-wide, it is lower resolution and prohibits granular spatial-molecular

organization of small tissue areas such as VMH and ARC. We addressed this limitation by employing Xenium, which resolves single nuclei in space, but at the expense of measuring only a fraction of the transcriptome. Our Xenium data cover only 366 genes, potentially preventing identification of cell types, including tanycytes and cells of VMH. In the former case, tanycyte labels were assigned to rare cells in the median eminence bordering VMH/ARC as would be expected[139], but also to cells intermingled with connective and vascular clusters in many samples (**Fig. S31**), illustrating the limitations of cell-resolution spatial transcriptomics as a method for cell type discovery. Separately, findings regarding *ESR1*, KNDy neurons, and sex-DE could be potentially affected by oral contraceptive use by female donors, though this information was neither systematically collected nor necessarily accurate (only available when reported by next of kin). At the same time, findings on *ESR1* distribution were consistent between sexes, and our smFISH experiment for *ESR1* and *KISS1* was performed on tissue from a male donor. While other recent human atlasing efforts have made headway in profiling a broader extent of HYP[43,44], they have not been powered to address sex differences in any specific area. Future studies should build upon this and other foundational molecular work in human HYP by profiling its domains in-depth—that is, at similar or greater sampling depths to ours here—in order to power studies contrasting the effects of sex or physiology (BMI, reproductive status, etc.) on HYP.

In sum, these data represent a unique transcriptomic resource focused on the human tuberal HYP, an area important for endocrine, reproductive, and behavioral functions. We provide a spatially-resolved single cell atlas of the adult human HYP and demonstrate that the molecular identities of VMH and ARC are generally similar between mice and humans. Further, we show that autosomal gene expression differs across the sexes in both VMH and ARC, often in a cell type-specific manner. Finally, we show that sex differences in gene expression are associated with ASDs and NDDs, disorders with greater prevalence in males. To augment ongoing HYP research, we generated interactive web-based data browsers for the scientific community to explore these data at bit.ly/libdVisHYP (Visium) and bit.ly/libdXenHYP (Xenium).

Acknowledgements

We thank the Joint High Performance Computing Exchange (JHPCE) at Johns Hopkins University for providing computing resources for these analyses. We thank the JHU Single Cell Transcriptomics core for Illumina sequencing and Xenium data generation. We thank the families of Connie and Stephen Lieber and Milton and Tamar Maltz for their generous support of this work. We thank the LIBD neuropathology team, particularly James Tooke and Amy Deep-Soboslay, for curation of the brain samples and assistance with tissue dissections, and the families of the brain donors for their generosity. We also would like to thank William Reay, PhD, for insightful discussion regarding retinoid genes/pathways in the brain and disease. We thank Madhavi Tippani for reviewing and testing code for smFISH puncta quantification and visualization in R. Schematic illustrations were generated using Biorender and Adobe Illustrator.

Funding

This work was supported by the Lieber Institute for Brain Development, 10X Genomics, and National Institutes of Health awards T32MH015330 (BM) and R35GM149323 (KDH).

Code and Data availability

From this study: The source data described in this manuscript, including FASTQ files from Visium and unfiltered sample-level *xeniumranger* 1.7 outputs (transcript-level data) compatible with Xenium Explorer software are respectively available through GEO accessions GSE280316 and GSE280460. Supplemental Data files, smFISH microscopy images, the Visium *SpatialExperiment* (including bundled H&E images) and Xenium *SpatialFeatureExperiment* objects as filtered, clustered, and analyzed, and Xenium hematoxylin and eosin images (not used in the study) are available through http://research.libd.org/globus/ via the Globus endpoint jhpce#HYP_suppdata. Source code for all processing/analyses and plotting is available on GitHub (https://github.com/LieberInstitute/spatial_HYP), with contents at the time of submission archived to Zenodo (https://zenodo.org/records/14285059). Data can be interactively visualized and annotated in a web browser through our Samui[140] portals for the Visium data (with clustering available at four different resolutions from *BayesSpace*) at bit.ly/libdVisHYP (permanent URL: research.libd.org/spatial_HYP) and for Xenium (with annotated cell clusters from *Banksy*, their higher-order cell groups, and domains) at bit.ly/libdXenHYP (permanent URL: research.libd.org/spatial_HYP). Documentation for the data portals can be accessed through research.libd.org/spatial_HYP.

Outside data/databases: Raw h5ad files with single-cell RNA-seq data from adult mouse HYP[141] were downloaded from https://allen-brain-cell-atlas.s3-us-west-2.amazonaws.com/expression_matrices/WMB-10Xv3/20230630/WMB-10Xv3-HPF-raw.h5ad and https://allen-brain-cell-atlas.s3-us-west-2.amazonaws.com/expression_matrices/WMB-10Xv2/20230630/WMB-10Xv2-HY-raw.h5ad, then were filtered to cells that received cluster labels found in https://allen-brain-cell-atlas.s3-us-west-2.amazonaws.com/metadata/WMB-10X/20230630/views/cell_metadata_with_cluster_annotation.csv . TF-target gene sets (Transcription_Factor_PPIs, ChEA_2022, Rummagene_transcription_factors, TF_Perturbations_Followed_by_Expression, TF-LOF_Expression_from_GEO, TRRUST_Transcription_Factors_2019) were obtained from the gene set libraries used by the *Enrichr* tool[76] at https://maayanlab.cloud/Enrichr/#libraries on 12/06/2023. Our assertion for low expression of transcription factor RNAs is based on an interactive plot from[77] illustrating the transcripts per million for each likely human TF in each tissue in data from the Human Tissue Atlas: http://humantfs.ccbr.utoronto.ca/ .

Non-Bioconductor/CRAN R Packages: scCoco and the Allen Brain Atlas API interface it utilizes, *cocoframer*, are respectively available at https://github.com/lsteuernagel/scCoco and https://github.com/AllenInstitute/cocoframer .

Conflict of Interest/Disclosures

Joel E. Kleinman is a consultant on a Data Monitoring Committee for an antipsychotic drug trial for Merck & Co., Inc.

Author Contributions

Conceptualization: BM, KM, KDH, KRM
Methodology: AC, BM, KM, KDH, KRM, SCH, SCP, YW
Software: AC, BM, KDH, RAM, SCH, YW
Validation: BM, HRD, KDH, KDM, SVB, YD
Formal Analysis: AC, BM, HRD, KDH, KRM, RAM, YW
Investigation: AC, HRD, KDM, SC, SCP, SVB, TMH, YD
Resources: JEK, TMH
Data Curation: BM, KDH, RAM
Writing-original draft: AC, BM, HRD, KM, KDH, KDM, KRM, SC, SVB
Writing-review and editing: BM, KM, KDH, KRM
Visualization: AC, BM, HRD, KDH, KRM, RAM, SC, SVB, YD
Supervision: KM, KDH, KRM, SCH
Project administration: KM, KDH, KRM
Funding Acquisition: KM, KDH, KRM

Main Figures and Legends

Figure 1

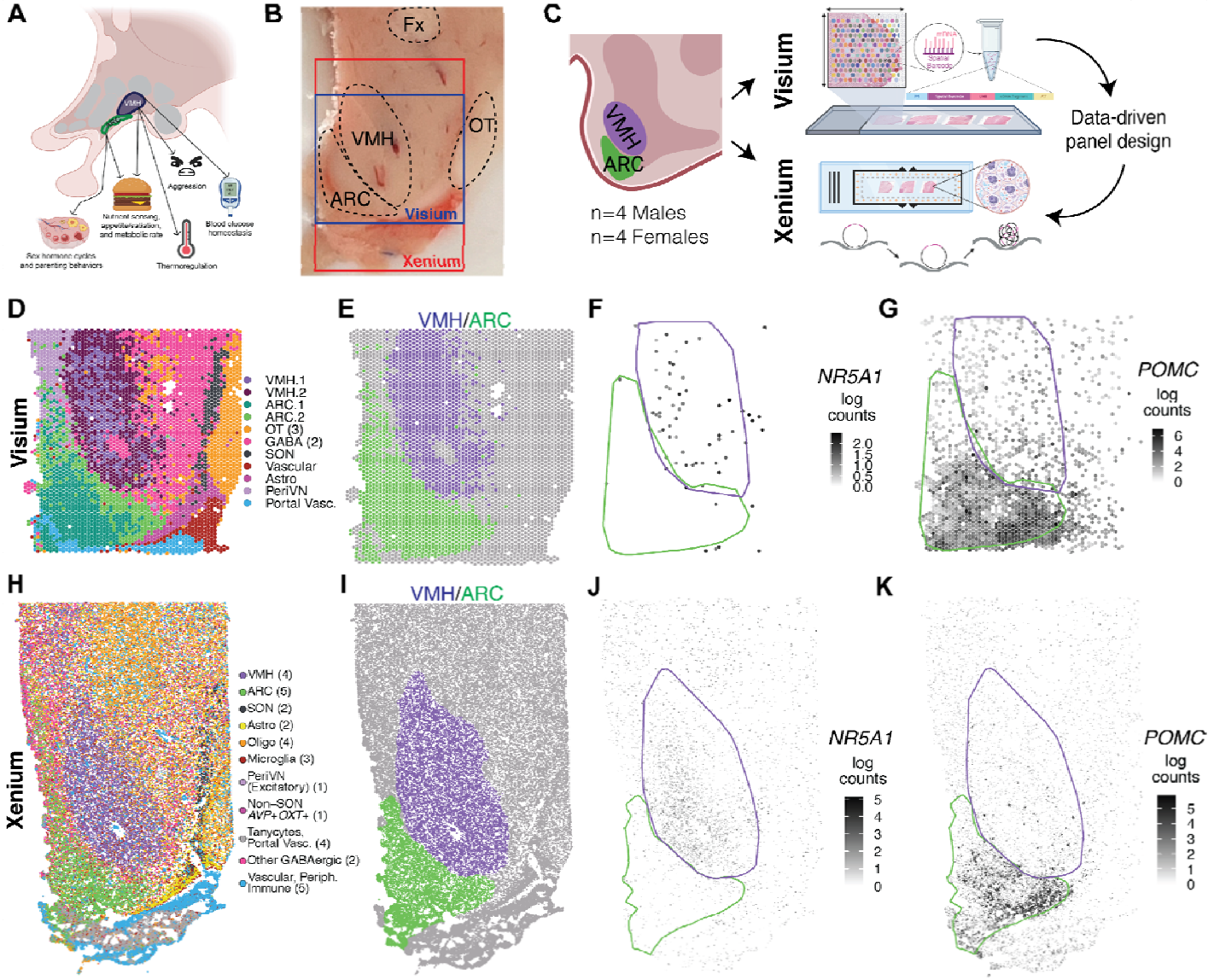

Figure 1. Study overview and identification of VMH and ARC domains. A) Sagittal view of the human HYP highlighting the arcuate (ARC) and ventromedial hypothalamus (VMH) and main functions. **B)** Tissue block (Br1225) sectioned for Visium and Xenium, indicating the ARC, VMH, and landmark white matter tracts (fornix [Fx] and optic tract [OT]). Boxes denote sampling areas for Visium (blue) and Xenium (red), which have different sized capture areas. **C)** Study design overview. We first utilized the Visium spatial gene expression platform which uses ~5,000 expression spots with unique spatial barcodes that correspond to an associated H&E image for a single capture area. We then performed *in situ* sequencing on the Xenium platform to acquire cellular resolution data for 266 pre-defined 'brain panel' genes and 100 custom genes based on domain marker, spatial, and sex-differential patterns from Visium. **D-G)** *BayesSpace* clustering of Visium data at *k*=15, with 3 optic tract clusters grouped together and likewise for 2 GABAergic neuron clusters (**D**) and subsequent identification of Visium (v)VMH and vARC domains (**E**), which respectively express known markers *NR5A1* (**F**) and *POMC* (**G**). Approximate boundaries of vVMH and vARC domains labeled in same colors as **E**. **H)** Xenium cell type clustering from the same donor as **D**, with clusters combined into Visium-analogous groups for visualization. Legend notes the number of cell type clusters in the visualized group. **I)** Xenium (x)VMH and xARC anatomical boundaries as inferred *k*-nearest neighbors-based determinations of how many spatially neighboring cells were from xVMH or xARC clusters. **J-K)** Expression of VMH marker *NR5A1* was restricted to the xVMH domain (**J**), and ARC marker *POMC* was enriched in xARC domain (**K**).

