Integrative single-cell characterization of hypothalamus sex-differential and obesity-associated genes and regulatory elements

Hai P. Nguyen[1,2,#], Candace S.Y Chan[1,2,#], Dianne Laboy Cintron[1,2], Rory Sheng[1,2] , Lana Harshman[1,2], Mai Nobuhara[1,2], Aki Ushiki[1,2], Cassidy Biellak[1,2], Kelly An[1,2] Gracie M. Gordon[1,2,3,4], Francois Mifsud[5], Abbey Blake[5], Eric J. Huang[6], Martin Hemberg[7,], Christian Vaisse[5], Nadav Ahituv[1,2,*]

[1] Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, CA, USA.

[2] Institute for Human Genetics, University of California San Francisco, San Francisco, CA, USA.

[3] Biological and Medical Informatics Graduate Program, University of California, San Francisco, CA, USA.

[4] Department of Rheumatology, University of California, San Francisco, CA, USA.

[5] Diabetes Center, University of California San Francisco, San Francisco, CA, USA.

[6] Department of Pathology, University of California San Francisco, San Francisco, CA, USA.

[7]Evergrande Center for Immunologic Diseases, Harvard Medical School and Brigham and Women's Hospital, Boston, MA, USA.

[#] The authors contributed equally.

*Corresponding author: nadav.ahituv@ucsf.edu

Abstract

Over 500 noncoding genomic loci are associated with obesity. The majority of these loci reside near genes that are expressed in the hypothalamus in specific neuronal subpopulations that regulate food intake, hindering the ability to identify and functionally characterize them. Here, we carried out integrative single-cell analysis (RNA/ATAC-seq) on both mouse and human male and female hypothalamus to characterize genes and regulatory elements in specific cell subpopulations. Utilizing both transcriptome and regulome data, we identify over 30 different neuronal and non-neuronal cell subpopulations and a shared core of transcription factors that regulate cell cluster-specific genes between mice and humans. We characterize several sex-specific differentially expressed genes and the regulatory elements that control them in specific cell subpopulations. Overlapping cell-specific scATAC peaks with obesity-associated GWAS variants, identifies potential obesity-associated regulatory elements. Using reporter assays and CRISPR editing, we show that many of these sequences, including the top obesity-associated loci (*FTO* and *MC4R*), are functional enhancers whose activity is altered due to the obesity-associated variant and regulate known obesity genes. Combined, our work provides a catalog of genes and regulatory elements in hypothalamus cell subpopulations and uses obesity to showcase how integrative single-cell sequencing can identify functional variants associated with hypothalamus-related phenotypes.

Introduction

Obesity dramatically increases the risk of hypertension, dyslipidemia, diabetes, COVID-19, and cardiovascular diseases and is a major public health concern[1,2]. Environmental and genetic factors are both involved in the onset and progression of weight gain[3], with genetic factors being a major component[4,5]. Both rare, monogenic obesity causing variants, and common variants, detected through large obesity genome wide association studies (GWAS), have been found to predispose to obesity. While rare gene mutations causing obesity have outlined a neuro-physiological pathway involved in body weight regulation, the majority of which are part of the hypothalamic leptin-melanocortin pathway[6], how common variants influence body weight remains poorly understood. GWAS have identified over 500 common loci associated with obesity[6-9]. The majority of obesity GWAS loci represent clusters of common variants in noncoding regions that likely alter gene expression and have been found to reside primarily near genes involved in central nervous system (CNS)-related processes[9-11]. In addition, a number of loci appear to directly influence neurons of the hypothalamic leptin-melanocortin system[10,12]. Combined, these studies single out neuronal subpopulations in the hypothalamus that are involved in food intake regulation as the major cause of this condition.

The hypothalamus is one of the most complex regions in the brain, composed of diverse neurons and glial cells that are spatially distributed in different regions or nuclei[13-15]. Extensive studies of the hypothalamus have demonstrated that individual nuclei have distinctive functions in regulating different homeostatic processes, such as thermoregulation, mating motivation, sweat, blood pressure, reward, sleep and hunger[16-22]. Most of these processes are controlled by neuronal subtypes in each nucleus. In addition, sex-dimorphism in the hypothalamus has been characterized[23-25], with the medial preoptic nucleus being a major sex-dimorphic region, releasing different hormones depending on sex[26-28]. However, the understanding of sex-differential gene expression and transcriptional regulation in the hypothalamus remains largely uncharacterized.

With advances in single-cell sequencing technologies, several studies have set out to characterize diverse cell populations in the whole hypothalamus or specific nuclei using single-cell RNA sequencing (scRNA-seq)[29-41]. These studies characterized various hypothalamus neuronal populations, such as GABAergic and glutamatergic neurons, non-neuronal cell types, including microglia, oligodendrocytes, and astrocytes. While these studies were able to identify various cell populations based on their transcriptional profiles, they lack the capacity to identify the regulatory elements driving these cell type-specific subpopulations, where many variants of the obesity GWAS or other hypothalamus-associated phenotypes could reside.

Integrative single-cell sequencing utilizes RNA-seq to identify the transcriptome and ATAC-seq to characterize the regulome in the same cell. Here, we utilized this technique to generate a combined scRNA- and sc-ATAC seq atlas of the adult female and male mouse and human hypothalamus. This atlas, composed of more than 50,000

cells, identified numerous non-neuronal and neuronal cells that play key roles in controlling various hypothalamus-associated physiological processes, including blood pressure, mating, nursing, and feeding. We also identified cell-type specific regulatory elements and linked them to their target genes in several cell clusters. Moreover, we uncover cell cluster-specific transcription factor (TF) activity and their regulated genes and identify cell cluster-specific core TFs shared between mice and humans. In addition, we annotated sex-associated gene expression and regulatory element differences, including hormones/neuropeptides in different neuronal subtypes that are regulated by TF expressed in a sex-dimorphic manner. Overlapping our cell-type specific regulatory elements with obesity GWAS variants, we identified eighteen potential enhancers for several obesity associated loci, including *FTO* and *MC4R*, the top two associated obesity variants[42,43]. Enhancer assays in mouse hypothalamus cells, identify ten of these to be functional enhancers and five to alter enhancer activity due to the obesity-associated variant. Genetic disruption or CRISPR inactivation (CRISPRi) mediated downregulation of these enhancers in cultured human neurons or astrocytes leads to a reduction of gene expression of their target genes. Taken together, this work provides a catalog of genes and regulatory elements in different cellular subpopulations of the hypothalamus in both males and females in mice and humans which can be utilized to identify causative sequences that lead to hypothalamus-associated phenotypes, such as obesity.

Results

Combined sc-RNA and ATAC profiling of the mouse hypothalamus

To characterize both the genes and regulatory elements of various hypothalamic cell types, we utilized the 10X Chromium multiome platform. We initially focused on mouse, utilizing three male and three female hypothalami dissected from 6 weeks old C57BL/6J mice (**Fig.1a**). We isolated nuclei via FACS and subjected them to 10X Chromium Multiome following established protocols (see **Methods**). Overall, we sequenced 72,504 cells. After quality control filtering (see **Methods**), we retained 13,455 from female and 13,329 from male mice, amounting to 26,784 mouse hypothalamus cells in our dataset (**Extended Data Fig.1a**). We next set out to characterize the various hypothalamus cell populations using our scRNA-seq data. Utilizing differential expressed genes and cell-type specific markers curated from literature and previous studies in hypothalamus, we first determined the various cell-type clusters[29-41]. We identified 37 distinctive cell populations, including 34 previously-described cell-types and four specialized neurons that were not described in previous studies. These include 33 different types of neurons and oligodendrocytes, astrocytes, and microglia cell clusters (**Fig.1b**).

For neuronal cells, we found 33 different neuronal cell populations, including glutamatergic, GABAergic, and specialized neuron clusters. We found 17 glutamatergic neuron populations that showed high expression levels of the glutamatergic marker, *Slc17a6*, which encodes vesicular glutamate transporter 2 (*Vglut2*). We also identified nine clusters of GABAergic neurons, expressing GABAergic marker *Slc32a*, which encodes the vesicular GABA transporter (*Vgat*) (**Fig.1c**). The GABAergic neurons also expressed the synthetic enzyme for GABA, *Gad1* (**Extended Data Fig.1b**). Furthermore, we found other specialized neurons secreting various neuronal peptides and hormones that regulate whole-body endocrine and physiological homeostasis. Among these cells, we identified Pomc/Cart neurons that distinctly express both proopiomelanocortin (*Pomc*) and cocaine and amphetamine regulated transcript (*Cartpt*) and Npy/Argp neurons, highly expressing neuropeptide Y (*Npy*) and agouti-related peptide (*Agrp*) (**Fig.1c**). Both cell types are known to control satiety by regulating melanocortin 4 receptor (*Mc4r*) neurons activity[44] (**Extended Data Fig.1b**). These neurons were identified in various nuclei, including paraventricular nucleus, accurate nucleus, and intermediolateral nucleus[45,46]. Similar to previous reports[47,48], we found that Pomc/Cart and Npy/Agrp neurons express leptin receptor (*Lepr*) (**Extended Data Fig.1b**). In addition, we identified oxytocin (*Oxt*)- and pro-melanin concentrating hormone (*Pmch*)- secreting neurons (**Fig.1c**). While *Oxt* is known for its role in female reproduction, breastfeeding, and childbirth[49], *Pmch* controls skin pigmentation and also has important roles in regulating motivated behaviors, such as feeding, drinking, and mating[50]. We also identified dopamine-producing neurons, which show high expression levels of dopamine synthetic enzymes, such as tyrosine hydroxylase (*Th*) and dopa decarboxylase (*Ddc*) (**Fig.1c, Extended Data Fig.1b**). We annotated hypocretin neuropeptide (*Hcrt*) or orexin -expressing neurons, which are reported to regulate arousal, wakefulness, and appetite[51,52]. Additionally, we found growth hormone-releasing hormone (*Ghrh*)- specific neurons (**Fig.1c**), which regulate growth hormone production in the anterior pituitary gland[53,54]. Analysis of other neuronal peptides, found

oxytocin neurons expressed *Avp* (vasopressin), regulating vascular tone, blood pressure and urine concentration by the kidney[55] (**Extended Data Fig.1c**). We found that *Ghrh* neurons highly expressed *Gal* (Galanin-like peptide), which is known to regulate feeding and reproduction[56] (**Extended Data Fig.1c**). In addition, glutamatergic neurons widely expressed *Adcyap1*, encoding for pituitary adenylate cyclase-activating polypeptide binding to its receptor in intestinal cells and this gene is associated with post-traumatic stress disorder (PTSD)[57] (**Extended Data Fig.1c**).

We also identified several non-neuronal cell populations. These include oligodendrocytes (ODC), oligodendrocyte precursor cells (OPC), which show high expression of suppression of tumorigenicity 18 (*St18*) and proline-rich 5 like (*Prr5l*) genes and mature oligodendrocytes, expressing platelet-derived growth factor receptor A (*Pdgfra*) and oligodendrocyte transcription factor 1 (*Olig1*) (**Fig.1c**). In addition, we found astrocytes that show high levels of astrocyte gene markers, such as angiotensinogen (*Agt*) expression and microglia that express the immune marker cathepsin S (*Ctss*) (**Fig.1c**).

We next examined our scATAC-seq data, obtained from cells that were jointly profiled with scRNA-seq. Dimensionality reduction was performed using latent semantic indexing and batch correction using Harmony[58] (see **Methods**). Peak calling on each cell-type from joint clustering recovered 359,963 chromatin accessible peaks from cells that passed quality control. Amongst them, 157,204 (43%) peaks were found in only one cluster, indicating robustness of clustering and annotation. Comparing differential gene expression and accessible peak profiles, found that all our cell populations had unique gene and chromatin signatures in a cell-specific manner (**Fig.1d**). We found 37,556 marker peaks that were enriched in specific cell-types, with many of them located near cell type-specific gene markers (5,400). For example, we found increased accessibility near *Slc176* in glutamatergic neuron cell types, *Slc32a6* in GABAergic neuronal populations, *Ddc* in dopamine neurons, and *Hcrt* in Hcrt neurons (**Extended Data Fig.1d**).

Combined scATAC and scRNA sequencing allows to identify potential *cis*-regulatory elements by annotating peaks where chromatin accessibility correlates with gene expression (see **Methods**). We found 245 peaks that were linked to the expression of 65 marker genes. Among the specialized neurons, *Pomc* expression was linked to two scATAC-seq peaks that were highly specific to Pomc/Cart neurons cells (**Fig.1e**). These elements overlapped with two enhancers reported in[59]. We also identified cell type-specific scATAC-seq peaks that were linked to *Cartpt* expression in neuronal populations, such as Pomc/Cart and *Pmch* neurons that highly express *Cartp* (**Fig.1e**). In addition, we annotated three *Npy* linked scATAC-seq peaks that were highly specific to *Npy*/*Argp* neurons. *Pomc* and *Cartpt* are key regulators of feeding, metabolic homeostasis and obesity[60,61] and their enhancer identification could be used to modulate their expression. We also found a *Hcrt* neuron specific scATAC-seq peak that was linked to *Hcrt* expression (**Fig.1e**) and a specific peak that was linked to *Th* expression in dopamine neurons (**Extended Data Fig.1e**). Moreover, there were multiple peaks that were linked to *Slc17a6* expression in GABAergic neurons

(**Extended Data Fig.1e**). Combined, these results show the ability of integrative single-cell analyses to identify specific hypothalamus cell subpopulations and link potential gene regulatory elements with their putative target genes.

Combined sc-RNA and ATAC profiling of the human hypothalamus

We next conducted a similar integrative single-cell experiment on three males and three female adult human hypothalami. We collected post-mortem hypothalami from patients ages 43-82 and a post-mortem interval (PMI) between 4-12 hours (**Supplementary Table 1, Extended Data Fig.2a**) and carried out scRNA/ATAC-seq, as described for mice. Due to the longer PMI of these samples, compared to mice, we initially obtained a lower number of cell populations. We thus carried out an additional scRNA/ATAC-seq experiment on the same samples subjecting them to NeuN (neuronal marker) FACS sorting to increase for neuronal populations. We combined both experiments and sequenced a total of 113,854 cells, which after filtering for high quality cells retained 34,257 cells (see **Methods**; **Extended Data Fig.2b**).

To improve the quality of our analyses, we removed estimated ambient RNA from our data using SoupX[62]. Joint analysis of scATAC and scRNA identified five broad clusters, which we annotated as mature oligodendrocytes, oligodendrocyte progenitor cells, astrocytes, microglia, and neurons (**Fig.2a**). Using differentially expressed genes, we found that neuronal clusters showed high expression for excitatory neuronal markers, including *GAD2*, which encodes for glutamate decarboxylase 2 and *NRG1*, encoding neuregulin 1 (**Fig.2b**). They also expressed other neuron markers, including stathmin 2 (*STMN2*) and synaptotagmin (*STY1*). Astrocytes expressed astrocyte gene markers, such as glial fibrillary acidic protein (*GFAP*) and *AGT* (**Fig.2b**). Similar to mice, OPC expressed oligodendrocyte markers, such as *PDGFRa* and chondroitin sulfate proteoglycan 4 (*CSPG4*). More mature oligodendrocytes expressed proteolipid protein 1 (*PLP1*) and myelin oligodendrocyte glycoprotein (*MOG*) (**Fig.2b**). We also found microglia cells with high expression of the pan-immune marker *PTPRC* (CD45), complement C1q B chain (*C1QB*), integrin alpha M (*ITGAM*), and integrin subunit alpha X (*ITGAX*) (**Fig.2b**). We found distinctive gene expression and scATAC profiles that are highly specific to each cell type (**Extended Data Fig.2c**).

Despite the neuron top-off experiment, we found that the majority of cells (69.7%) isolated from human samples consisted of mature oligodendrocytes and had a much smaller neuronal population (1.7%). This low neuronal population recovery is consistent with other recent multiomic single-cell assays on post-mortem human midbrain or frontal lobe[63,64]. To increase the resolution of neuronal cells and other smaller populations, we subclustered these cells without the mature ODC, retaining 10,356 cells. Removal of the ODC cluster and re-clustering led to the identification of GABAergic and glutamatergic neuronal clusters, comprising 145 cells and 271 cells respectively (4.01%). In addition, we found two neural progenitor clusters, three astrocyte clusters, four microglia clusters, three myelinating oligodendrocytes (MODC) and three OPC clusters originating from the OPC population (**Fig.2c**). After subsetting our dataset, we found that 38.2% of peaks were specific to one cluster.

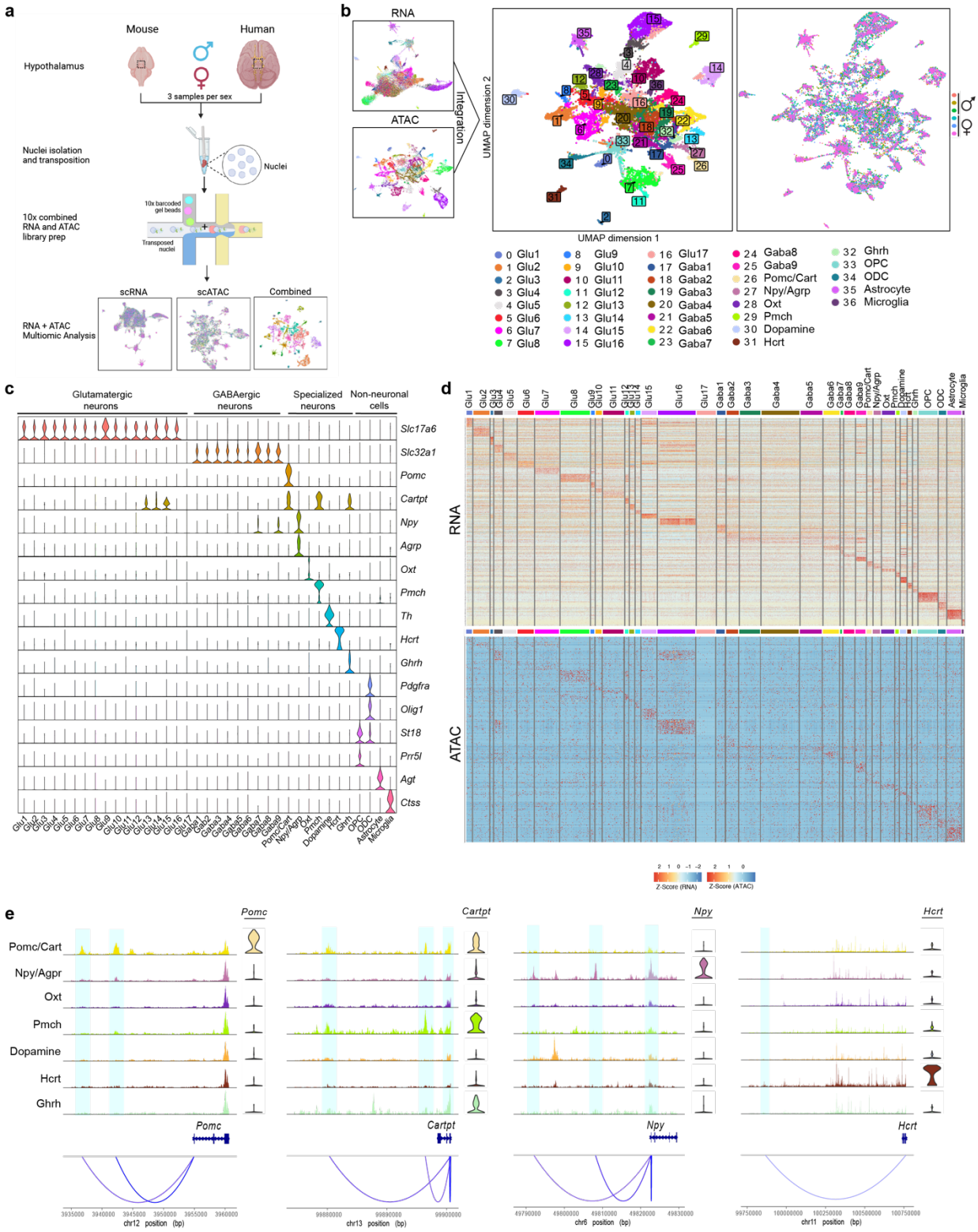


Fig 1. Integrative single-cell RNA and ATAC atlas of adult mouse hypothalamus. a, Schematic of preparation of mouse and human hypothalamus of three male and three female samples and 10X multiomic library prep. **b,** Uniform Manifold Approximation and Projection (UMAP) of single-cell RNA,

ATAC, and integrative data (center). Clusters were identified based on scRNA-seq data. Right panel shows the UMAP by subject. **c,** Violin plot of gene markers used to identify all clusters. **d,** Heatmap of the top five differentially expressed genes and the top ten ATAC peaks across all cell populations. **e,** scATAC peaks to gene linkage plots for *Pomc*, *Cartpt*, *Npy*, and *Hcrt* with genome tracks of specialized neuron populations, including Pomc/Cart, Npy/Agrp, Oxt, Pmch, Dopamine, Hcrt, and Ghrh cell clusters. Gene expression violin plots for each cell type are shown on the right. The linked scATAC peaks to gene promoters are highlighted in light blue.