Figure 2

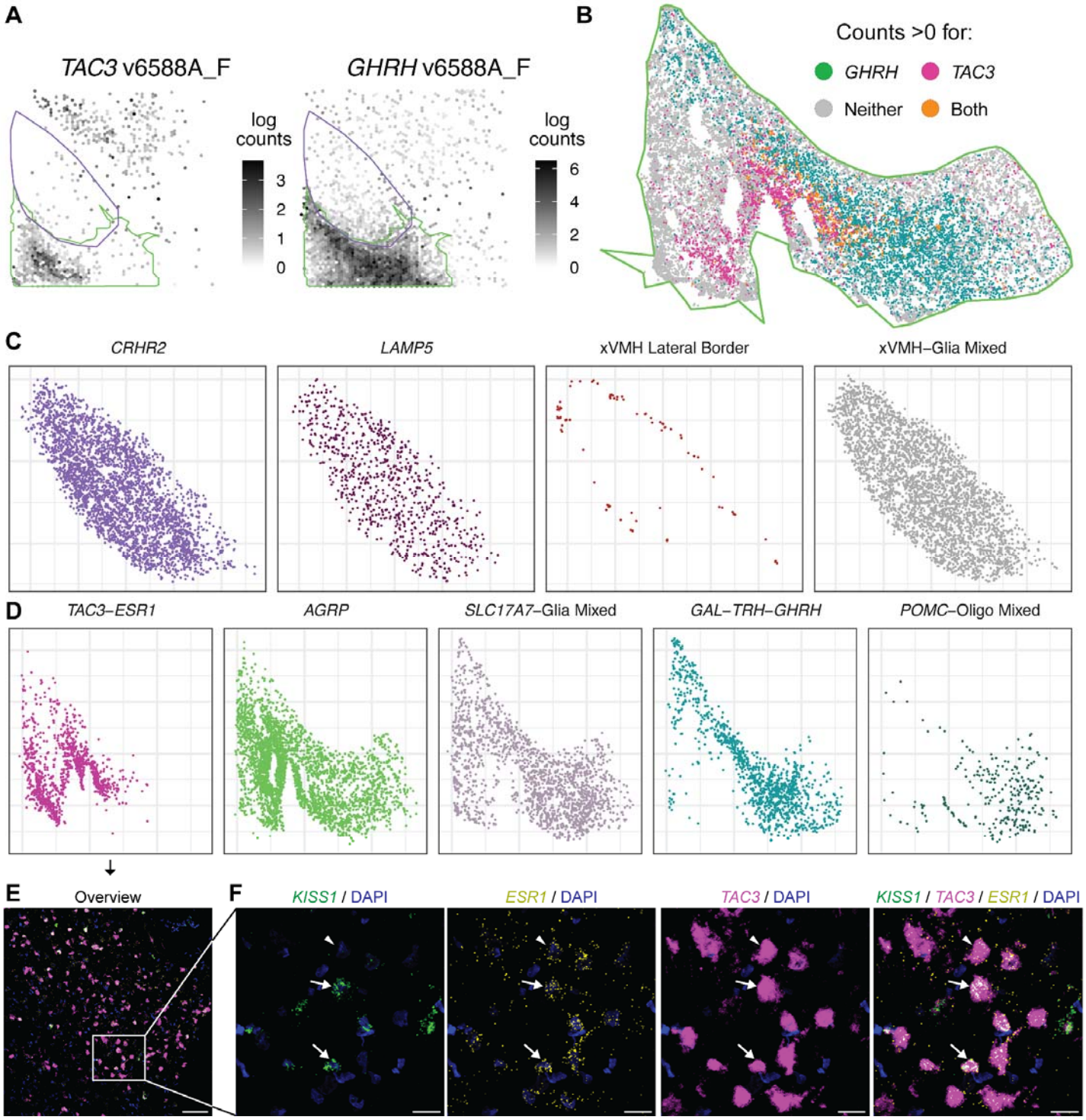
A
TAC3 v6588A_F
GHRH v6588A_F
log counts
3
2
1
0
log counts
6
4
2
0
B
Counts >0 for:
GHRH
TAC3
Neither
Both
C
CRHR2
LAMP5
xVMH Lateral Border
xVMH–Glia Mixed
D
TAC3–ESR1
AGRP
SLC17A7–Glia Mixed
GAL–TRH–GHRH
POMC–Oligo Mixed
E
Overview
F
KISS1 / DAPI
ESR1 / DAPI
TAC3 / DAPI
KISS1 / TAC3 / ESR1 / DAPI

Figure 2. Identification and spatial localization of VMH and ARC cell types using Xenium. A) Analysis of spatially variable genes (SVGs) in Visium suggested spatial organization of cell types in ARC exemplified by non-overlapping expression of *TAC3* and *GHRH* in vARC. **B)** Expression of *TAC3* and *GHRH* in individual cells of xARC shown for x6588B_F measured by Xenium. Expression of these Visium SVGs in xARC is almost entirely non-overlapping and corresponds to distinct cell clusters (see panel **D**). **C)** Identification of 4 xVMH cell clusters using *Banksy* shown for x6588B_F. Only cells within the boundaries of the xVMH domain are shown. **D)** Identification of 5 xARC cell clusters using *Banksy* shown for x6588B_F. Only cells within the boundaries of the xVMH domain are shown. **E)** Confocal image showing expression of *ESR1*, *TAC3*, and *KISS1* in an adjacent tissue section by smFISH with an area of ARC expressing all 3 genes (scale bar 100µm). **F)** Zoom in of the area boxed in **E** showing each gene with DAPI separately (scale bar 25µm). Arrows indicate cells expressing all three genes of interest. Arrowhead indicates an example cell with only *TAC3* and *ESR1* expression. Sample nomenclature denotes “v” for Visium or “x” for Xenium, followed by donor ID (BrXXXX), “A” or “B” for tissue replicates, and underscore “M” for male or “F” for female.

Figure 3

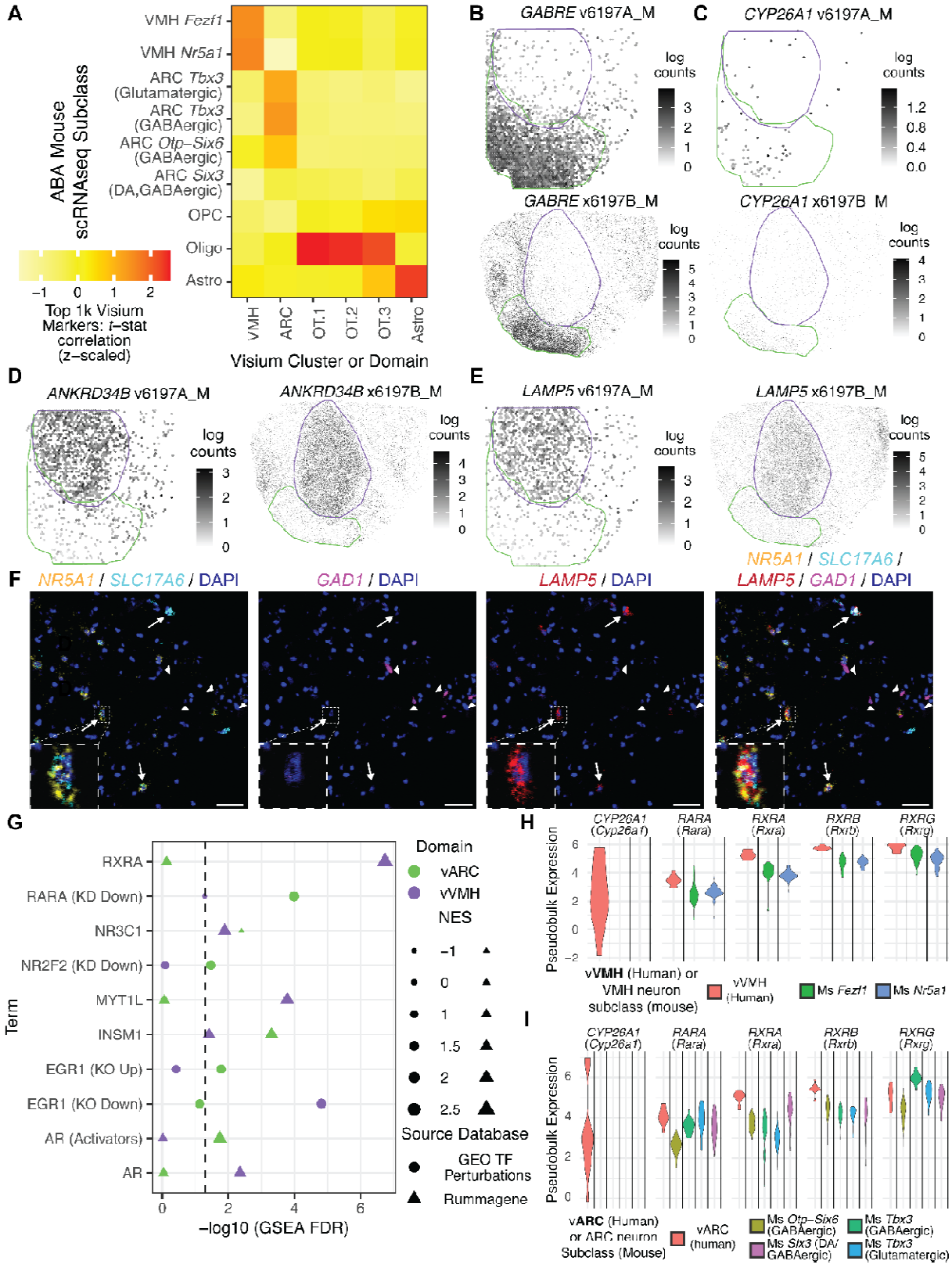
A
ABA Mouse scRNAseq Subclass
VMH Fezf1
VMH Nr5a1
ARC Tbx3 (Glutamatergic)
ARC Tbx3 (GABAergic)
ARC Otp–Six6 (GABAergic)
ARC Six3 (DA,GABAergic)
OPC
Oligo
Astro
VMH
ARC
OT.1
OT.2
OT.3
Astro
Visium Cluster or Domain
Top 1k Visium Markers: t–stat correlation (z–scaled)
B
GABRE v6197A_M
C
CYP26A1 v6197A_M
GABRE x6197B_M
CYP26A1 x6197B M
log counts
D
ANKRD34B v6197A_M
ANKRD34B x6197B_M
E
LAMP5 v6197A_M
LAMP5 x6197B_M
F
NR5A1 / SLC17A6 / DAPI
GAD1 / DAPI
LAMP5 / DAPI
LAMP5 / GAD1 / DAPI
G
Term
RXRA
RARA (KD Down)
NR3C1
NR2F2 (KD Down)
MYT1L
INSM1
EGR1 (KO Up)
EGR1 (KO Down)
AR (Activators)
AR
–log10 (GSEA FDR)
Domain
vARC
vVMH
NES
Source Database
GEO TF Perturbations
Rummagene
H
CYP26A1 (Cyp26a1)
RARA (Rara)
RXRA (Rxra)
RXRB (Rxrb)
RXRG (Rxrg)
Pseudobulk Expression
vVMH (Human) or VMH neuron subclass (mouse)
vVVH (Human)
Ms Fezf1
Ms Nr5a1
I
vARC (Human) or ARC neuron Subclass (Mouse)
vARC (human)
Ms Otp–Six6 (GABAergic)
Ms Tbx3 (GABAergic)
Ms Six3 (DA/ GABAergic)
Ms Tbx3 (Glutamatergic)

Figure 3. **Conserved and divergent features of human and mouse VMH and ARC. A)** Using a recent single cell atlas of adult mouse brain[86], we compared marker gene statistics of our Visium domains to those of cell 'subclasses' in HYP dissections from mouse (see *Methods*). Plot shows z-normalized *t*-statistic correlations (Spearman) between the top 1,000 Visium markers for a given domain and their orthologs in a given mouse subclass. **B)** Spotplots showing selectively high expression of *GABRE,* encoding a GABA receptor subunit, in vARC (top) and xARC (bottom) denoted by green boundary. **C)** Spotplots showing retinoid pathway gene *CYP26A1* is strongly enriched in vARC (top) and xARC (bottom) denoted by green boundary. **D)** Spotplots showing high expression of a previously unreported VMH marker gene, *ANKRD34B*, in vVMH (left) and xVMH (right) denoted by purple boundary. **E)** Spotplots showing expression of *LAMP5*, an inhibitory neuron marker gene in the cortex, which is strongly enriched in vVMH (left) and xVMH (right) despite the generally excitatory identity of VMH. **F)** Confocal image showing expression of *NR5A1*, *SLC17A6*, *GAD1*, and *LAMP5* by smFISH in an adjacent tissue section from Br1735, at the intersection of the ventrolateral edge of VMH with ARC and areas lateral to VMH/ARC. We confirm that excitatory *SLC17A6+/NR5A1+* VMH neurons express *LAMP5*, unlike inhibitory (*GAD1+*) neurons. Arrows indicate *LAMP5+* excitatory VMH neurons; arrowheads indicate *GAD1+* inhibitory neurons ventral/lateral to VMH. Scale bar 50µm. (smFISH targeting GAD2 instead of GAD1 is shown in **Fig. S25**). **G)** GSEA analysis of vVMH- and vARC-enriched genes for aggregated mammalian TF-target gene pairs reveals potential roles for the androgen receptor (*AR*) and its co-TFs, as well as retinoid receptors (*RXRA*, *RARA*) in human ARC/VMH-biased expression patterns. NES = normalized enrichment score. **H-I)** Using the same mouse data in **A**, we examined the expression of *CYP26A1* and retinoid receptors as implicated by our GSEA analysis. Compared to mouse ("Ms") VMH subclasses (top) or mouse ("Ms") ARC subclasses (bottom), several retinoid receptors appear to be more highly expressed in vVMH (human) and vARC (human); additionally, *CYP26A1* is only appreciably expressed in human VMH and ARC.