Similar to the mouse data, glutamatergic neurons had high expression levels of *SLC17A6* while GABAergic neurons showed high *SLC32A1* expression (**Fig.2d**). In glutamatergic neurons, we found oxytocin-expressing neurons with high expression of *OXT* and *AVP* (**Fig.2d**, **Extended Data Fig.2d**). We also found *POMC*- and *CARTPT*-expressing neurons, as well as *NPY*-expressing cells (**Extended Data Fig.2d**) and *PMCH*, *HCRT*, and *GHRH* neurons (**Extended Data Fig.2d**). GABAergic neurons showed high expression levels of *COL1A1, GRIP2*, and hydroxytryptamine receptor 2c (*HTR2C*), and dopamine receptor D1 (*DRD1*) (**Fig.2d**). In the GABAergic neuronal population, we also found neurons expressing neuropeptides, such as vasoactive intestinal peptide (*VIP*), proenkephalin (*PENK*) which plays a role in pain perception and response to stress[65], tachykinin (*TAC1*) which is a ghrelin target and is involved in energy balance and food intake[66], and thyrotropin releasing hormone (*TRH*) which stimulates the release of thyroid stimulating hormone (*TSH*) and prolactin from the anterior pituitary[67] (**Extended Data Fig.2d**). Interestingly, we identified two neural progenitor (NP) clusters that showed high expression of neural progenitor gene marker, Vimentin (*VIM*) and stem cell/progenitor marker, transcription factor *HES1* (**Fig.2d**).

Using markers from previous single-cell studies on oligodendrocytes[68], we identified three OPC and three myelinating ODC (MODC) populations. We found that OPC clusters had high expression levels of *PDGFRA*, protocadherin related 15 (*PCDH15*) and versican (*VCAN*) (**Fig.2d**). As OPC differentiate to myelin oligodendrocytes, they expressed G protein-coupled receptor 17 (*GPR17*), a Gi-coupled GPCR, that acts as an intrinsic timer of oligodendrocyte differentiation and myelination[69]. The MODC population showed high expression of transcription factor 7 like 2 (*TCF7L2*) and oligodendrocytic myelin paranodal and inner loop protein (*OPALIN*) (**Fig.2d**). We also identified four astrocyte clusters with high expression of *AGT* and *GFAP*. Examination of gene expression profiles of these astrocyte clusters found that each had a unique function. Genes that were highly expressed in MODC1 were involved in glutathione transport, plasma membrane repair and mitochondria distribution. In contrast, the MODC2 cluster showed high expression for genes involved in amino acid transport. Genes expressed in MODC3 were involved in glutamine transport and detoxification of copper ions, cholesterol and inorganic compounds. In addition, MODC4 expressed genes were involved in response to salt stress and calcium transport (**Extended Data Fig.2e**). Among the non-neuronal cell types, we also identified four microglia populations with high *ITGAM* and *ITGAX* expression (**Fig.2d**). In addition, we found that CX3CR1 and CD86, gene markers for activate/proinflammatory microglia, were expressed most highly in the microglia 1 cluster and at lower levels in microglia 2 and microglia 3 clusters. Microglia cluster 4 showed high expression for CD163, a gene marker for quiescent/anti-inflammatory microglia[70] (**Fig.2d**, **Extended Data Fig.2f**).

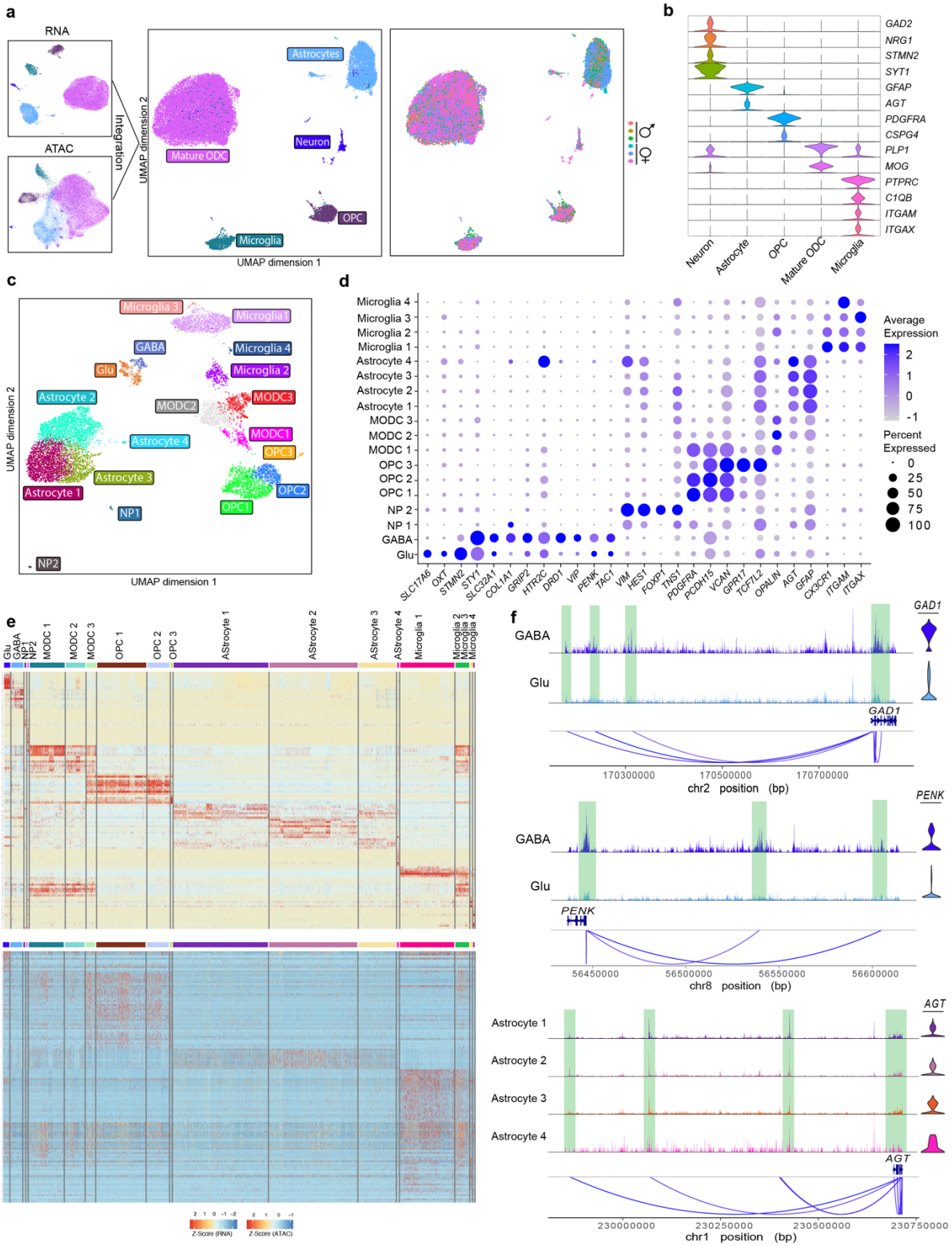


Fig 2. Integrative single-cell RNA and ATAC atlas of adult human hypothalamus. a, Uniform Manifold Approximation and Projection (UMAP) of single-cell RNA, ATAC, and integrative data of human hypothalamus samples. Right panel shows the UMAP by subject. **b,** Violin plot of gene markers used to identify cell clusters. **c,** UMAP of integrative data after removal of mature oligodendrocytes. Clusters were identified based on scRNA-seq data. **d,** Dot plots of gene markers used to identify cell clusters. **e,**

Heatmap of the top five differentially expressed genes and the top ten ATAC peaks across all cell populations. **f,** scATAC peaks to gene linkage plots for *GAD1*, *PENK*, and *AGT* in GABAergic and glutamatergic neurons and astrocytes with the genome tracks. Gene expression violin plots are shown on the right. The linked scATAC peaks to gene promoters are highlighted in light green.

We observed that all our cell populations had unique gene and chromatin profiles in a cell-specific manner (**Fig.2e**). We next examined our scATAC-seq data and found 65,560 peaks (38.2%) peaks found in only one cluster. 1,271 were differentially accessible, of which 30.9% (394) of differentially accessible peaks were located near cell type-specific gene markers. For example, we found higher number of reads in ATAC peaks near *GAD2* and *OXT* in GABAergic neurons, *GFAP* in the astrocyte populations and *CDPG4* in the MODC clusters (**Extended Data Fig.2g**). We next analyzed 46 cell marker genes to see whether chromatin accessibility in their locus correlated with gene expression. We examined chromatin accessibility up to one million base pairs away from the marker gene, which is roughly the median topologically associated domain (TAD) length[71]. Amongst the 46 marker genes analyzed, we identified 490 regions where accessibility correlated with gene expression. These regions were on average 289,499 (median 189,142) base pairs away from the transcription start site (TSS) of their postulated target gene. For example, we found that the expression of *GAD1*, a GABAergic neuronal marker, was linked to four scATAC-seq peaks in GABAergic neurons (**Fig.2f**). We also found three scATAC-seq peaks linked to the expression of proenkephalin (PENK). In addition, in the four astrocyte clusters, we found the expression of angiotensinogen (*AGT*), which regulates blood pressure, to be linked to seven scATAC-seq peaks (**Fig.2f**). Similar to our mouse data, using the integrative single-cell analyses allowed us to identify specific cell subpopulations in the human hypothalamus and potential gene regulatory elements with their putative target genes.

Motif enrichment analyses identify cell-type specific regulatory pathways

We next set out to identify the major transcription factors (TFs) that regulate the various hypothalamus cell types, initially starting with mice. We identified motifs enriched in accessible regions for each cell-type using chromVAR[72], finding distinct profiles of TF binding sites (TFBS) for each cell cluster. For example, for the glutamatergic neuron populations, we found significant TFBS enrichment of hypothalamic transcription factor nescient helix-loop-helix 2 (NHLH2), a known regulator of metabolic activity and puberty in the paraventricular nucleus (PVN)[73,74](**Fig.3a**). We also found enrichment for the forkhead box P1 (FOXP1) motif and the NeuroD family of basic helix-loop-helix 1 (NEUROD1) motif, both known to regulate glutamatergic neurogenesis[75,76](**Fig.3a**). In addition, we found the TFBS for cutlike homeobox 1 (CUX1), which controls neuronal proliferation, dendrite branching and synapse formation to be enriched in GABAergic neurons[77](**Fig.3a**). Dopamine neurons showed enrichment for FOXA1/FOXA2 motifs, known to be associated with dopamine biosynthesis[78]. For non-neuronal populations, we identified binding motif enrichment of known cluster-specific TFs, such as *SOX9* in OPCs[79], *BHLH21*, encoded by *Olig1* in oligodendrocyte and Spi-B transcription factor (Spi-1/PU.1 related; *SPIB*) known to regulate microglia/macrophages[80](**Fig.3a**).

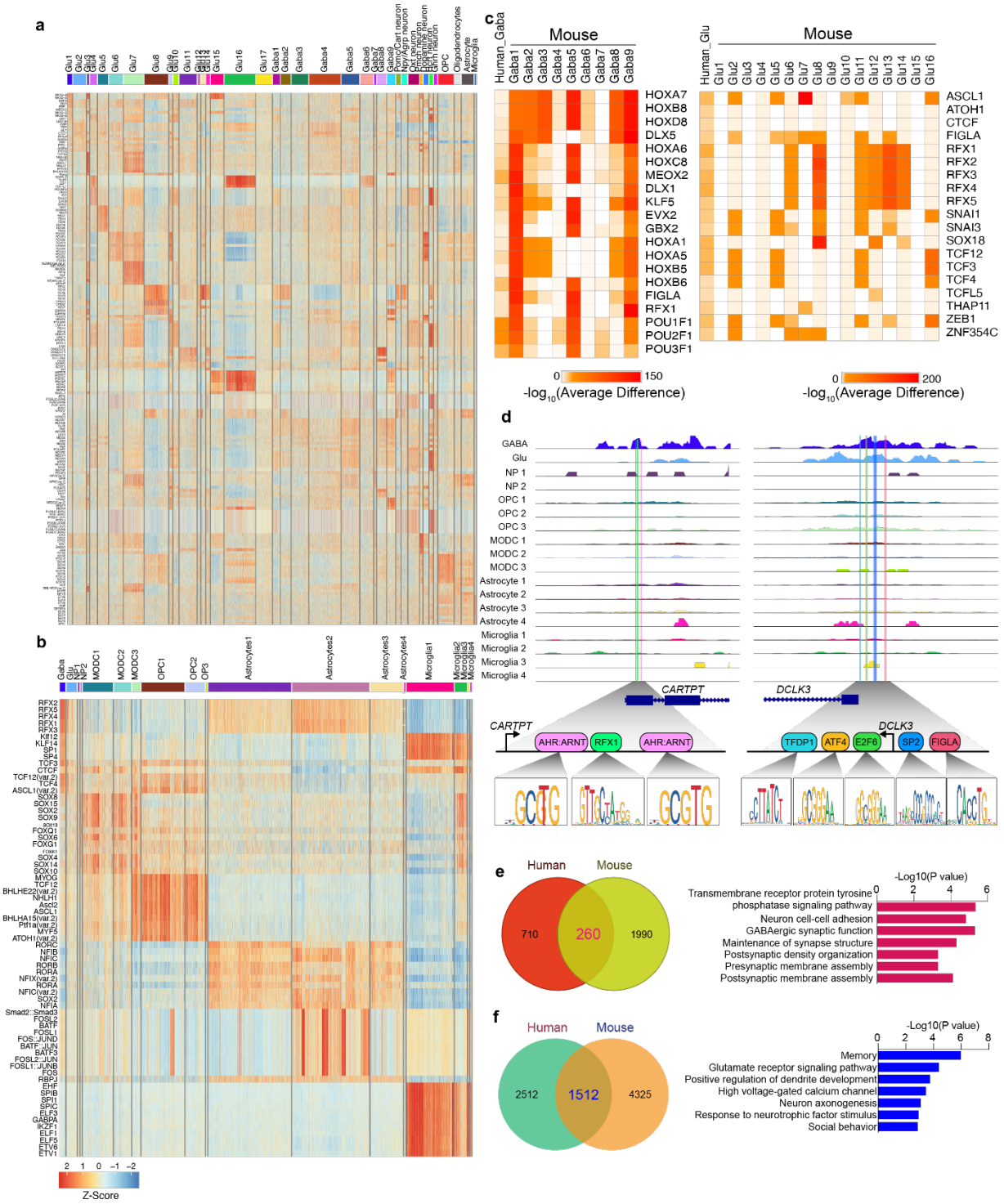


Fig 3. Motif enrichment analyses identify cell-type specific regulatory pathways. a-b, Heatmap of the top five differentially enriched transcription factors binding sites (TFBS) across cell populations labeled with different colors in the mouse (**a**) or human (**b**) hypothalamus. **c,** Heatmap of the top TFBS enriched in GABAergic and glutamatergic neuronal populations in mouse and human hypothalamus. **d,** Genome tracks of *CARTPT* and *DLCK3* showing logo plots for various TFBS in their promoters. **e-f,** Venn diagrams showing the overlap of TFBSs enriched in GABAergic (**e**) or glutamatergic (**f**) neurons of mouse and human hypothalamus along with gene ontology analysis of enriched TFBS in these neurons.

In humans, similar to mice, we identified unique TFBS profiles in a cell type-specific manner. For glutamatergic and GABAergic neurons, we observed an enrichment for the NHLH2 TFBS (**Fig.3b**). We found that both mouse and human neuronal populations have highly enriched motifs for regulatory factor binding to the X-box (RFX) family, including *RFX1* and *RFX3*. *SOX9* TFBS and its family members were enriched in OPC and MODC populations (**Fig.3b**). We also found a TFBS enrichment for transcription factor 3 (*TCF3*), which is known to maintain OPC[81]. In addition, we found *BHLH21* TFBSs were highly enriched in astrocytes along with nuclear factor IA (NIFA), which is known to induce astrocyte formation[82] (**Fig.3b**). Similar to mouse data, microglia cells showed an enrichment for *SPIB* TFBSs (**Fig.3b**).

We next compared mouse and human TFBSs whose motifs were enriched in a cell-cluster specific manner. In both mouse and human GABAergic and glutamatergic neuronal populations, we observed a large number of shared TFBSs, such as *HOXA7*, *MEOX2*, *NHLH2*, *RFX1* and *RFX3* (**Extended Data Fig.3a,b**). Analysis of mouse and human GABAergic neurons for the TFBSs that have the most enriched motif scores, found many shared TFBS to be enriched in these neurons, such as distal-less homeobox 5 (*DLX5*), well-known to regulate GABAergic neurons and control behavior and metabolism[83], even-skipped homeobox 1 (*EVX1*), which is required for GABAergic neuron development[84], and *POU3F1* which regulates neural differentiation[85](**Fig.3c**). In the glutamatergic neuronal populations, we found mouse and human cells to show TFBS enrichment for the RFX TF family, achaete-scute family BHLH transcription factor 1 (*ASCL1*), that is required for dopaminergic neurons differentiation from iPSC cells[86] and *THAP11* or RONIN, known regulators of neural stem cell differentiation[87](**Fig.3c**). In the glutamatergic neurons, we also found TFBS enrichment for TFs that have not been reported to have specific function in neuronal cells, such as folliculogenesis specific BHLH transcription factor (*FIGLA*), which has a known function in postnatal oocyte-specific gene expression and zinc finger E-box binding homeobox 1 (*ZEB1*), which is known to suppress the T-lymphocyte-specific IL2 gene[88] (**Fig.3c**). In addition, we found that many other non-neuronal clusters in mouse and human hypothalamus shared many TFBSs that potentially regulate genes in these cell types (**Extended Data Fig.3c**). Combined, these results suggest that hypothalamus cell populations in mice and humans utilize a similar cell type-specific regulatory machinery.

We next tested whether the TFBS that were enriched in specific cell clusters are associated with TFs expression in the corresponding clusters. We compared TFBS enrichment to the corresponding TF expression, finding several TFs to be both differentially expressed and enriched in chromatin accessible regions (**Extended Data Fig.3d**). However, we did not find statistically significant correlation ($r > 0.5$, Pearson correlation) between TF expression and motif enrichment, possibly as TF concentration is not the sole determinants of DNA binding[89]. However, a small number of TFs which had enriched TFBS in a cluster-specific manner had high expression in the corresponding clusters. For example, *KLF5*, *FIGLA* and *RFX1* were found to be highly expressed and have enriched TFBS scores in GABAergic neurons (**Fig.3d**). We also found *THAP11* TFBS to be enriched and with high expression in glutamatergic neurons.

We also examined TFBSs corresponding to scATAC peaks in the promoters of marker genes of neuronal populations. Examination of the *CARTPT* promoter, a glutamatergic neuronal marker, found multiple TFBSs for hydrocarbon receptors (AHR:ARNT), which are known to regulate and determine neuronal fate[90] and *RFX1* (**Fig.3d**), whose expression were high in glutamatergic neurons in mouse and rat[91,92]. In addition, we found multiple TFBSs for *ATF4*, *E2F6*, *FIGLA*, *SP2*, and *TFDP1* in the promoter of doublecortin like kinase 3 (*DCLK3*) (**Fig.3d**), which is highly expressed in GABAergic neurons (**Extended Data Fig.3e**) and is known to have a protective role in these cells[93].