Figure 4

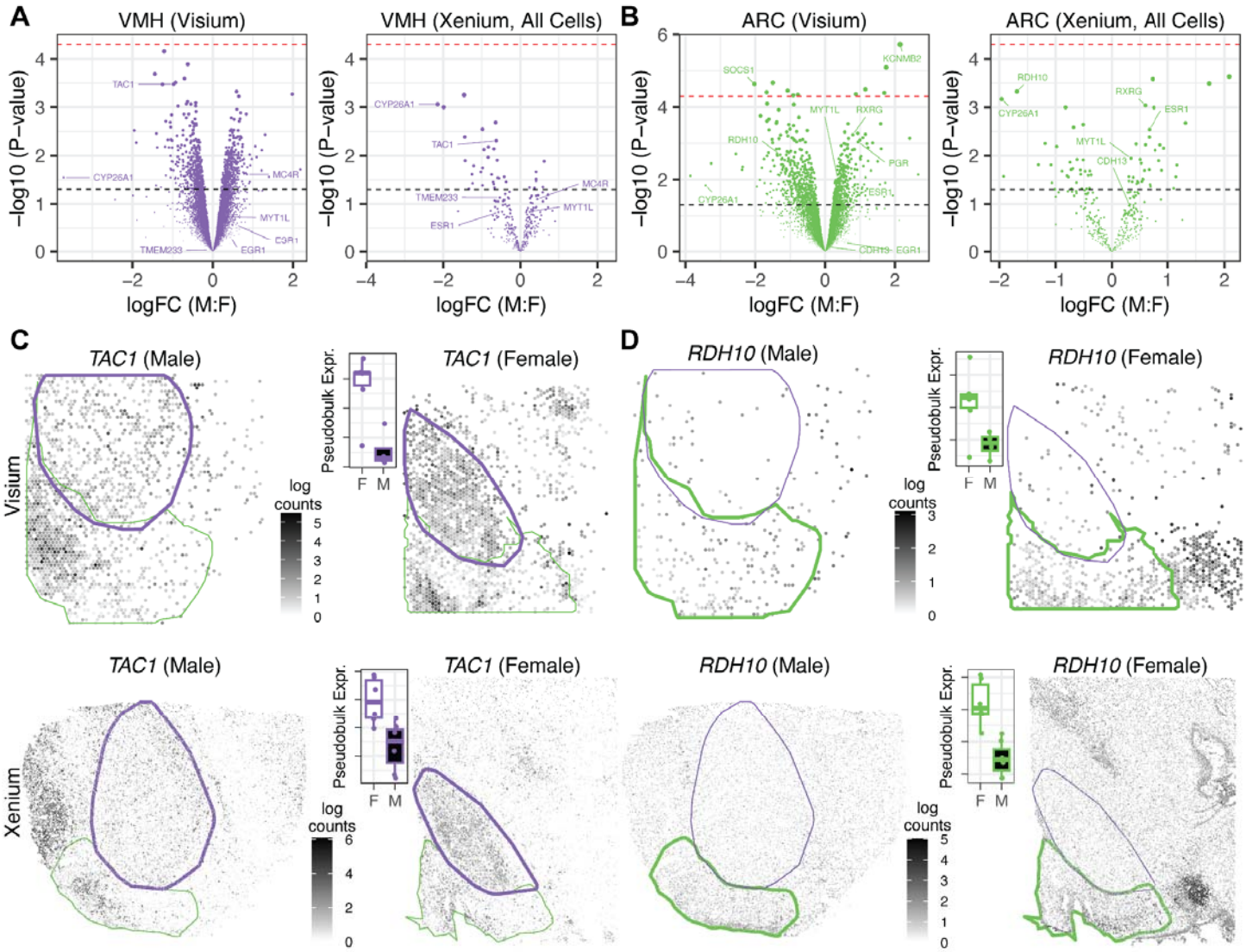

Figure 4. Sex-differential expression (sex-DE) at the domain level in the human VMH and ARC. A) Volcano plots illustrating autosomal sex-DE within vVMH and xVMH (all cells of all clusters within xVMH). The black line indicates nominal $p<0.05$; the red line indicates FDR<0.05. **B)** Volcano plot illustrating sex-DE within vARC or xARC (all cells of all clusters within xARC). **C)** Spotplots of Visium (upper row) and Xenium (bottom row) samples from one representative male donor (Br6197) and female donor (Br6588), illustrating female-preferential expression of *TAC1* in VMH (thicker domain outline). Boxplot insets show pseudobulk domain-level expression of *TAC1* from each Visium sample (top) or Xenium sample (bottom), stratified by sex. **D)** Visium (upper row) and Xenium (bottom row) samples from the same representative donors as in C, illustrating female-preferential expression of *RDH10* in ARC (thicker domain outline). Boxplot insets show pseudobulk domain-level expression of *RDH10* in each Visium sample (top) or Xenium sample (bottom), stratified by sex.

Figure 5

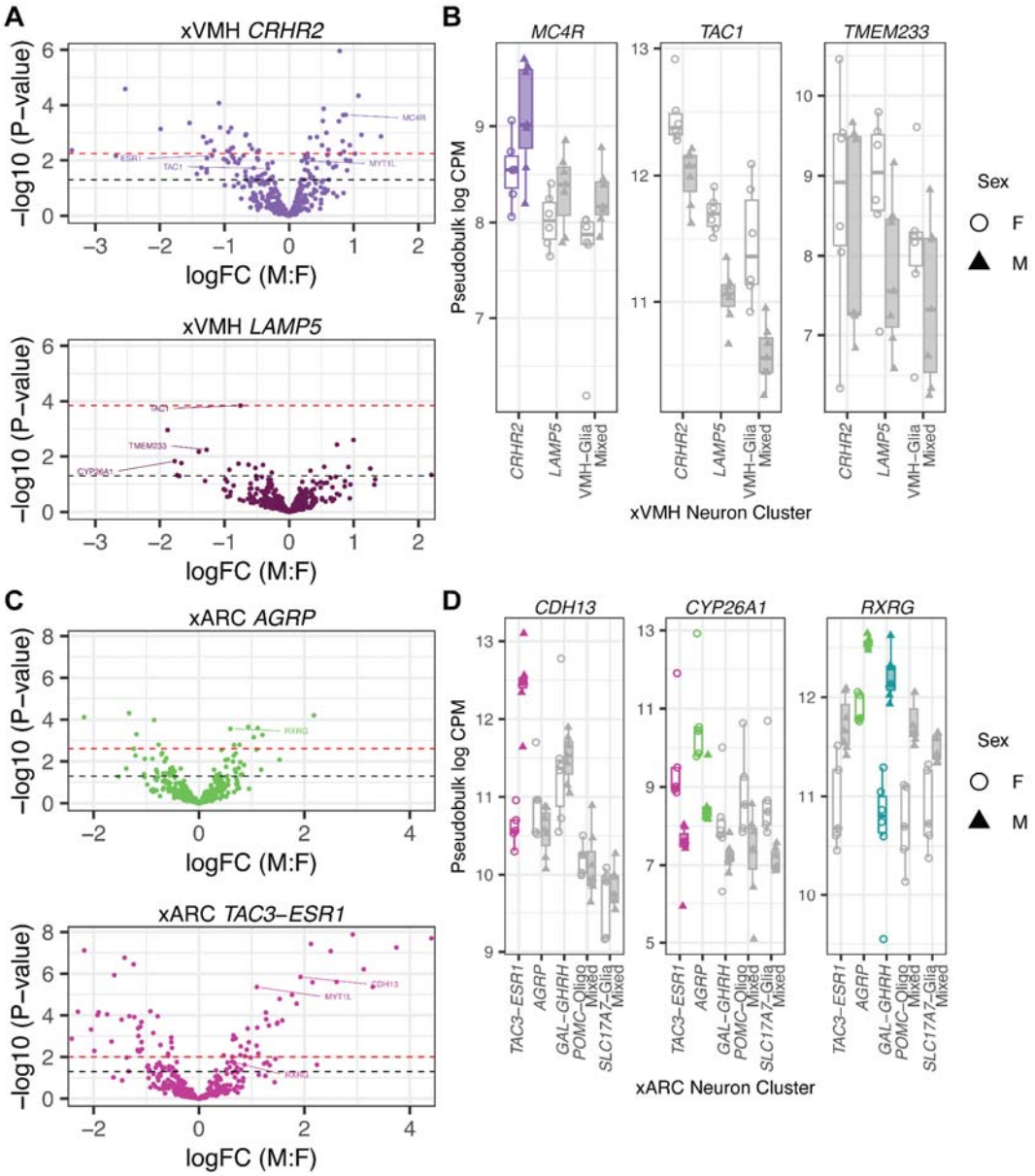

Figure 5. Sex-DE within Xenium-resolved cell types of the human xVMH and xARC. A) Volcano plots of sex-DE for selected xVMH cell clusters. **B)** Top candidate sex-DE genes from vVMH measured using Xenium. Each point represents the pseudobulk intra-xVMH gene expression for a cluster in one sample (color legend shared with panel A), stratified by sex. Sex-DE was significant (FDR<0.05) in clusters shown in color. **C)** Volcano plots of sex-DE for selected xARC cell clusters highlighting that sex-DE is most prominent in terms of significance and effect size in *TAC3-ESR1* cells. **D)** Top candidate sex-DEGs identified in vARC measured using Xenium. Each point represents the pseudobulk intra-xARC gene expression for a cluster in one sample (color legend shared with panel C), stratified by sex. sex-DE was significant (FDR<0.05) for clusters shown in color.