We next examined genes that might could be regulated by TFs with cell-type enriched TFBS in both mice and humans. For each TF, we searched for the nearest genes (within 2kb) of motif-containing cell-type accessible peaks (see **Methods**). We then compared the list of genes predicted to be regulated by TFs in neuron populations between humans and mice. In GABAergic neurons, we identified 970 regulated genes in humans and 2,250 genes in mice (**Fig.3e**). For GABAergic neurons, we found 260 genes that overlap between humans and mice, which following gene ontology (GO) analyses[94] were found to be enriched for neuron-specific pathways, including neuron cell-cell adhesion, GABAergic synaptic function, and postsynaptic density organization (p-value < 0.05, hypergeometric distribution test) (**Fig.3e**). Interestingly, the human specific pathways in GABAergic neurons were found to be enriched for circulatory system development, cytoskeleton organization, and cell projection organization (**Extended Data Fig.3f**) while the mouse specific pathways in GABAergic neurons were enriched for cell-substrate adhesion, extracellular structure organization and cell-matrix adhesion (**Extended Data Fig.3g**). These results suggest that these species-specific pathways in GABAergic neurons are involved in general cell maintenance. In glutamatergic neurons, we identified 4,200 genes in humans and 5,837 genes in mice, with both species sharing 1,512 genes (**Fig.3f**). GO analysis of shared genes found them to be involved in memory, glutamate receptor signaling pathway and high voltage-gated calcium channels (**Fig.3f**). Human specific glutamatergic neuron-specific regulated genes were found to be involved in neuron projection, axon extension, and actin filament organization, while mice specific genes were involved in heparan sulfate proteoglycan biosynthesis, ERBB2 signaling pathway and response to forskolin (**Extended Data Fig.3h,i**). These data indicate that mouse and human hypothalamus cell populations share many common TFs that regulate similar cluster-specific gene networks and have networks unique to each species. In addition, our datasets provide the ability to identify TFs that are involved in cell-type specific hypothalamus gene-regulatory networks.

Characterization of sex-differential gene and regulatory activity

The hypothalamus is a neuro-endocrine organ that releases various neuropeptides in a sex-specific manner. We thus annotated sex-differential gene expression across the cell populations in the human hypothalamus. We first identified genes that were differentially expressed between males and females in corresponding cell types. We detected sex-differential expressed genes across different cell types, but their numbers are small within some clusters (10 out of 18 clusters) having less than 10 sex-differentially expressed genes (**Fig.4a**). We next examined various neuropeptides in the neuronal

populations. As a positive control[95,96], we first examined the estrogen receptor (*ESR1*) and androgen receptor (*AR*), finding *ESR1* to be more highly expressed in glutamatergic neurons of females than males and *AR* to have higher expression in male GABAergic neurons (**Fig.4b**). When examining all neuropeptide genes, we found various genes to be differentially expressed in neuronal clusters between sexes. We found a number of neuropeptide genes to be expressed at higher levels in males than females, including *AVP*, *GHRH*, *GPR*, *NPY*, *HCRT*, and *TRH*. In contrast, we found *PENK* and *VIP* expression to be higher in females than males (**Fig.4b**). Analysis of these genes in mouse found them to be similarly sex-differentially expressed (**Extended Data Fig.4a**). Our findings are in line with previous studies. For example, *AVP* is more highly expressed in male GABAergic and glutamatergic neurons than in females and previous studies have shown that *AVP*, which regulates various social behaviors, including aggression and mating in males, has a higher expression in male hypothalamus than females[97,98]. We also found higher expression of *VIP* in female GABAergic neurons than males. In the hypothalamus, *VIP* is regulates the release of prolactin, stimulating milk production in the mammary gland[99,100] and *VIP* expression is found to be higher in the female brain than males[100,101]. In addition, we found novel sex-differentially expressed neuropeptides, including *GHRH*, *GPR*, *HCRT*, and *TRH* (**Fig.4b**), whose physiological process has been indicated to be sex-dimorphic. However, their cell-type specific expression in neurons have not been reported. For example, growth hormone secretion has been demonstrated to be sex-dimorphic, and higher in males in both mice and humans[102]. Our data shows that *GHRH* (growth hormone releasing hormone) is more highly expressed specifically in glutamatergic neurons, leading to higher secretion of *GH* (**Fig.4b**).

To further dissect the molecular mechanism of sex-differentially expressed genes, we examined potential sex-differential accessible chromatin regions located in the promoters of the sex-differentially expressed genes. We found 54 sex-differentially expressed genes that had sex-differential scATAC peaks located near their promoters. For example, in glutamatergic neurons, we observed increased accessibility at the promoter region and increased gene expression for the cyclin dependent kinase 5 regulatory subunit 1 (*CDK5R1*)(**Fig.4c**), which encodes for P35, a neuron-specific activator of cyclin-dependent kinase 5 (*CDK5*) and is required for central nervous development[103]. In addition, cyclin-dependent kinase 2 (*CDK2*) had higher expression in female neural progenitor clusters, accompanied by a female-specific scATAC peak at its promoter (**Fig.4c**). We also found *OLIG2*, an oligodendrocyte gene marker and known activator of myelin-associated genes, to have higher expression in males and increased accessibility in its promoter in male ODC compared to females (**Fig.4c**). Oligodendrocytes are known to have proliferation and differentiation sex-differences, which leads to sex-dimorphic risk to neurodegenerative diseases[104]. Sex-differential expression of *OLIG2,* the key transcription factor of oligodendrocytes might be responsible for sex-dimorphic differentiation of oligodendrocytes.

We identified several sex-differential scATAC peaks linked to corresponding sex-differentially expressed genes. For example, in GABAergic neurons, a scATAC peak was linked to proenkephalin (*PENK*) expression in males, but not in females (**Fig.4d**).

This gene plays a role in pain perception and response to stress and was shown to be more highly expressed in males versus female rats[105], similar to what we observed in humans. In microglia populations, *C1QB*, which is involved in synaptic pruning, showed higher expression and stronger scATAC peaks in males versus females in line with previous analyses of the developing brain[106]. In addition, we found a scATAC peak that was linked to the expression of the platelet derived growth factor receptor alpha (*PDGFRa*), which plays an important role in cell growth and division, in the female OPC population (**Fig.4d**).

Characterization of regulatory elements overlapping obesity-associated variants

We next set out to utilize our single-cell catalog of human hypothalamus cell-type specific genes and regulatory elements, to characterize obesity associated GWAS noncoding variants. We compiled a list of 508 index SNPs associated with obesity from the NHGRI-EBI GWAS catalog[107]. We performed linkage disequilibrium (LD) expansion on all the lead SNPs to include nearby variants with high probability of coinheritance (r^2 >0.8). The majority of these SNPs (>97%) localize to noncoding regions of the genome. We then checked for overlap of these SNPs with scATAC-seq peaks (**Fig.5a**), finding 102 SNPs that overlap 35 marker peaks. These SNPs comprised some of the most highly associated loci with obesity, including the top two loci (*FTO* and *MC4R*)[108,109]. Among the 35 peaks, we had two scATAC peaks linked to *FAIM2*, two scATAC peaks linked to *FTO*, and two scATAC peaks linked to *MC4R*. We did not observe a specific cell-type that was more strongly associated with obesity, having overlapping scATAC peaks from various cell types including oligodendrocytes, oligodendrocyte precursor cells, astrocytes and neurons.

We next tested the enhancer activity of 18 obesity-associated scATAC-seq peaks that overlapped with the strongest obesity-associated SNPs. These peaks are found to be specific to cell clusters, including OPC, MFO, astrocytes. We cloned these sequences into an enhancer assay vector that contains a minimal promoter followed by the luciferase reporter gene and transfected them into mouse neuronal hypothalamus cells. We found 10 out of 18 (55.5%) sequences to have luciferase activity that was significantly higher than the empty vector negative control (**Fig.5b**). These include sequences near the genes *ADCY3, FAIM2, FTO, GIPR, GNAT2, GPRC5B, LEMD2, MC4R*, *SDCCAG8, SH2B1, SULT1A1, TDH,* and *VPS45.* (**Fig.5b**). Using a similar assay, we also tested the enhancer activity of sequences from scATAC astrocyte-specific peaks (*FTO*, *SDCCAG8*, *LPP*) in primary human astrocytes, obtaining similar results to the neuronal cells (**Extended Data Fig.5a**). We next used site-directed mutagenesis to introduce the obesity-associated SNPs to the sequences that showed enhancer activity and tested them for differential activity compared to the unassociated variants. We found that 5 out of the 10 sequences, near the genes *FAIM2*, *FTO*, *GIPR*, *MC4R* and *VPS45* showed significant differential enhancer activity with the obesity-associated variant/s compared to the reference variant (**Fig.5c**).

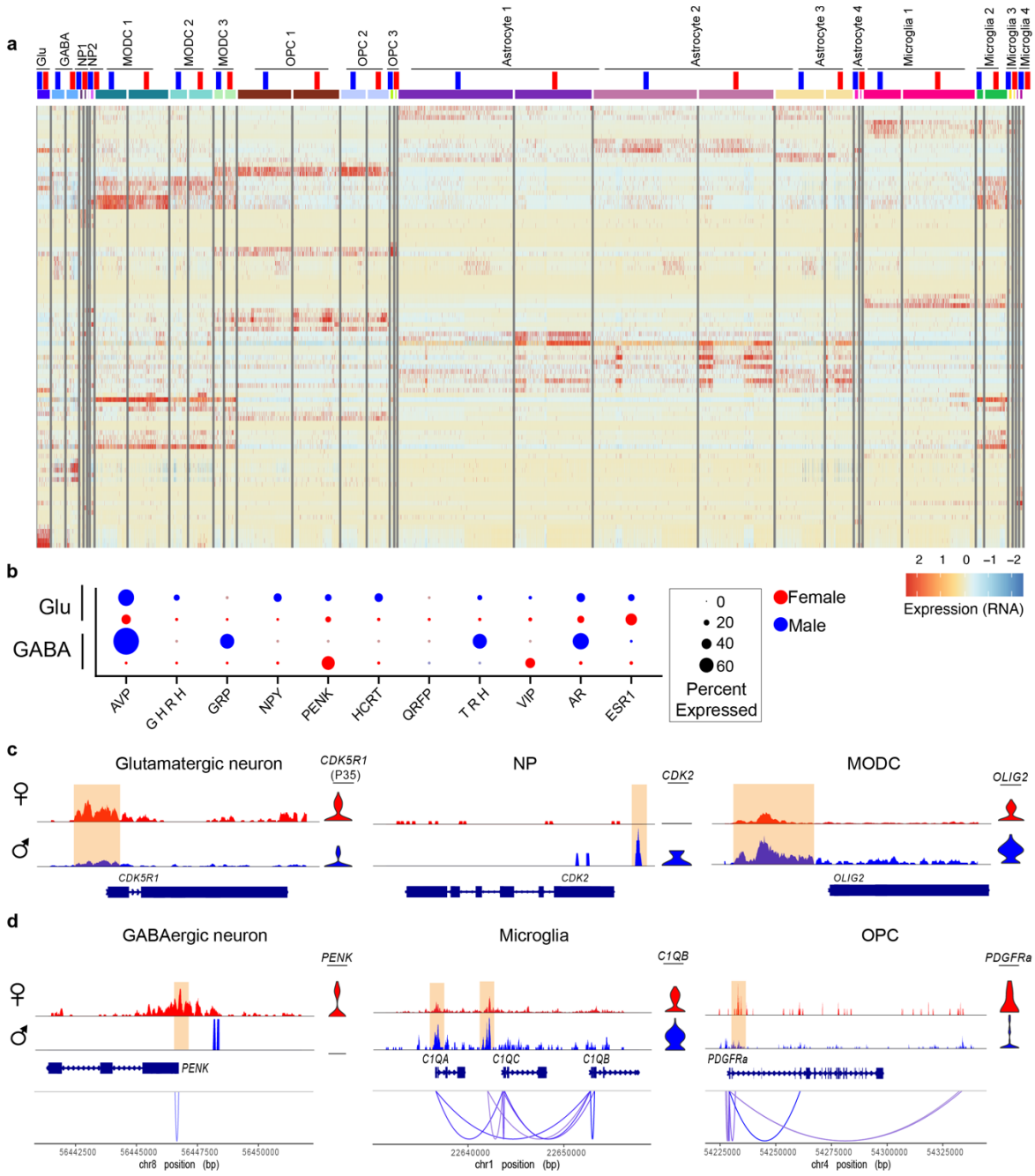


Fig 4. Human sex-differentially expressed genes and regulatory elements. a, Heatmap of the top five sex-differentially expressed genes for the corresponding cell populations in the human hypothalamus. Female cell populations are labeled in red and male cell populations in blue. **b,** Dot plot showing the expression of neuropeptides that were sex-differentially expressed in GABAergic and glutamatergic neurons. **c,** Genome tracks of sex-differentially expressed genes with sex-differential scATAC peaks at their promoters, including *CDK5R1* (P53) in glutamatergic neurons, *CDK2* in NP, and *OLIG2* in MODC clusters. **d,** Linkage plots showing sex-specific scATAC peaks linked to sex-differentially expressed genes, including *PENK* in GABAergic neurons, *C1QB* in microglia, and *PDGFRa* in OPC clusters.

The SNP that showed differential enhancer activity for *MC4R*, is the second top obesity-associated GWAS SNP and also happens to be the lead SNP, rs17782313. It is located 187k bp from *MC4R* in a sequence that is not conserved in mice. Using TRANSFAC[110], PROMO[111], and JASPAR[112], analysis of the risk allele found it to potentially disrupt binding of GATA binding protein 1 (GATA1), which was shown to bind to the 5'UTR of the *MC4R* gene[113](**Fig.5d**). To characterize its function and test whether it regulates *MC4R*, we set out to knockout this sequence in neurons using CRISPR-Cas9. As *MC4R* is not expressed in many cell types, we first differentiated WTC11-NGN2 iPSC cells[114] to functional glutamatergic neurons (see **Methods**) and tested whether it is expressed in these cells. We found *MC4R* to be significantly expressed in these cells (~8-fold compared to iPSC cells; **Fig.5f**). We then transfected these iPSC cells with two gRNAs targeting both ends of the postulated *MC4R* enhancer along with Cas9 protein. Single-cell colonies were FACS-isolated and screened by genotyping for homozygous cells, finding two colonies. These colonies, along with wild-type cells, were differentiated to neurons and analyzed for *MC4R* expression via qRT-PCR. We found that the two independent *MC4R* enhancer knockout cells lines had significantly lower *MC4R* expression (60%) compared to wild-type cells (**Fig.5f**). Since there are no other genes in the *MC4R* TAD boundary, we did not measure expression for other genes. Taken together, these results showcase that the obesity associated *MC4R* SNP reduces enhancer activity, and that removal of this enhancer leads to reduced *MC4R* expression, suggesting that it regulates this gene.

We next examined the function of the *FTO* enhancer. This sequence was previously shown to regulate its neighboring genes, Iroquois homeobox 3 and 5 (*IRX3*/*5*), in the hypothalamus[115] and adipocytes[116]. TFBS analysis for the obesity-associated SNP, rs8043757, suggested that it decreases binding affinity for the high mobility group forkhead box P3 (*FOXP3*), which was shown to regulate gonadotropin expression in pituitary in mice[117](**Fig.5e**). To further characterize its target genes in a cell-type specific manner, we attempted to delete the *FTO* enhancer in astrocytes. However, cells failed to proliferate and form colonies, likely due to the known role of *FTO* in the proliferation of neuronal cells[118]. We thus used CRISPR inactivation (CRISPRi) instead. We designed four gRNAs targeting the identified enhancer and tested them for their ability to downregulate expression by co-transfecting human astrocytes with a nuclease deficient Cas9 (dCas) fused to the KRAB transcriptional repressor. Examination by qRT-PCR found *FTO* to have significantly lower expression for three of the four gRNAs, with the lowest reduction for gRNA-4 (**Fig.5g**). We also analyzed the expression of *FTO* enhancer target genes, *IRX3* and *IRX5*, finding both to have decreased expression levels with all gRNAs, with the highest impact for gRNA-4 (**Fig.5g**). In line with the enhancer being specific to astrocytes, we found that both *IRX3* and *IRX5* were highly expressed in astrocyte populations (**Extended Data Fig.5b**). In summary, we found obesity-associated SNPs located in the enhancer regions of *MC4R* and *FTO* could potentially disrupt binding of TFs to these enhancers, lead to lower enhancer activity and subsequently lower expression of these obesity-associated genes. Combined, this work showcases the ability of integrative scRNA and scATAC-seq to identify gene regulatory elements whose alteration is associated with hypothalamus related phenotypes, such as obesity.

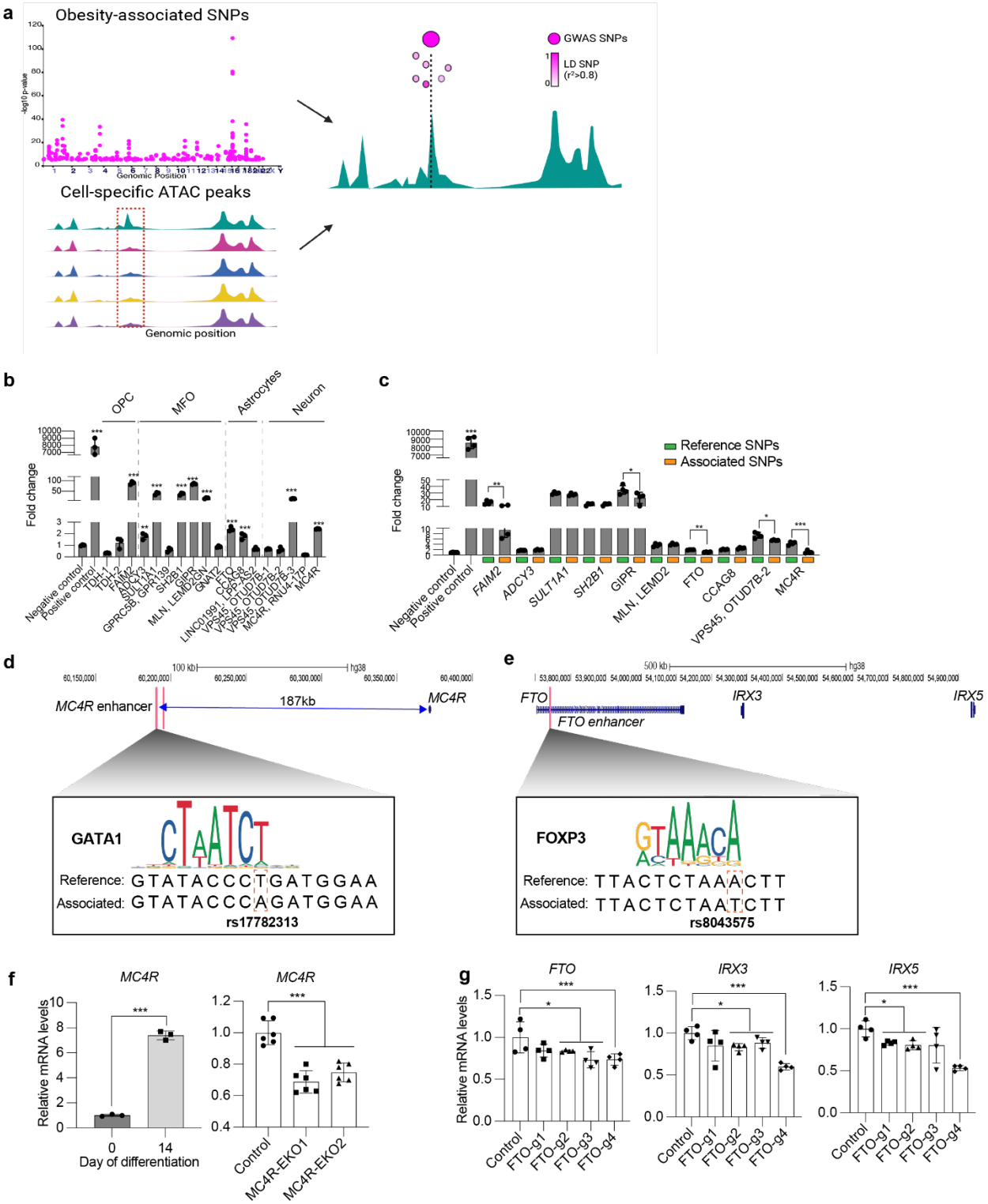


Fig 5. Integrative single cell sequencing data identifies cell-type specific gene regulatory elements overlapping obesity-associated variants in the human hypothalamus. a, Schematic of cluster-specific scATAC peaks analyses used to prioritize obesity-associated SNPs. **b-c,** Luciferase assays in mouse hypothalamus neuronal cells for sequences that overlap obesity-associated SNPs (**b**) and the reference or obesity-associated SNPs (**c**). pGL4.13 (Promega) with an SV40 early enhancer was used as

a positive control and the pGL4.23 empty vector as a negative control. **d,** Genomic snapshot of the *MC4R* locus and two identified enhancer regions with logo plots showing TFBS inference due to the obesity-associated SNP, rs17782313 **e,** Genomic snapshot of the *FTO* locus along with the *IRX3* and *IRX5* genes and the *FTO* enhancer with logo plots showing TFBS inference of HOXD10 due to the obesity-associated SNP, rs8043757. **f,** qRT-PCR of *MC4R* (left) in iPSC and differentiated neurons and control and two *MC4R* enhancer KO cell line (right). Data are represented as mean ± standard deviation, (***) P-value ≤0.001 **g,** qRT-PCR of *FTO, IRX3,* and *IRX5* in human astrocytes treated with CRISPRi and different gRNA targeting the *FTO* enhancer. Data are represented as mean ± standard deviation, (*) P-value≤0.001 (***) P-value≤0.001.