Figure 6

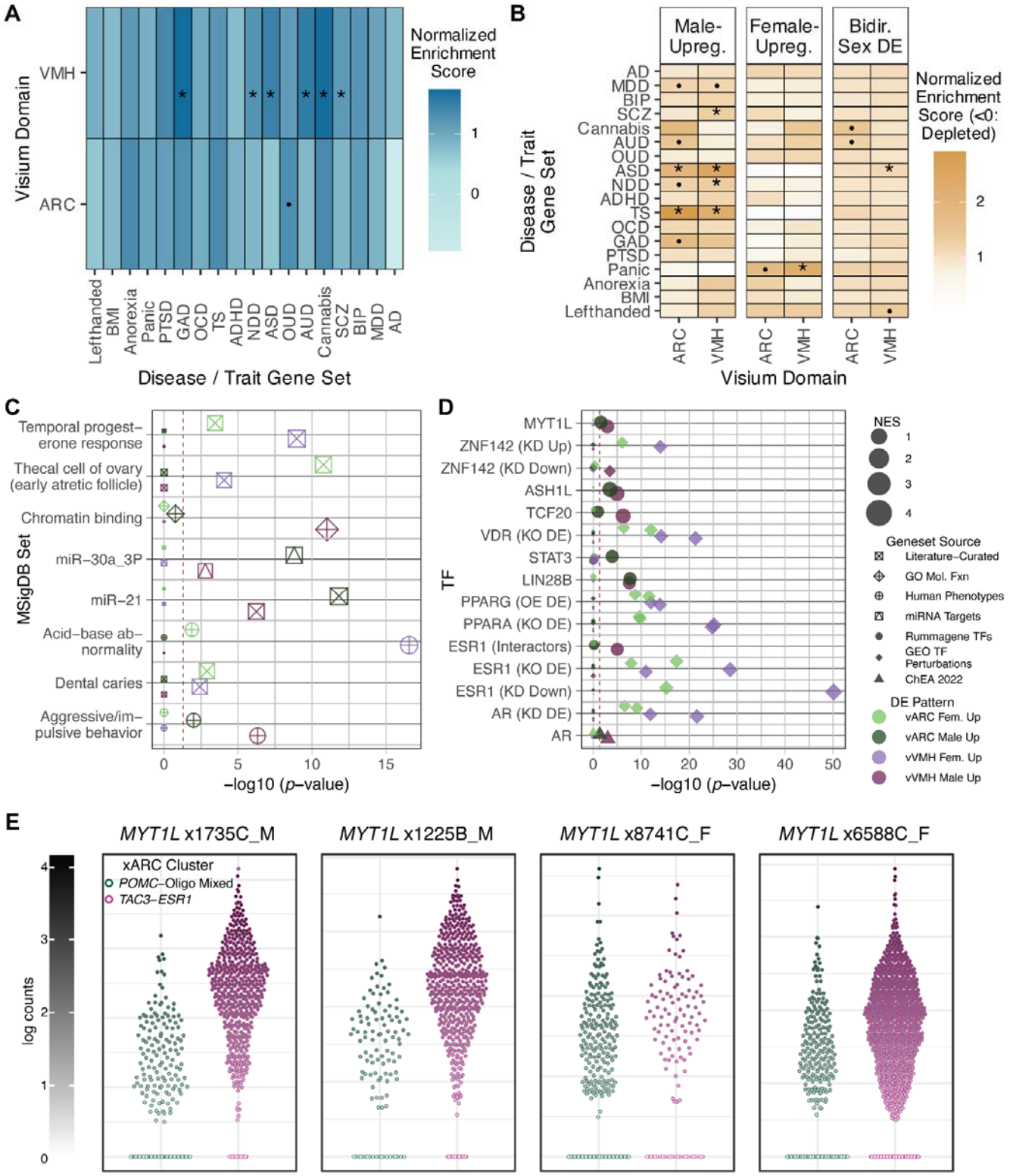
A
Visium Domain
VMH
ARC
Normalized Enrichment Score
1
0
Lefthanded
BMI
Anorexia
Panic
PTSD
GAD
OCD
TS
ADHD
NDD
ASD
OUD
AUD
Cannabis
SCZ
BIP
MDD
AD
Disease / Trait Gene Set
B
Male-Upreg.
Female-Upreg.
Bidir. Sex DE
Disease / Trait Gene Set
Normalized Enrichment Score (<0: Depleted)
2
1
ARC
VMH
Visium Domain
C
MSigDB Set
Temporal progest-erone response
Thecal cell of ovary (early atretic follicle)
Chromatin binding
miR-30a_3P
miR-21
Acid-base ab-normality
Dental caries
Aggressive/im-pulsive behavior
-log10 (p-value)
D
TF
MYT1L
ZNF142 (KD Up)
ZNF142 (KD Down)
ASH1L
TCF20
VDR (KO DE)
STAT3
LIN28B
PPARG (OE DE)
PPARA (KO DE)
ESR1 (Interactors)
ESR1 (KO DE)
ESR1 (KD Down)
AR (KD DE)
AR
-log10 (p-value)
NES
Geneset Source
Literature-Curated
GO Mol. Fxn
Human Phenotypes
miRNA Targets
Rummagene TFs
GEO TF Perturbations
ChEA 2022
DE Pattern
vARC Fem. Up
vARC Male Up
vVMH Fem. Up
vVMH Male Up
E
MYT1L x1735C_M
MYT1L x1225B_M
MYT1L x8741C_F
MYT1L x6588C_F
xARC Cluster
POMC-Oligo Mixed
TAC3-ESR1
log counts

Figure 6. Sex-DE patterns of vARC/vVMH are enriched in ASD/NDD risk genes, including disease-associated TFs. A) GSEA enrichment of vVMH/vARC markers, weighted by one-vs-all marker *t* statistics in neuropsychiatric disorder- and BMI-associated genes. * FDR<0.05; • uncorrected *p*<0.05. **B)** GSEA enrichment of genome-wide vVMH/vARC sex-DE, weighted by signed DE *t* statistics, in neuropsychiatric disorder- and BMI-associated genes. Tests of enrichment for patterns of male and female upregulation are shown, along with a test considering sex-DE in both directions. * FDR<0.05; • uncorrected *p*<0.05. **C)** GSEA enrichment of Visium sex-DE for ontologic and functional terms indexed in MSigDB, with points sized based on normalized enrichment score (NES). Legend shared with panel D. **D)** GSEA enrichment of Visium sex-DE for TF targets/TF interactors, with points sized based on normalized enrichment score (NES). Multiple points are illustrated for a sex-domain pair for GEO TF perturbation sets reporting multiple experimental replicates. **E)** ASD-associated TF, MYT1L, whose targets are overrepresented in genes with male-preferential vARC and vVMH expression (**D**), is also male upregulated in xARC, predominantly in the *TAC3-ESR1* population. Cell-level expression is plotted for two samples per sex (sex denoted in sample name) for xARC clusters *TAC3-ESR1* and *POMC*-Oligo Mixed). AD: Alzheimer's disease; ADHD: attention deficit-hyperactivity disorder; ASD: autism spectrum disorder; AUD: alcohol use disorder; BIP: bipolar disorder; BMI: body mass index (quantitative trait); Cannabis: cannabis use disorder; GAD: generalized anxiety disorder; MDD: major depressive disorder; NDD: neurodevelopmental disorder; OCD: obsessive-compulsive disorder; OUD: opioid use disorder; PTSD: post-traumatic stress disorder; SCZ: schizophrenia; TS: Tourette's syndrome.

Supplemental Material

Supplemental Figures

Figure S1. **Identification of VMH and ARC boundaries in Visium samples.** Left: Representative image of HYP tissue blocks containing VMH and ARC. Boxes denote the approximate portion of tissue used for Visium/Xenium assays. Middle: *BayesSpace* clustering at k=15 identifying 2 VMH and 2 ARC clusters. Right: Spotplots showing log transformed normalized expression (logcounts) of *FEZF1* (VMH marker) and *POMC* (ARC marker) per donor and sample(s). Approximate Visium (v)VMH and Visium (v)ARC domain boundaries are illustrated for reference—note that these boundaries, in the case of Visium, are for illustrative purposes only and were *not* used for any analyses.

Figure S2. **Spots removed from each Visium sample by local outlier quality control with *SpotSweeper* followed by manual filtering. A)** Filtering of local outliers for UMIs per spot and subsequently for low UMI content (under 210 UMIs/spot). **B)** Filtering of local outliers for number of unique genes per spot and subsequently for low number of unique genes/spot (under 126 genes/spot). **C)** Filtering for proportion of mitochondrial reads on the basis of local outliers, and subsequently for spots exceeding a numerical threshold (50%).

Figure S3. **UMAP of Xenium cell data annotated with major cell type clusters, sex, donor, or experimental batch; heatmap of top markers for all retained Xenium clusters. A**) UMAP illustrating Xenium cell type clusters grouped by the overarching cell type (e.g., astrocyte) and/or inferred domain of origin (e.g., periventricular nucleus). **B)** Same UMAP dimensions as in A, colored by brain donor. **C**) Same UMAP dimensions as in A, colored by experimental batch. **D**) Same UMAP dimensions as in A, colored by donor sex. **E**) Heatmap displaying average log counts top cluster marker genes for each of the 33 Xenium clusters retained in downstream analyses. All UMAPs illustrate dimensionality reduction following normalization of non-log adjusted count data, which was used as input for Banksy clustering (as its authors suggest[89]).

Figure S4. **Quality control metrics for the final Visium dataset.** Number of unique genes per spot, UMIs per spot, and % mitochondrial reads per spot are shown in columns. **A)** Samplewise medians for each metric. **B)** Spot-level values per sample. **C)** Spot values per sample only among spots labeled xVMH. **D)** Spot values per sample only among spots labeled ARC. Each violin in B-D shows a dotted line to demarcate the sample median, and solid lines to demarcate 25th and 75th %iles. All plots reflect the filtered dataset as used in downstream analyses.

Figure S5. **Characterization of VMH and ARC clusters and identification of marker genes with BayesSpace at *k*=15**. **A)** Spotplots of individual VMH, ARC, and optic tract (OT) white matter (WM) clusters at *k*=15 for all Visium samples. **B)** Heatmap of top marker genes for VMH, ARC, and WM clusters using one vs. all marker gene analyses. **C)** Heatmap of top marker genes using pairwise comparisons between VMH clusters and ARC clusters.

Figure S6. **Gene set enrichment analysis (GSEA) of vARC- and vVMH-enriched genes.** GSEA enrichment of vARC/vVMH enriched genes (vs. all other tissue spots) for ontologic and functional terms indexed in MSigDB, with points sized based on normalized enrichment score (NES). Full results of this analysis at *k*=15 (with vVMH and vARC as well as their individual clusters), *k*=20, and *k*=31 are available in **Data S3**.

Figure S7. **Quality control metrics from the final Xenium dataset.** Number of transcripts detected per cell, segmented cell area, segmented nucleus area, and number of unique genes per cell are shown for various

subsets of the Xenium data. Across all 41 cell clusters initially resulting from *Banksy* analysis, shown are **A)** samplewise medians and **B)** the distributions of these cell-level metrics within samples. After excluding eight sample/donor-specific *Banksy* clusters, shown are **C)** samplewise medians and **D)** the distributions of these cell-level metrics within samples. **E)** Violin plots of these metrics are shown for each xARC neuron cluster. Clusters are named along the top and metrics are listed along the right margin. **F)** Violin plots of these metrics are shown for each xVMH neuron cluster. Clusters are named along the top and metrics are listed along the right margin.

Figure S8. **Anatomically consistent VMH and ARC definitions across transcriptomic platforms.** Visium and Xenium samples are shown with spots or cells, respectively, colored according to their domain label. Left column displays Visium samples, and the right column Xenium samples. Each row shows all samples from a given donor (aspect ratios are not to scale). Sample nomenclature denotes "v" for Visium or "x" for Xenium, followed by donor ID (BrXXXX), "A" through "D" for tissue replicates, and "_M" for male or "_F" for female. Tissue samples from donor Br8667 (v8667A_F / x8667B_F) contained limited ARC. Aspect ratios are not to scale between platforms or among Xenium samples; several Xenium samples cover a larger dorsal-ventral span than is accommodated by Visium.

Figure S9. **BayesSpace clustering of Visium data at *k*=20. A)** Spotplots of individual VMH, ARC, and optic tract (OT) white matter (WM) clusters at *k*=20 for all Visium samples. **B)** Heatmap of top marker genes for VMH, ARC, and WM clusters using one vs. all marker gene analyses. **C)** Heatmap of top marker genes using pairwise comparisons between VMH clusters and ARC clusters.

Figure S10. **BayesSpace clustering of Visium data at *k*=31. A)** Spotplots of individual VMH, ARC, and optic tract (OT) white matter (WM) clusters at *k*=31 for all Visium samples. **B)** Heatmap of top marker genes for VMH, ARC, and WM clusters using one vs. all marker gene analyses. **C)** Heatmap of top marker genes when comparing each VMH cluster to all other VMH clusters combined, or each ARC cluster to all other ARC clusters combined.

Figure S11. **Spatially variable genes (SVGs) within vVMH: *SST*.** Spotplots of *SST* in vVMH for each sample. *SST* is a within-vVMH SVG that appears ventrally biased. Sample titles indicate whether spatial variation was significant. Expression is only shown for spots labeled as vVMH for readability.

Figure S12. **Spatially variable genes (SVGs) within vVMH: *PGR.*** Spotplots of *PGR* in vVMH for each sample. *PGR* is a within-vVMH SVG that appears ventrally to ventro-medially biased. Sample titles indicate whether spatial variation was significant. Expression is only shown for spots labeled as vVMH for readability.

Figure S13. **Spatially variable genes (SVGs) within vARC: *GAL***. Spotplots of *GAL* in vARC for each sample. *GAL* is a within-vARC SVG that appears biased to lateral ARC (along long medial edge of VMH) and center of ARC (ventral to VMH). Sample titles indicate whether spatial variation was significant. Expression is only shown for spots labeled as vARC for readability.

Figure S14. **Spatially variable genes (SVGs) within vARC: *NMU***. Spotplots of *NMU* in vARC for each sample. *NMU* is a within-vARC SVG with ventral localization in the center portion of ARC. Sample titles indicate whether spatial variation was significant. Expression is only shown for spots labeled as vARC for readability.

Figure S15. **Spatially variable genes (SVGs) within vARC: *IGFBP3*.** Spotplots of *IGFBP3* in vARC for each sample. *IGFBP3* is a within-vARC SVG that appears ventromedially biased. Sample titles indicate whether spatial variation was significant. Expression is only shown for spots labeled as vARC for readability.

Figure S16. **Spatially variable genes (SVGs) within vARC: *TAC3*.** Spotplots of *TAC3* in vARC for each sample. *TAC3* is a within-vARC SVG (ventral localization in center ARC, similar to *NMU*), for each sample. Sample titles indicate whether spatial variation was significant. Expression is only shown for spots labeled as vARC for readability.

Figure S17. **Spatially variable genes (SVGs) within vARC: *GHRH*.** Spotplots of *GHRH* in vARC for each sample. *GHRH* is a within-vARC SVG that appears ventrally biased. Sample titles indicate whether spatial variation was significant. Expression is only shown for spots labeled as vARC for readability.

Figure S18. **Spatially variable genes (SVGs) within vARC: *AGRP*.** Spotplots of *AGRP* in vARC for each sample. *AGRP* is a within-vARC SVG that shows a dorsal-to-ventral gradient. Sample titles indicate whether spatial variation was significant. Expression is only shown for spots labeled as vARC for readability.

Figure S19. **Spatially variable genes (SVGs) within vARC: *DIRAS3*.** Spotplots of *DIRAS3* in vARC for each sample. *DIRAS3* is a within-vARC SVG that shows a central band of expression in ARC with exclusion from lateralmost area. Sample titles indicate whether spatial variation was significant. Expression is only shown for spots labeled as vARC for readability.

Figure S20. **Marker gene heatmaps of selected xVMH and xARC cell clusters. A)** Top gene markers of xVMH, xARC, and white matter (WM) cell clusters as identified using one-vs.-all analysis across all 33 retained Xenium clusters. **B)** Top gene markers for each xVMH or xARC neuron cluster (constrained to cells with those labels falling within their domain boundaries), compared only to other xVMH or xARC neuron clusters, respectively.

Figure S21. **Spatial organization of xVMH cell clusters plotted within the xVMH domain for all samples.** For each sample, plots illustrate the location of each cell from *CRHR2*, *LAMP5*, xVMH-Glia Mixed, and xVMH Lateral Border clusters. Only cells within the xVMH domain area are plotted. Samples from donor Br8741 contained a sample-specific VMH cluster (not shown) overlapping those displayed; such cells were excluded from all analyses.

Figure S22. **Spatial organization of xARC cell clusters plotted within the xARC domain for all samples.** For each sample, plots illustrate the location of each cell from *AGRP*, *GAL-TRH-GHRH*, *POMC*-Oligo Mixed, *SLC17A7*-Glia Mixed, and *TAC3-ESR1* clusters. Only cells within the xARC domain area are plotted. Note that x8667B_F contains very little ARC tissue (similar to v8667A_F), and no *TAC3-ESR1* cells.

Figure S23. **VMH-ARC smFISH of *ESR1*, and *TAC3*, and *KISS1*. A**) Top row illustrates smFISH cells from two Xenium-adjacent serial tissue sections from donor Br1225 that were programmatically defined (*Methods*) as positive for the indicated gene (*ESR1*, *TAC3*, or *KISS1*). Boxes indicate the approximate tissue area shown below from raw fluorescence confocal images in the ARC region for the same gene (and DAPI). Scale bars 100µm. **B**) smFISH cells programmatically defined as triple-positive for *ESR1*, *TAC3*, and *KISS1*. **C)** Proportion of segmented cells in the ARC region of interest (*Methods*) across both tissue sections expressing ≥1 of *ESR1*, *TAC3*, or *KISS1*.

Figure S24. **Expression of *CYP26A1* in each xARC neuron population**. Pseudobulk expression of *CYP26A1* in each xARC neuron cluster. Each point represents the pseudobulk expression for a given cluster in one sample.

Figure S25. ***LAMP5* is expressed in *SLC17A6*+ excitatory neurons of VMH.** Representative confocal image showing expression of *GAD2*, *LAMP5*, *NR5A1*, and *SLC17A6* in the VMH by smFISH on an adjacent tissue from Br1735**.** Each gene is shown individually overlaid with nuclear DAPI staining along with a merge of all channels. Arrows indicate example cells co-expressing *NR5A1*, *LAMP5*, and *SLC17A6,* but negative for *GAD2*; arrowheads indicate example *GAD2*-expressing cells. Lipofuscin autofluorescence (without DAPI) is shown in the bottom left image. Scale bar: 100µm.

Figure S26. **Shared patterns of sex-DE in vVMH and vARC; reproducibility of sex-DE in ARC and VMH across platforms. A)** Visium domain-domain correlations among sex-DE *t* statistics or **B)** logFCs. **C)** *p*-value densities for sex-DE analysis considering the union of domain sex-DEGs as called using an uncorrected *p* value threshold of 0.05 or **D)** 0.01. **E-F)** Conditional *p* value density plots for (*i.e.*, vVMH *p* values for genes under a given threshold, plotted for the same genes in vARC, or vice versa). **G)** Visium-Xenium correlation of domain-level sex-DE analyses among all shared genes or **H)** Genes earmarked for sex-DE follow-up with Xenium during panel design based on preliminary Visium sex-DE analyses.

Figure S27. **Marker and sex-DE correlations between adult mouse ARC/VMH single cell RNA-seq populations and human vARC/vVMH domains and xARC/xVMH neuron populations.** Correlations are shown for varying numbers of top marker genes or sex-DEGs. **A)** Mouse subclasses from the respective domains were collapsed into singular “VMH” or “ARC” populations for Visium (i.e., domain-level) comparisons. Marker and sex DE correlations with the collective mouse subclasses are plotted together for each Visium domain. **B)** Each Xenium ARC neuron cluster was tested against each mouse subclass from ARC in terms of markers (left plot of each a pair) or sex DE (right plot of each pair). **C)** Each Xenium VMH neuron cluster was tested against each mouse subclass from VMH in terms of markers (left plot of each a pair) or sex DE (right plot of each pair).

Figure S28. **Proportion of Visium spots and Xenium xVMH/xARC neurons expressing sex hormone receptors (*AR*, *ESR1*, *ESR2*, *PGR*) by domain or Xenium cluster-within-domain. A)** Proportion of spots vVMH/vARC spots (Visium) or xVMH- or xARC-specific neurons within their respective domains (Xenium) in which at least one count of a given gene (color-coded) was detected. *PGR* and *ESR2* were not assayed with Xenium. Note that *ESR1* and *AR* were both detected in a ~2-fold greater proportion of Xenium cells of a domain relative to Visium spots of that domain; this indicates improved detection of genes with generally low expression (*e.g.*, those encoding TFs) via imaging-based transcriptomics. **B)** Each plot shows the proportion of cells from a given Xenium cluster with at least one count of *AR* or *ESR1*. Only cells within the Xenium domain (*x*-axis) were considered.

Figure S29. **Schematic overview of sex DE analysis procedure for Xenium clusters.** Sex-DE analysis was performed at the level of Xenium clusters by first subsetting to all cells within the boundaries of the xVMH domain or the xARC domain. Then, cells of each type are tested separately for sex-DE within that domain. Thus, broadly distributed cell clusters/types (e.g., glia) are tested for sex-DE in each domain. Meanwhile, sex-DE testing of domain-specific neuronal clusters (e.g., *TAC3-ESR1* in xARC) is filtered to the cells given that label *and* found in their domain (i.e. in the expected anatomic space). This simplified schematic only shows one domain-specific cluster each for xARC and xVMH.

Figure S30. **Sex-DE volcano plots of the 5 xARC neuron clusters and the 4 xVMH neuron clusters.** Volcano plots are shown for each xARC/xVMH neuron cluster. Each cluster was analyzed only considering cells falling within their corresponding domain. Black and red dashed lines indicate cutoffs for $p<0.05$ and FDR<0.05, respectively.

Figure S31. **Potential tanycyte clusters identified by Xenium variably map to anatomically-expected zones.** Each Xenium sample is shown, with cells in clusters including possible tanycytes colorized (legend, bottom right). Example areas with the cells in anatomically-expected locations and with the cells in arrangements resembling vascular/mural/connective tissue (*i.e.*, less likely to be bonafide tanycytes) are indicated with arrows.

Supplemental Table Legends

Table S1. Experimental metadata

a. **Sheet A**: Visium sample processing information and sequencing metrics (by sample, with rotations and reflections columns included for how spatial data was manipulated).
b. **Sheet B**: Xenium sample processing information and per-sample readout statistics from the Xeniumranger 1.7-resegmented results. Mean, median, and standard deviation are presented for i. total counts per cell (including non-targeting/control probes), ii. total transcript counts per cell, and iii. number of cells. Each value is given for the Xenium data i. as initially imported into R (only the default *xeniumranger* filter of transcripts with a "quality score" of ≥ 20 having been applied), in the initially clustered data (with 41 clusters), and in the final data (only cells in the 33 non-sample-specific clusters). Finally, information on how the coordinates were rotated/reflected is provided as in Sheet A.
c. **Sheet C**: Sample demographics.

Table S2. **Highly-variable and spatially variable genes for Visium.**

a. **Sheet A**-Table of gene name and whether it was a gene in the HVG 10%ile / HVG 20%ile / nnSVG 10%ile / nnSVG 20%ile set.
b. **Sheet B**-Full output of *nnSVG* analyses per gene per sample (including features not in either feature set from sheet A).
c. **Sheet C** *nnSVG* per-sample ranks and mean ranks (taken across nominally significant samples) for all genes achieving nominal significance in at least one sample.

Table S3. Marker analyses of Visium and Xenium clusterings.

a. **Sheet A**: Inferred predominant cell population for each Visium *k*=15 cluster.
b. **Sheets B-E**: *spatialLIBD* one-vs-all ("enrichment") outputs for *k* = 15 (collapsed VMH/ARC), 15, 20, 31, respectively. VMH and ARC clusters are not collapsed in the latter three sheets.
c. **Sheet F:** Nomenclature reference for Xenium analyses. Cluster assignments from Banksy (X1, X2…X41), their generic annotations ('ARC 1', 'ARC 2', etc.), and for ARC/VMH neuron clusters, the descriptive annotations as used in illustrations ('ARC-AGRP', etc.). Donor/sample-specific clusters that were excluded in downstream analyses are not listed.
d. **Sheet G**: *spatialLIBD* one-vs-all ("enrichment") outputs across all 33 retained Xenium clusters.
e. **Sheet H**: *spatialLIBD* one-vs-all ("enrichment") outputs when only considering cells within the xARC domain and labeled as one of the 5 xARC clusters.
f. **Sheet I**: *spatialLIBD* one-vs-all ("enrichment") outputs when only considering cells within the xVMH domain and labeled as one of the 4 xVMH clusters.
g. **Sheet J**: *spatialLIBD* one-vs-all ("enrichment") outputs when only considering **nuclear counts** for cells within the xARC domain and labeled as one of the 5 xARC clusters.
h. **Sheet K**: *spatialLIBD* one-vs-all ("enrichment") outputs when only considering **nuclear counts** for cells within the xVMH domain and labeled as one of the 4 xVMH clusters.