Discussion

In this study, we utilized integrative single-cell sequencing to characterize mouse and human hypothalamus. Integrative single-cell analysis allowed us to identify various neuronal and non-neuronal cell types, uncover cell-specific regulatory elements and their target genes, and elucidate the transcriptional regulation networks of these cells in the hypothalamus. This atlas identified numerous non-neuronal and neuronal cells that play key roles in controlling hypothalamus-associated physiological processes. By using both female and male hypothalamus, we were able to identify many sex-differential expressed genes and regulatory elements. This integrative data also allowed us to identify and validate cell-type specific regulatory elements that encompass obesity-associated SNPs, including *FTO* and *MC4R*, which alter their enhancer activity. Deletion of the obesity associated *MC4R* enhancer or CRISPRi targeting of the *FTO* enhancer in astrocytes validated their target genes.

Using gene markers from previous scRNA-seq studies[29-41], we identified many neuronal cell populations in mice (N=33), including glutamatergic, GABAergic, specialized neuron clusters, OPC, ODC, astrocytes, and microglia, with the majority of cells (90%) being neurons. We were also able to identify four specialized neuron cell types that were not identified in these studies. In contrast, in the human hypothalamus, neurons made up less than 5% of the total cells while the majority of cells were oligodendrocyte associated. This is in line with two recent studies in which multiomic single-cell sequencing was performed using post-mortem human midbrain or frontal lobe, showed similar recovery of neuronal and oligodendrocyte populations[63,64]. Similarly, the recent study by Siletti et al,. in which ~3.4 million nuclei from 106 sections of three human adult brains were sequenced, observed a large number of oligodendrocytes (~600K) despite the fact that they isolated neurons by FACS with NeuN (similar to what was done here) aiming to collect 90% neurons and 10% non-neuronal cells[119]. This might be due to human neurons being sensitive to nuclear permeabilization or samples being collected post-mortem when nucleus neurons are the most sensitive to hypoxia[120] as well as the age of donors (**Supplemental Table 1**).

Tissue- or cell-specific gene regulatory networks control cell fate specification and drive dynamic processes, such as cell differentiation. Integrative single-cell sequencing allowed us to identify regulatory elements and link them to their regulated genes in specific clusters and further map TFs activity in a cell type-specific manner. For example, in the mouse hypothalamus, we identified two regulatory elements linked to *Pomc* expression only in the *Pomc*/*Cartpt* neurons which have been reported to be enhancers of this gene[59]. In the human hypothalamus, we identified four regulatory elements linked to *GAD1* expression, one of which is located ~50kb from its promoter and has already been characterized as its enhancer[121]. We also identified numerous cell-specific scATAC peaks linked to expressions of cell marker genes, including *SLC32A1* in GABAergic neurons, *SLC16A7* in glutamatergic neurons, *TH* and *DDC* in dopamin neurons, and *PDGFRa* in OPC. These regulatory elements are all novel and can provide deep understanding of the transcriptional regulation of these cell type-specific marker genes. Furthermore, by performing TFBS enrichment analyses on

scATAC peaks, we identified many cell type-specific networks of TF activity. Comparison of cell-types across mice and human hypothalamus allowed us to identify shared core TFs that could serve as cell cluster-specific regulatory machinery. We also found genes that are regulated by these core TFs and had cell type-specific functions, further confirming that each cell type requires distinctive gene regulatory networks to regulate its function and activities.

We identified numerous genes that are sex-differentially expressed across various cell clusters in mouse and human hypothalamus. We found that sex effects are small, but ubiquitous across cell populations. When examining neuropeptides, we found several neuropeptide genes that were sex-differentially expressed, some of which were previously reported. For example, we found *NPY* to be more highly expressed in male glutamatergic neurons compared to females, similar to studies showing *NPY* brain expression levels are lower in females than males, contributing to higher anxiety and anxiogenic phenotype in males[122,123]. Interestingly, we also found novel sex-differentially expressed neuropeptides, such as *HCRT* in glutamatergic neurons, T*RH* in GABAergic neurons. *HCRT* was shown to have sex-differential effects on body weight and body composition when ablated in mice, however, its expression was not reported[124]. This sex-differential effects might be due to its sex-differential expression levels in the hypothalamus. Similarly, *TRH*, which stimulates the release of thyroid stimulating hormone (*TSH*) and prolactin from the anterior pituitary[67], was not shown to be differentially expressed between sexes in GABAergic neurons. Furthermore, we found sex-specific regulatory elements that are linked to sex-differentially expressed genes. For examples, in GABAergic neurons, we found a male-specific scATAC peak was linked to the expression of proenkephalin (*PENK*), which was shown to be involved in pain perception and response to stress in males rats[105]. Sex-dimorphic expression of some of these key neuropeptides might be involved in sex-differential processes, such as reproductive and social behaviors, and physiological responses to environment cues.

Our integrative single-cell datasets allowed the identification of novel genes, regulatory elements and pathways associated with hypothalamus related phenotypes. Here, as an example, we showcase the advantage of this approach to identify obesity-associated regulatory elements. Numerous obesity GWAS loci reside near hypothalamus expressed genes[9,10], several of which are associated with the hypothalamic leptin-melanocortin system[10], and have been difficult to detect due to the inability to identify gene regulatory elements in distinct neuronal subpopulations. We were able to identify multiple hypothalamus cell type specific scATAC peaks that overlap with obesity-associated SNPs, including the top two obesity associated loci, *FTO* and *MC4R*. This allowed us to pinpoint the cell types in which these potential obesity-associated regulatory elements are involved, finding them to be distributed in various cell populations, including OPC, ODC, astrocyte, and neurons. Interestingly, we found no potential obesity-associated enhancers in the microglia population. Utilizing enhancer assays, we found that more than half of these overlapping scATAC peaks have enhancer activity in either neurons or astrocytes. Additional cell types and or primary cells could also be used to dissect the function of these sequences. Amongst the active sequences, we found that half of them altered enhancer activity. Interestingly, in all

these cases this led to a reduction rather than an increase in enhancer activity. Our CRISPR experiments for *MC4R* and *FTO* in iPSC-derived neurons and primary astrocytes further confirmed the target genes of these enhancers, showcasing the utility of these datasets.

Methods

Nuclei isolation and library preparation

For mice, the whole hypothalamus of three male and three female 6 weeks-old *Mc4r*$^{t2aCre/t2aCre}$ *x Gt(ROSA)26Sor*$^{tm2(CAG\text{-}NuTRAP)Evdr}$ C57BL/6J mice generated by Dr. Vaisse[44] were subjected to nuclei isolation following the 10X Genomics established protocol (CG000375-Rev A). In brief, flash-frozen hypothalamus samples were lysed and homogenized in 0.1% NP40 lysis buffer supplemented with RNase inhibitor (Sigma, 3335402001). The nuclei mixture was filtered through a 70um strainer and then stained with 7-AAD (7-aminoactinomycin D) in PBS with 1% BSA and RNase inhibitor. Nuclei were FACS-sorted using the ARIA Fusion Cell Sorter with a 100um nozzle. Nuclei were washed and pre-metallized with 0.01% digitonin lysis buffer supplemented with RNase inhibitor. For each sample, 20,000 nuclei were subjected for GEX and ATAC libraries prepared using Chromium Next GEM Single Cell Multiome ATAC + Gene Expression (CG000338 Rev B). Both libraries were sequenced using HiSeq 4000 with 25,000 paired reads per nucleus. For human samples, three male and three female hypothalami were collected post-mortem at UCSF medical center following the UCSF human research protection program institutional review board protocol number 21-34261. The donors' age ranged from 43-82 (**Supplementary Table 1**). The human samples were subjected to the same protocols as mouse hypothalamus. Due to the large size of the human sample, the hypothalamus was split in half and subjected to the lysis step and pooled together prior to FACS. 40,000 nuclei per sample were used for the library prep.

Data preprocessing and quality control

Demultiplexed scRNA- and scATAC-seq fastq files were generated using CellRanger Arc mkfastq (10x Genomics, 7.0.1). scRNA and scATAC-seq data were aligned to pre-built reference genomes GRCh38 and GRCm38 from 10x Genomics. Barcoded count matrices of gene expression and ATAC data were generated using CellRanger ARC pipeline (version 2.0.0) (10X Genomics). Count and peak matrices and fragment files from CellRanger were analyzed using the 'Seurat' package (version 4.1.1)[125]. Quality cells were selected using the following quality control metrics: TSS enrichment greater than 1; nucleosome signal greater than 2; fraction of mitochondrial genes less than 30%; between 100 and 7500 genes detected in each cell; and between 1000 and 30000 peaks with at least one readcount detected in each cell. For the human dataset, contaminating ambient RNA was detected using SoupX and the corrected count matrix used for downstream analysis. Gene annotations from Ensembl Database v79 (mouse) and Ensembl Database v86 were used.

scRNA-seq data analysis

scRNA-seq data of nuclei were analyzed using 'Seurat' after quality control filtering. Gene expression data were log-normalized and multiplied by 10,000 using the NormalizeData and ScaleData functions. The top 3,000 variable features were identified using FindVariableFeatures function. Batch-correction between samples was performed using Harmony[126]. Dimensionality reduction was performed with PCA and the top 50 principal components were kept. Nearest neighbors used Harmony embeddings, and

clustering was performed using resolution = 0.6. Differentially expressed genes were identified using[127] Seurat's FindAllMarkers function[125], using Wilcoxon Rank Sum test, filtering for a minimum log2 fold change threshold of 0.25, and minimum fraction of 0.1 cells in the population.

scATAC-seq data analysis

scATAC-seq data were processed using 'Signac' R package (version 1.7.0). We apply term frequency-inverse document frequency normalization to ATAC peaks, followed by feature selection and dimensionality reduction using singular value decomposition. Batch-correction was performed using Harmony. 50 dimensions for ATAC and 50 dimensions RNA PCA dimensions were used to construct the weighted nearest neighbor graph. Peak calling was performed using MACS2[128], and peaks not found on standard chromosomes and within blacklist regions were pruned from the peak set. Differentially accessible regions were identified using FindAllMarkers function, using likelihood ratio test, filtering for a minimum log2 fold change of 0.25, minimum fraction of 0.05 cells in the population, including number of ATAC counts as latent variable. Linked peaks were identified using Signac's LinkPeaks function using peaks found in at least 3 cells, distance cutoff of 1 million basepairs, p-value cutoff of 0.05. Nearest genes to scATAC-seq peaks were determined using ClosestFeatures function.