Table S4. Aggregated disease-gene sets and TF-target gene sets.

a. **Sheet A**: Source list delineating GWASes and databases from whence genes for each diagnosis or trait were collected.

b. **Sheets B-E**: GWAS, TWAS, DisGeNET, and union gene sets per diagnosis from above.
c. **Sheet F**: FUMA gene mapping results for PGC Panic disorder (run for preparing these gene sets).
d. **Sheet G**: FUMA parameters for PGC Panic disorder gene mapping.
e. **Sheet H:** Aggregated TF-target pairs from all databases queried.
f. **Sheet I**: Curated primary cilia gene sets obtained from[95].

Table S5. **Genes included on Xenium assay.** For custom genes in the assay, notes taken for gene prioritization (based on examinations of preliminary Visium analyses) are included.

Table S6. **Pairwise marker analyses between k=15, 20, or 31 VMH clusters and between k=15, 20, or 31 ARC clusters.** *scran* pairwise marker analyses between individual VMH or ARC clusters for *k*=15, 20, and 31. Sheet name denotes the BayesSpace cluster being tested for marker status and comparator cluster. (VMH1/VMH2 and ARC1/ARC2 are the VMH/ARC clusters from k=15 clustering). Note that these metrics can be interpreted bidirectionally as each comparison is between pairs of clusters; that is, for example, that a gene highly enriched in the tested cluster relative to the comparator is necessarily highly depleted in the comparator cluster compared to the tested cluster.

Table S7. Results of *nnSVG* analysis performed only on Visium spots classified as ARC or only classified as VMH.
a. **Sheet A**-Full output of *nnSVG* analyses of ARC, per gene per sample.
b. **Sheet B**-Full output of *nnSVG* analyses of VMH, per gene per sample.
c. **Sheet C**-ARC-only *nnSVG* per-sample ranks and mean ranks (taken across nominally significant samples) for all genes achieving nominal significance in at least one sample in this analysis.
d. **Sheet D**-VMH-only *nnSVG* per-sample ranks and mean ranks (taken across nominally significant samples) for all genes achieving nominal significance in at least one sample in this analysis.

Table S8. Visium marker gene spatial registration results in mouse: *scCoco* results for all mouse hypothalamic areas mapped from Allen ISH atlas; registration to mouse HYP subclasses from single-cell RNA-seq[86].
a. **Sheets A-D**: scCoco domain-to-mouse brain atlas HYP region scores using 50-150 top human markers per domain from the k=15 (collapsed VMH, ARC), k=15 using individual domains, k=20 domains, k=31 domains respectively.
b. **Sheet E:** Spearman correlations between (i.e., spatial registration of) mouse HYP cell “subclasses” and each Visium cluster. Mouse VMH and ARC (“ARH”) subclasses were aggregated into single domains for comparability to Visium domains. Results are included for Visium clusterings at k=15 (with collapsed VMH, ARC clusters), k=15 using individual clusters, k=20, or k=31. Correlations were tested using variable numbers of top *n* marker genes (ranked by *t*-statistic) per human cluster. Only genes with 1-1 mouse orthologs and present in both datasets were used.

Table S9. Visium pseudobulk sex-DE results for k=15, k=15 (collapsed), k=20, k=31, and domain sex-DE pattern enrichments in disease-gene sets.
a. **Sheets A-D:** Pseudobulked sex-DE results for Visium cluster assignments using each k-value, including single k=15 domains and the collapsed vVMH/vARC domains as primarily reported in the main text.
b. **Sheet E:** Collated results from all Visium sex-DE GSEA analyses for BMI and neuropsychiatric disorders. All BayesSpace classifications (k=15 (collapsed VMH/ARC), k=15, k=20, and k=31) and disease-gene sets/subsets (GWAS publication genes only, TWAS genes only, DisGeNET (“dgn”) genes only, union of genes from these sources) are presented in a single table.

Table S10. Spearman correlations of sex-DE in mouse HYP subclasses identified in[86] with that of Visium domains/clusters or Xenium clusters (within xARC domain boundaries or within xVMH boundaries).

a. **Sheet A:** sex-DE correlations between Visium domains and mouse subclasses using only genes with 1-1 mouse orthologs present in both datasets, across varying numbers of top human DEGs (by absolute value of *t*-statistic for a given Visium domain/cluster). Spearman correlations between sex-DE in mouse HYP cell "subclasses" and each Visium cluster. Mouse VMH and ARC subclasses were aggregated into single domains for comparability to Visium domains. Results are included for Visium clusterings at k=15 (with collapsed VMH, ARC clusters), k=15 using individual clusters, k=20, or k=31. Correlations were tested using variable numbers of top *n* sex-DEGs (ranked by |*t*-statistic|) per human cluster. Only autosomal genes with 1-1 mouse orthologs and present in both datasets were used.
b. **Sheet B:** sex-DE correlations between Xenium clusters and mouse subclasses. Spearman correlations between sex-DE in mouse HYP cell "subclasses" and each Xenium cluster, only considering cells from each cluster within the xARC or within the xVMH domain (indicated by the column "*Xenium Domain Considered*"). Each mouse VMH and ARC subclass was analyzed independently for comparability to Xenium. Only autosomal genes with 1-1 mouse orthologs and present in both datasets were used.

Table S11. GSEA results: Visium domain-enriched gene overlap with disease/trait-associated genes. Collated results for all domains from all BayesSpace classifications for source data-subsetted disease/trait-gene sets (**A**) union gene set of all sources, **B**) GWAS publication genes only, **C**) TWAS genes only, **D**) DisGeNET genes only) for each diagnosis/trait using BayesSpace k=15 (collapsed VMH/ARC), k=15, k=20, and k=31 clusterings.

Table S12. **Xenium pseudobulk sex-DE results for each cluster, considering only cells within boundaries of xARC or of xVMH; sex-DE when pseudobulking all cells in the xARC or xVMH domain; confirmatory xenium sex-DE analyses.** Positive logFC indicates male expression greater than female.

c. **Sheet A:** Pseudobulk sex-DE results by Xenium cluster only considering cells within xARC.
d. **Sheet B:** Pseudobulk sex-DE results by Xenium cluster only considering cells within xVMH.
e. **Sheet C**: Pseudobulk sex-DE results at the level of Xenium domains (all cells within xARC or all cells within xVMH).
f. **Sheet D**: Pseudobulk sex-DE results by Xenium cluster, only considering transcripts within segmented nuclei in xARC.
g. **Sheet E**: Pseudobulk sex-DE results by Xenium cluster, only considering transcripts within segmented nuclei in xVMH.
h. **Sheet F**: Pseudobulk sex-DE results by Xenium cluster, only considering cells within a shrunken xARC domain (so as to exclude the domain boundary, where xVMH-xARC populations would likely comingle).
i. **Sheet G**: Pseudobulk sex-DE results by Xenium cluster only considering cells within a shrunken xVMH domain (so as to exclude the domain boundary, where xVMH-xARC populations would most likely comingle).
j. **Sheet H**: Correlation of sex-DE *t*-statistics between all pairs of Xenium xARC/xVMH cluster analyses (primary domain-level analysis, primary clusterwise-within-a-domain analysis, nuclear-counts, shrunken domain counts) amongst themselves and Visium domain or *k*=15 VMH/ARC clusters. The analysis and cluster of the first comparator are listed in "*Cluster/Domain 1 Analysis*" and "*Cluster/Domain 1*", respectively. Individual Xenium clusters were only compared to themselves and to Visium clusters/domains.

Table S13. **Each gene assayed on Visium and/or Xenium, whether it is a TF tested in the TF-target GSEA analyses or in 1 or more disease-gene sets, and information about its expression and sex-DE from Visium and Xenium**. Each gene measured in Visium and/or Xenium is listed. Columns indicate whether

the gene was assayed with Xenium; which of the disease(s) gene lists the gene appears in; whether the gene is expressed in vVMH/vARC overall and whether it is expressed in each Visium domain on a single-sex basis; whether the gene is nominally sex-DE or significantly sex-DE in each Visium domain; each Xenium cluster within xARC wherein the gene was sex-DE, if any; and each cluster within xVMH wherein the gene was sex-DE, if any.

Supplemental Data Files

Data File S1. **Visium PCA and UMAP, before and after HARMONY dimensionality reduction.** Plots depicting PCs 1 vs 2 and 1 vs 3 or UMAP with dimensionality reductions on the 10%ile nnSVG feature set, without and with HARMONY batch correction. Plot legends indicate the variable by which points (Visium spots) have been colored, including BayesSpace cluster assignments (ascertained using HARMONY dimensionality reduction on the 10%ile SVG feature set), demographic characteristics, and technical variables.

Data File S2. GSEA results for Visium domain/cluster-enriched genes in MSigDb. File is provided in an R data format (*.RDS*). Composed of an R list with names corresponding to which domain assignments were tested. Each entry of the list holds a dataframe of collated, unabridged results produced by *fgsea::fgseaMultilevel*, with a column added to mark whether or not that row was one of the 'collapsed' results following independent pathway analysis with *fgsea::collapsePathways(). fgsea* output tables for each Visium cluster in a BayesSpace run were collapsed together into one table, with a column added to indicate the analyzed Visium cluster. For a cluster-by-BayesSpace run-by-gene set combination, the gene set name was checked against the corresponding output from *collapsePathways()* for labeling as in the 'collapsed' gene set results. Additional information on columns output by *fgseaMultilevel* can be found in the package's documentation.

Data File S3. GSEA results for Visium domain/cluster-enriched genes in aggregated TF-target gene sets. File is provided in an R data format (*.RDS*). Contains a list, named according to the BayesSpace *k* value and cluster tested (*k*=15, 20, 31, or *k*=15 with ARC and VMH collapsed into singular domains). Each entry of the list holds a dataframe of collated results produced by *fgsea::fgseaMultilevel*, with a column added to mark whether or not that row was one of the 'collapsed' results following independent pathway analysis with *fgsea::collapsePathways().* Only gene sets with a positive enrichment score (*i.e.*, overrepresented in the tested Visium genes) are presented. Each table was collated as described for File S2.

Data File S4. Results of sex-DE analysis by cell subclass in the mouse single-cell HYP data from Yao et al[86]. DE result tables for each subclass or aggregate subclass (for Visium comparisons–relabeling all VMH subclasses to a single VMH cell type and likewise for ARC subclasses) are collated here with a column added to indicate the subclass/aggregate subclass analyzed. Further information on how these data were obtained is in *Methods* section 3.7.11.

Data File S5. GSEA results for Visium sex-DE pattern enrichment in MSigDb. File is provided in an R data format (*.RDS*). Contains a list named as described for file S2, but with two result tables produced by *fgsea::fgseaMultilevel* for every cluster tested: one "signed", indicating a GSEA run including direction of DE to identify enrichments among male-upregulated or female-upregulated genes, and one "unsigned", indicating a GSEA run with the absolute values of DE statistics to identify terms enriched for sex effects bidirectionally. For signed test results, the DE effect direction explicitly tested (enrichment in male-upregulated genes or in female-upregulated genes) is indicated, which corresponds to a positive or negative value, respectively, for enrichment score and normalized enrichment score.

Data File S6. GSEA results for Visium sex-DE pattern enrichment in TF-target gene sets. File is provided in an R data format (*.RDS*). Contains a list named as described for file S5. Each entry of the list holds a dataframe of collated, unabridged results produced by *fgsea::fgseaMultilevel*, with columns added to indicate whether the TF qualified as expressed in male and/or female samples (at least three samples from a given sex with the TF detected in ≥1% of spots in a cluster, and the TF detected in at least 2.5% of spots overall for that cluster and sex). For signed test results, the DE effect direction explicitly tested (enrichment in male-upregulated genes or in female-upregulated genes) is indicated, which corresponds to a positive or negative value, respectively, for enrichment score and normalized enrichment score.

Bibliography

1. Mei, L., Osakada, T. & Lin, D. Hypothalamic control of innate social behaviors. *Science* **382**, 399–404 (2023).
2. Fong, H., Zheng, J. & Kurrasch, D. The structural and functional complexity of the integrative hypothalamus. *Science* **382**, 388–394 (2023).
3. Kim, D.-W. *et al.* Multimodal analysis of cell types in a hypothalamic node controlling social behavior. *Cell* **179**, 713-728.e17 (2019).
4. Calvigioni, D. *et al.* Esr1+ hypothalamic-habenula neurons shape aversive states. *Nat. Neurosci.* **26**, 1245–1255 (2023).
5. Qiu, J. *et al.* Estrogenic-dependent glutamatergic neurotransmission from kisspeptin neurons governs feeding circuits in females. *eLife* **7**, (2018).
6. Herber, C. B. *et al.* Estrogen signaling in arcuate Kiss1 neurons suppresses a sex-dependent female circuit promoting dense strong bones. *Nat. Commun.* **10**, 163 (2019).
7. Unger, T. J., Calderon, G. A., Bradley, L. C., Sena-Esteves, M. & Rios, M. Selective deletion of Bdnf in the ventromedial and dorsomedial hypothalamus of adult mice results in hyperphagic behavior and obesity. *J. Neurosci.* **27**, 14265–14274 (2007).
8. van Veen, J. E. *et al.* Hypothalamic estrogen receptor alpha establishes a sexually dimorphic regulatory node of energy expenditure. *Nat. Metab.* **2**, 351–363 (2020).
9. Ramos-Pittol, J. M. *et al.* Dax1 modulates ERα-dependent hypothalamic estrogen sensing in female mice. *Nat. Commun.* **14**, 3076 (2023).
10. Cortes, L. R., Sturgeon, H. & Forger, N. G. Sexual differentiation of estrogen receptor alpha subpopulations in the ventromedial nucleus of the hypothalamus. *Horm. Behav.* **151**, 105348 (2023).
11. MacKay, H. *et al.* Sex-specific epigenetic development in the mouse hypothalamic arcuate nucleus pinpoints human genomic regions associated with body mass index. *Sci. Adv.* **8**, eabo3991 (2022).
12. Affinati, A. H. *et al.* Cross-species analysis defines the conservation of anatomically segregated VMH neuron populations. *eLife* **10**, (2021).
13. Satoh, N. *et al.* Sympathetic activation of leptin via the ventromedial hypothalamus: leptin-induced increase in catecholamine secretion. *Diabetes* **48**, 1787–1793 (1999).

14. Vander Tuig, J. G., Knehans, A. W. & Romsos, D. R. Reduced sympathetic nervous system activity in rats with ventromedial hypothalamic lesions. *Life Sci.* **30**, 913–920 (1982).

15. Lin, W. *et al.* Whole-brain mapping of histaminergic projections in mouse brain. *Proc Natl Acad Sci USA* **120**, e2216231120 (2023).

16. Chalmers, D. T., Lovenberg, T. W. & De Souza, E. B. Localization of novel corticotropin-releasing factor receptor (CRF2) mRNA expression to specific subcortical nuclei in rat brain: comparison with CRF1 receptor mRNA expression. *J. Neurosci.* **15**, 6340–6350 (1995).

17. Viskaitis, P. *et al.* Modulation of SF1 neuron activity coordinately regulates both feeding behavior and associated emotional states. *Cell Rep.* **21**, 3559–3572 (2017).

18. Iigaya, K., Okazaki, S., Minoura, Y. & Onimaru, H. Interaction between novel oscillation within the ventromedial hypothalamus and the sympathetic nervous system. *Neuroscience* **343**, 213–221 (2017).

19. Lewis, J. E. *et al.* Relaxin/insulin-like family peptide receptor 4 (Rxfp4) expressing hypothalamic neurons modulate food intake and preference in mice. *Mol. Metab.* **66**, 101604 (2022).

20. Liu, X., Zhu, Z., Kalyani, M., Janik, J. M. & Shi, H. Effects of energy status and diet on Bdnf expression in the ventromedial hypothalamus of male and female rats. *Physiol. Behav.* **130**, 99–107 (2014).

21. Yang, T. *et al.* Social Control of Hypothalamus-Mediated Male Aggression. *Neuron* **95**, 955-970.e4 (2017).

22. Yang, C. F. *et al.* Sexually dimorphic neurons in the ventromedial hypothalamus govern mating in both sexes and aggression in males. *Cell* **153**, 896–909 (2013).

23. Lee, H. *et al.* Scalable control of mounting and attack by Esr1+ neurons in the ventromedial hypothalamus. *Nature* **509**, 627–632 (2014).

24. Lee, C. H. *et al.* Primary cilia mediate early life programming of adiposity through lysosomal regulation in the developing mouse hypothalamus. *Nat. Commun.* **11**, 5772 (2020).

25. Osterstock, G. *et al.* Ghrelin stimulation of growth hormone-releasing hormone neurons is direct in the arcuate nucleus. *PLoS ONE* **5**, e9159 (2010).

26. Wu, Q., Clark, M. S. & Palmiter, R. D. Deciphering a neuronal circuit that mediates appetite. *Nature* **483**, 594–597 (2012).

27. Goto, T., Hagihara, M. & Miyamichi, K. Dynamics of pulsatile activities of arcuate kisspeptin neurons in

aging female mice. *eLife* **12**, (2023).

28. Li, X.-F. *et al.* Kisspeptin signalling in the hypothalamic arcuate nucleus regulates GnRH pulse generator frequency in the rat. *PLoS ONE* **4**, e8334 (2009).
29. Mansano, N. da S. *et al.* Fasting modulates gabaergic synaptic transmission to arcuate kisspeptin neurons in female mice. *Endocrinology* **164**, (2023).
30. Pahl, M. C. *et al.* Cis-regulatory architecture of human ESC-derived hypothalamic neuron differentiation aids in variant-to-gene mapping of relevant complex traits. *Nat. Commun.* **12**, 6749 (2021).
31. Velasco, E. R. *et al.* PACAP-PAC1R modulates fear extinction via the ventromedial hypothalamus. *Nat. Commun.* **13**, 4374 (2022).
32. Hildebrandt, M. G., Stage, K. B., Kragh-Soerensen, P. & Danish University Antidepressant Group. Gender differences in severity, symptomatology and distribution of melancholia in major depression. *Psychopathology* **36**, 204–212 (2003).
33. Singh, T. & Williams, K. Atypical depression. *Psychiatry Edgmont Pa Townsh* **3**, 33–9 (2006).
34. Gold, P. W. & Chrousos, G. P. Organization of the stress system and its dysregulation in melancholic and atypical depression: high vs low CRH/NE states. *Mol. Psychiatry* **7**, 254–275 (2002).
35. Cristancho, M. A., O'Reardon, J. P. & Thase, M. E. Atypical depression in the 21st century: diagnostic and treatment issues. *Psychiatric Times* 42–46 (2011).
36. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Psychiatry* **9**, 137–150 (2022).
37. McKibben, L. A. & Dwivedi, Y. Early life and adult stress promote sex dependent changes in hypothalamic miRNAs and environmental enrichment prevents stress-induced miRNA and gene expression changes in rats. *BMC Genomics* **22**, 701 (2021).
38. Trabzuni, D. *et al.* Widespread sex differences in gene expression and splicing in the adult human brain. *Nat. Commun.* **4**, 2771 (2013).
39. Oliva, M. *et al.* The impact of sex on gene expression across human tissues. *Science* **369**, (2020).
40. Fass, S. B. *et al.* Relationship between sex biases in gene expression and sex biases in autism and Alzheimer's disease. *medRxiv* (2023) doi:10.1101/2023.08.29.23294773.

41. GTEx Consortium *et al.* Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017).

42. Trabzuni, D. *et al.* Quality control parameters on a large dataset of regionally dissected human control brains for whole genome expression studies. *J. Neurochem.* **119**, 275–282 (2011).

43. Herb, B. R. *et al.* Single-cell genomics reveals region-specific developmental trajectories underlying neuronal diversity in the human hypothalamus. *Sci. Adv.* **9**, eadf6251 (2023).

44. Tadross, J. A. *et al.* Human HYPOMAP: A comprehensive spatio-cellular map of the human hypothalamus. *BioRxiv* (2023) doi:10.1101/2023.09.15.557967.

45. Nguyen, H. P. *et al.* Integrative single-cell characterization of hypothalamus sex-differential and obesity-associated genes and regulatory elements. *BioRxiv* (2022) doi:10.1101/2022.11.06.515311.

46. Chandra, V. *et al.* Structure of the intact PPAR-gamma-RXR- nuclear receptor complex on DNA. *Nature* **456**, 350–356 (2008).

47. Lipska, B. K. *et al.* Critical factors in gene expression in postmortem human brain: Focus on studies in schizophrenia. *Biol. Psychiatry* **60**, 650–658 (2006).

48. Maynard, K. R. *et al.* Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat. Neurosci.* **24**, 425–436 (2021).

49. Huuki-Myers, L. *et al.* Integrated single cell and unsupervised spatial transcriptomic analysis defines molecular anatomy of the human dorsolateral prefrontal cortex. *BioRxiv* (2023) doi:10.1101/2023.02.15.528722.

50. Ramnauth, A. D. *et al.* Spatiotemporal analysis of gene expression in the human dentate gyrus reveals age-associated changes in cellular maturation and neuroinflammation. *BioRxiv* (2024) doi:10.1101/2023.11.20.567883.

51. Yalcinbas, E. A. *et al.* Transcriptomic analysis of the human habenula in schizophrenia. *BioRxiv* (2024) doi:10.1101/2024.02.26.582081.

52. Maynard, K. R. *et al.* dotdotdot: an automated approach to quantify multiplex single molecule fluorescent in situ hybridization (smFISH) images in complex tissues. *Nucleic Acids Res.* **48**, e66 (2020).

53. Tippani, M. *et al.* VistoSeg: Processing utilities for high-resolution images for spatially resolved transcriptomics data. *Biol. Imaging* **3**, e23 (2023).

54. Righelli, D. *et al.* SpatialExperiment: infrastructure for spatially-resolved transcriptomics data in R using Bioconductor. *Bioinformatics* **38**, 3128–3131 (2022).

55. Zappia, L., Lun, A. & Cannoodt, R. zellkonverter: Conversion Between scRNA-seq Objects. *Bioconductor* (2020) doi:10.18129/b9.bioc.zellkonverter.

56. Pardo, B. *et al.* spatialLIBD: an R/Bioconductor package to visualize spatially-resolved transcriptomics data. *BMC Genomics* **23**, 434 (2022).

57. Guo, B., Huuki-Myers, L. A., Grant-Peters, M., Collado-Torres, L. & Hicks, S. C. escheR: unified multi-dimensional visualizations with Gestalt principles. *Bioinformatics Advances* **3**, vbad179 (2023).

58. Coombes, K. R., Brock, G., Abrams, Z. B. & Abruzzo, L. V. Polychrome□: Creating and Assessing Qualitative Palettes with Many Colors. *J. Stat. Softw.* **90**, (2019).

59. Totty, M., Hicks, S. C. & Guo, B. SpotSweeper: spatially-aware quality control for spatial transcriptomics. *BioRxiv* (2024) doi:10.1101/2024.06.06.597765.

60. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).

61. Lun, A. T. L., McCarthy, D. J. & Marioni, J. C. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. [version 2; peer review: 3 approved, 2 approved with reservations]. *F1000Res.* **5**, 2122 (2016).

62. Weber, L. M., Saha, A., Datta, A., Hansen, K. D. & Hicks, S. C. nnSVG for the scalable identification of spatially variable genes using nearest-neighbor Gaussian processes. *Nat. Commun.* **14**, 4059 (2023).

63. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).

64. Ding, S.-L. *et al.* Comprehensive cellular-resolution atlas of the adult human brain. *J. Comp. Neurol.* **524**, 3127–3481 (2016).

65. Dyjack, N., Baker, D. N., Braverman, V., Langmead, B. & Hicks, S. C. A scalable and unbiased discordance metric with H. *Biostatistics* **25**, 188–202 (2023).

66. Zhao, E. *et al.* Spatial transcriptomics at subspot resolution with BayesSpace. *Nat. Biotechnol.* **39**, 1375–1384 (2021).

67. Haghverdi, L., Lun, A. T. L., Morgan, M. D. & Marioni, J. C. Batch effects in single-cell RNA-sequencing

data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* **36**, 421–427 (2018).

68. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102**, 15545–15550 (2005).

69. Mootha, V. K. *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).

70. Korotkevich, G. *et al.* Fast gene set enrichment analysis. *BioRxiv* (2016) doi:10.1101/060012.

71. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).

72. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).

73. Keenan, A. B. *et al.* ChEA3: transcription factor enrichment analysis by orthogonal omics integration. *Nucleic Acids Res.* **47**, W212–W224 (2019).

74. Han, H. *et al.* TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res.* **46**, D380–D386 (2018).

75. Clarke, D. J. B. *et al.* Rummagene: Mining Gene Sets from Supporting Materials of PMC Publications. *BioRxiv* (2023) doi:10.1101/2023.10.03.560783.

76. Xie, Z. *et al.* Gene Set Knowledge Discovery with Enrichr. *Curr. Protoc.* **1**, e90 (2021).

77. Lambert, S. A. *et al.* The human transcription factors. *Cell* **172**, 650–665 (2018).

78. Gamazon, E. R. *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nat. Genet.* **47**, 1091–1098 (2015).

79. Barbeira, A. N. *et al.* Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat. Commun.* **9**, 1825 (2018).

80. Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).

81. Hu, Y. *et al.* A statistical framework for cross-tissue transcriptome-wide association analysis. *Nat. Genet.* **51**, 568–576 (2019).

82. Lu, M. *et al.* TWAS Atlas: a curated knowledgebase of transcriptome-wide association studies. *Nucleic Acids Res.* **51**, D1179–D1187 (2023).

83. Piñero, J., Saüch, J., Sanz, F. & Furlong, L. I. The DisGeNET cytoscape app: Exploring and visualizing

disease genomics data. *Comput. Struct. Biotechnol. J.* **19**, 2960–2967 (2021).

84. Border, R. *et al.* No Support for Historical Candidate Gene or Candidate Gene-by-Interaction Hypotheses for Major Depression Across Multiple Large Samples. *Am. J. Psychiatry* **176**, 376–387 (2019).

85. Lein, E. S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).

86. Yao, Z. *et al.* A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *Nature* **624**, 317–332 (2023).

87. Schilder, B. orthogene. *Bioconductor* (2021) doi:10.18129/b9.bioc.orthogene.

88. Moses, L. *et al.* Voyager: exploratory single-cell genomics data analysis with geospatial statistics. *BioRxiv* (2023) doi:10.1101/2023.07.20.549945.

89. Singhal, V. *et al.* BANKSY unifies cell typing and tissue domain segmentation for scalable spatial omics data analysis. *Nat. Genet.* **56**, 431–441 (2024).

90. Peters Couto, B. Z., Robertson, N., Patrick, E. & Ghazanfar, S. MoleculeExperiment enables consistent infrastructure for molecule-resolved spatial omics data in bioconductor. *Bioinformatics* **39**, (2023).

91. Hellier, V. *et al.* Female sexual behavior in mice is controlled by kisspeptin neurons. *Nat. Commun.* **9**, 400 (2018).

92. Hanna, L., Kawalek, T. J., Beall, C. & Ellacott, K. L. J. Changes in neuronal activity across the mouse ventromedial nucleus of the hypothalamus in response to low glucose: Evaluation using an extracellular multi-electrode array approach. *J. Neuroendocrinol.* **32**, e12824 (2020).

93. Meek, T. H. *et al.* Functional identification of a neurocircuit regulating blood glucose. *Proc Natl Acad Sci USA* **113**, E2073-82 (2016).

94. Tong, Q. *et al.* Synaptic glutamate release by ventromedial hypothalamic neurons is part of the neurocircuitry that prevents hypoglycemia. *Cell Metab.* **5**, 383–393 (2007).

95. Vasquez, S. S. V., van Dam, J. & Wheway, G. An updated SYSCILIA gold standard (SCGSv2) of known ciliary genes, revealing the vast progress that has been made in the cilia research field. *Mol. Biol. Cell* **32**, br13 (2021).

96. Steuernagel, L. *et al.* HypoMap-a unified single-cell gene expression atlas of the murine hypothalamus. *Nat. Metab.* **4**, 1402–1419 (2022).

97. Zhou, X. *et al.* Deciphering the spatial-temporal transcriptional landscape of human hypothalamus development. *Cell Stem Cell* **29**, 328-343.e5 (2022).

98. Göcz, B. *et al.* Transcriptome profiling of kisspeptin neurons from the mouse arcuate nucleus reveals new mechanisms in estrogenic control of fertility. *Proc Natl Acad Sci USA* **119**, e2113749119 (2022).

99. Ransom, J., Morgan, P. J., McCaffery, P. J. & Stoney, P. N. The rhythm of retinoids in the brain. *J. Neurochem.* **129**, 366–376 (2014).

100. Herring, C. A. *et al.* Human prefrontal cortex gene regulatory dynamics from gestation to adulthood at single-cell resolution. *Cell* **185**, 4428-4447.e28 (2022).

101. BRAIN Initiative Cell Census Network (BICCN). A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature* **598**, 86–102 (2021).

102. Kruse, S. W. *et al.* Identification of COUP-TFII orphan nuclear receptor as a retinoic acid-activated receptor. *PLoS Biol.* **6**, e227 (2008).

103. Shearer, K. D. *et al.* Photoperiodic regulation of retinoic acid signaling in the hypothalamus. *J. Neurochem.* **112**, 246–257 (2010).

104. Liu, C. *et al.* The MYH9 cytoskeletal protein is a novel corepressor of androgen receptors. *Front. Oncol.* **11**, 641496 (2021).

105. Rocks, D. *et al.* Egr1 is a sex-specific regulator of neuronal chromatin, synaptic plasticity, and behaviour. *BioRxiv* (2023) doi:10.1101/2023.12.20.572697.

106. Krishna Kumar, K. *et al.* Negative allosteric modulation of the glucagon receptor by RAMP2. *Cell* **186**, 1465-1477.e18 (2023).

107. García-Juárez, M. *et al.* Apelin-13 facilitates lordosis behavior following infusions to the ventromedial hypothalamus or preoptic area in ovariectomized, estrogen-primed rats. *Neurosci. Lett.* **773**, 136518 (2022).

108. Morishita, M., Higo, S., Iwata, K. & Ishii, H. Sex and interspecies differences in ESR2-expressing cell distributions in mouse and rat brains. *Biol. Sex Differ.* **14**, 89 (2023).

109. Sjöstedt, E. *et al.* An atlas of the protein-coding genes in the human, pig, and mouse brain. *Science* **367**, (2020).

110. Zhang, L. *et al.* Progesterone receptor distribution in the human hypothalamus and its association with

suicide. *Acta Neuropathol. Commun.* **12**, 16 (2024).

111. Knoedler, J. R. *et al.* A functional cellular framework for sex and estrous cycle-dependent gene expression and behavior. *Cell* **185**, 654-671.e22 (2022).

112. Farooqi, I. S. *et al.* Clinical spectrum of obesity and mutations in the melanocortin 4 receptor gene. *N. Engl. J. Med.* **348**, 1085–1095 (2003).

113. Hug, C. *et al.* T-cadherin is a receptor for hexameric and high-molecular-weight forms of Acrp30/adiponectin. *Proc Natl Acad Sci USA* **101**, 10308–10313 (2004).

114. Kichaev, G. *et al.* Leveraging polygenic functional enrichment to improve GWAS power. *Am. J. Hum. Genet.* **104**, 65–75 (2019).

115. Chung, C.-M. *et al.* A genome-wide association study reveals a quantitative trait locus of adiponectin on CDH13 that predicts cardiometabolic outcomes. *Diabetes* **60**, 2417–2423 (2011).

116. Huang, J. *et al.* Genomics and phenomics of body mass index reveals a complex disease network. *Nat. Commun.* **13**, 7973 (2022).

117. Chen, J. *et al.* A MYT1L syndrome mouse model recapitulates patient phenotypes and reveals altered brain development due to disrupted neuronal maturation. *Neuron* **109**, 3775-3792.e14 (2021).

118. Heras, V. *et al.* Hypothalamic miR-30 regulates puberty onset via repression of the puberty-suppressing factor, Mkrn3. *PLoS Biol.* **17**, e3000532 (2019).

119. Metzger, E., Müller, J. M., Ferrari, S., Buettner, R. & Schüle, R. A novel inducible transactivation domain in the androgen receptor: implications for PRK in prostate cancer. *EMBO J.* **22**, 270–280 (2003).

120. Metzger, E. *et al.* Phosphorylation of histone H3 at threonine 11 establishes a novel chromatin mark for transcriptional regulation. *Nat. Cell Biol.* **10**, 53–60 (2008).

121. Satterstrom, F. K. *et al.* Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* **180**, 568-584.e23 (2020).

122. Proskorovski-Ohayon, R. *et al.* ZNF142 mutation causes sex-dependent neurologic disorder. *J. Med. Genet.* (2024) doi:10.1136/jmg-2023-109447.

123. Yoon, M. The role of PPARalpha in lipid metabolism and obesity: focusing on the effects of estrogen on PPARalpha actions. *Pharmacol. Res.* **60**, 151–159 (2009).

124. Wang, X. & Kilgore, M. W. Signal cross-talk between estrogen receptor alpha and beta and the

peroxisome proliferator-activated receptor gamma1 in MDA-MB-231 and MCF-7 breast cancer cells. *Mol. Cell. Endocrinol.* **194**, 123–133 (2002).

125. Chu, R. *et al.* The cross-talk between estrogen receptor and peroxisome proliferator-activated receptor gamma in thyroid cancer. *Cancer* **120**, 142–153 (2014).
126. Siljee, J. E. *et al.* Subcellular localization of MC4R with ADCY3 at neuronal primary cilia underlies a common pathway for genetic predisposition to obesity. *Nat. Genet.* **50**, 180–185 (2018).
127. Sternson, S. M., Shepherd, G. M. G. & Friedman, J. M. Topographic mapping of VMH --> arcuate nucleus microcircuits and their reorganization by fasting. *Nat. Neurosci.* **8**, 1356–1363 (2005).
128. Hernández, V. S. *et al.* A comprehensive chemotyping and gonadal regulation of seven kisspeptinergic neuronal populations in the mouse brain. *BioRxiv* (2024) doi:10.1101/2024.07.23.604881.
129. Rance, N. E. & Bruce, T. R. Neurokinin B gene expression is increased in the arcuate nucleus of ovariectomized rats. *Neuroendocrinology* **60**, 337–345 (1994).
130. Navarro, V. M. *et al.* Regulation of gonadotropin-releasing hormone secretion by kisspeptin/dynorphin/neurokinin B neurons in the arcuate nucleus of the mouse. *J. Neurosci.* **29**, 11859–11866 (2009).
131. Vetrini, F. *et al.* De novo and inherited TCF20 pathogenic variants are associated with intellectual disability, dysmorphic features, hypotonia, and neurological impairments with similarities to Smith-Magenis syndrome. *Genome Med.* **11**, 12 (2019).
132. Blanchet, P. *et al.* MYT1L mutations cause intellectual disability and variable obesity by dysregulating gene expression and development of the neuroendocrine hypothalamus. *PLoS Genet.* **13**, e1006957 (2017).
133. Labonté, B. *et al.* Sex-specific transcriptional signatures in human depression. *Nat. Med.* **23**, 1102–1111 (2017).
134. Issler, O. *et al.* The long noncoding RNA FEDORA is a cell type- and sex-specific regulator of depression. *Sci. Adv.* **8**, eabn9494 (2022).
135. Issler, O. *et al.* Sex-Specific Role for the Long Non-coding RNA LINC00473 in Depression. *Neuron* **106**, 912-926.e5 (2020).
136. Spencer, S. *et al.* Distribution of catecholamine-containing neurons in the normal human hypothalamus.

Brain Res. **328**, 73–80 (1985).

137. Dudas, B. *et al.* Distribution and morphology of the catecholaminergic neural elements in the human hypothalamus. *Neuroscience* **171**, 187–195 (2010).

138. Grattan, D. R. 60 YEARS OF NEUROENDOCRINOLOGY: The hypothalamo-prolactin axis. *J. Endocrinol.* **226**, T101-22 (2015).

139. Mu, W. *et al.* Hypothalamic Rax+ tanycytes contribute to tissue repair and tumorigenesis upon oncogene activation in mice. *Nat. Commun.* **12**, 2288 (2021).

140. Sriworarat, C. *et al.* Performant web-based interactive visualization tool for spatially-resolved transcriptomics experiments. *Biol. Imaging* **3**, e15 (2023).

141. Yao, Z. *et al.* A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *BioRxiv* (2023) doi:10.1101/2023.03.06.531121.