Motif enrichment analysis

Motif analysis was performed on accessible peaks using ChromVar[72]. Motif position weight matrices were downloaded from the JASPAR 2020 database. Differential testing on chromVar Z-scores was performed using the FindAllMarkers function (Wilcoxon Rank Sum test) in Signac with minimum fraction of cells as 0.1. Peaks containing motifs were identified using the `motifmatchr` version 1.18.0[129]. GO terms were identified using GENEontology.

SNPs selection and overlapped with cell type-specific ATAC peaks

Indexed SNPs associated with obesity were collected from the NHGRI-EBI GWAS Catalog[107]. LD expansion was then performed using LDLink[130] to identify LD linked SNPs with r2>=0.8. LD linked SNPs were filtered to exclude variants in coding regions based on annotations in dbSNP to obtain a final set of noncoding SNPs. The coordinates of these SNPs were intersected with differentially accessible ATAC peaks. P-values for SNP enrichment in ATAC peaks were computed using the binomial test as implemented in SciPy. We set the number of successes as SNPs that overlapped the ATAC peaks. We set the number of trials as number of SNPs from obesity collected as described above. The coordinates of these SNPs were intersected with differentially accessible ATAC peaks, found using Seurat's 'FindAllMarkers' function using likelihood ratio test with log2fold change threshold > 0.15. 'motifbreakR' package[131] was used to predict effects of SNPs on motifs on its surrounding region.

Luciferase assay

The scATAC-seq peaks that overlapped obesity-associated SNPs were PCR amplified from human genomic DNA (see primers in **Supplementary Table 2**) and cloned into the

pGL4.23 plasmid (Promega, E84111). The associated SNPs were then introduced into these plasmids by PCR amplification with primers containing the associated variants. The PCR products were then treated with DpnI to remove WT and the constructs with the associated SNPs were confirmed with Sanger sequencing. The These unassociated and associated constructs (200ng) along with Renilla luciferase (20ng), to correct for the transfection efficiency, were transfected into mouse hypothalamus Pomc neuronal cells (mHypoA-POMC/GFP-1 from Sanbio #CLU500) or human astrocytes (CCF-STTG, ATCC) grown in 96-well plates of [132] using X-tremeGene (Sigma, 6366244001) following the manufacturer's protocol. An empty pGL4.23 vector was used as negative control and pGL4.13 (Promega, E668A), which has an SV40 early enhancer, as a positive control. Forty-eight hours post transfection, cells were lysed, and luciferase activity was measured using the Dual-Luciferase Reporter Assay System (Promega, E1910). Six technical replicates were performed for each condition.

Neuronal differentiation

WTC11 iPSC cells (a generous gift from Li Gan (Gladstone Institute)) were maintained in Matrigel-coated plates with mTeSR1 basal media (StemCell, 85850). Cells were subcultured using Accutase and plated with mTeSR1 media supplemented with 10uM Rock inhibitor (Fisher Scientific, NC1286855). To differentiate these cells to mature neurons, we followed the previously published protocol in [114]. In brief, cells were plated in Matrigel-coated 6-well plate in pre-differentiation media (see recipe in **Supplementary Table 3**) and Rock inhibitor. Media was changed every day for the next 2 days. On day 4, cells were then sub-cultured into poly-L-Ornithine-coated plates (Sigma, P3665) in maturation media (**Supplementary Table 3**) supplemented with 2ug/mL doxycycline (Sigma, D3447). Half of the media was then replaced with maturation media after 7 days. Cells were collected at day 14 of differentiation.

MC4R enhancer knockout cells were generated by transfecting WTC11-NGN2 iPSC cells in 6-well plates with 6.25ug Cas9 protein (Fisher Scientific, A36498) and 800ng sgRNAs (IDT), and 0.5ug GFP plasmid (Addgene, 13031) using Lipofectamine CRISPRMax Cas9 transfection reagent (Fisher Scientific, CMAX00-003) following the manufacturer's protocol. After 48 hours, GFP+ cells were isolated into 96 well-plates into single clones using FACS (BD FACSAria Fusion) (see gRNAs in **Supplementary Table 2**). These colonies were then genotyped by PCR with primers flanking 200bp the deletion site. The knockout genotype has 400 bp PCR product while WT has 3500 bp PCR product (see primers in **Supplementary Table 2**). Cells were then subjected to the neuron differentiation protocol described above.

RNA Isolation, cDNA synthesis and qRT-PCR

Total RNA was extracted from cells using RNeasy Plus kit (Qiagen, 74106). Reverse transcription was performed with 1µg of total RNA using qScript cDNA Synthesis Kit (Quantabio, 95047) following the manufacturer's protocol. qRT-PCR was performed on QuantStudio 6 Real Time PCR system (ThermoFisher) using Sso Fast (Biorad, 1705205). Statistical analysis was done using ddct method with GAPDH primers as control (see primer sequences in **Supplementary Table 2**). Gene expression results were generated using mean values for over 4-6 biological replicates.

CRISPRi study

gRNAs targeting the *FTO* enhancer were designed using CRISPick (Broad Institute)[133] (see gRNAs in **Supplementary Table 2**). gRNAs were cloned into an AAV vector (pAAV-U6-sasgRNA-CMV-mCherry-WPREpA) and co-transfected into human astrocyte cells (CCF-STTG1, ATCC) along with dCas9-KRAB (kind gift from Dr. Alejandro Lomniczi OHSU). After two days, cells were lysed, and RNA and cDNA were prepared as mentioned above.

Acknowledgments We would like to thank Dr. Alejandro Lomniczi for kindly providing the pCVM-sadCas9-KRAB vector. This work was funded in part by the National Institute of Diabetes and Digestive and Kidney Disease (NIDDK) R01DK116738 (C.V. and N.A), R01DK106404 and R01DK060540 (CV), the UCSF Nutrition and Obesity Research Center P30 DK098722, the California Institute for Regenerative Medicine (CIRM) postdoctoral fellowship (H.P.N.), the UCSF Hillblom Center for the Biology of Aging and Bakar Aging Research Institute Graduate Fellowship (C.S.Y.C), a Societe Francaise du Diabete and a fondation Servier fellowship (FM), Vera Long Fellowship (AB), and start-up funds from the Evergrande Center (M.H).

Author contributions H.P.N., N.A. and CV conceived and designed the study. H.P.N., D.L.R., R.S., L.H., M.N., A.U., C.B., K.A. performed experiments., H.C., C.S.Y.C., M.G.G., M.H. analyzed genomic data, F.M., A.D., and C.V. provided mice hypothalami, E.H. provide human samples, H.P.N., C.S.Y.C. and N.A. wrote the manuscript with contributions from all authors.

Competing interests N.A. is a cofounder and on the scientific advisory board of Regel Therapeutics. N.A. receives funding from BioMarin Pharmaceutical Incorporate.

Data and code availability Raw and processed data from this study have been deposited in the NCBI short read archive (SRA) as SRA12255727and SRA12255698. Code used for scRNA-and scATAC-seq analysis is available at https://github.com/candacechan/sc_hypothalamus.

References

1 Berrington de Gonzalez, A. *et al.* Body-mass index and mortality among 1.46 million white adults. *The New England journal of medicine* **363**, 2211-2219, doi:10.1056/NEJMoa1000367 (2010).

2 Hales, C. M., Fryar, C. D., Carroll, M. D., Freedman, D. S. & Ogden, C. L. Trends in Obesity and Severe Obesity Prevalence in US Youth and Adults by Sex and Age, 2007-2008 to 2015-2016. *JAMA* **319**, 1723-1725, doi:10.1001/jama.2018.3060 (2018).

3 Swinburn, B. A. *et al.* The global obesity pandemic: shaped by global drivers and local environments. *The Lancet* **378**, 804-814, doi:10.1016/s0140-6736(11)60813-1 (2011).

4 Silventoinen, K., Rokholm, B., Kaprio, J. & Sørensen, T. I. A. The genetic and environmental influences on childhood obesity: a systematic review of twin and adoption studies. *International Journal of Obesity* **34**, 29-40, doi:10.1038/ijo.2009.177 (2009).

5 Stunkard, A. J. A Twin Study of Human Obesity. *JAMA: The Journal of the American Medical Association* **256**, 51, doi:10.1001/jama.1986.03380010055024 (1986).

6 Ranadive, S. A. & Vaisse, C. Lessons from extreme human obesity: monogenic disorders. *Endocrinol Metab Clin North Am* **37**, 733-x, doi:10.1016/j.ecl.2008.07.003 (2008).

7 Speliotes, E. K. *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature genetics* **42**, 937-948, doi:10.1038/ng.686 (2010).

8 Thorleifsson, G. *et al.* Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nature Genetics* **41**, 18-24, doi:10.1038/ng.274 (2008).

9 Loos, R. J. The genetics of adiposity. *Curr Opin Genet Dev* **50**, 86-95, doi:10.1016/j.gde.2018.02.009 (2018).

10 Ghosh, S. & Bouchard, C. Convergence between biological, behavioural and genetic determinants of obesity. *Nature Reviews Genetics* **18**, 731-748, doi:10.1038/nrg.2017.72 (2017).

11 Loos, R. J. F. & Yeo, G. S. H. The genetics of obesity: from discovery to biology. *Nat Rev Genet* **23**, 120-133, doi:10.1038/s41576-021-00414-z (2022).

12 Inoue, F. *et al.* Genomic and epigenomic mapping of leptin-responsive neuronal populations involved in body weight regulation. *Nat Metab* **1**, 475-484, doi:10.1038/s42255-019-0051-x (2019).

13 Burbridge, S., Stewart, I. & Placzek, M. in *Comprehensive Physiology* 623-643 (Wiley, 2016).

14 Daniel, P. M. Anatomy of the hypothalamus and pituitary gland. *J Clin Pathol Suppl (Assoc Clin Pathol)* **7**, 1-7, doi:10.1136/jcp.s1-7.1.1 (1976).

15 Flament-Durand, J. Observations on Pituitary Transplants Into the Hypothalamus of the Rat. *Endocrinology* **77**, 446-454, doi:10.1210/endo-77-3-446 (1965).

16 Bligh, J. THE THERMOSENSITIVITY OF THE HYPOTHALAMUS AND THERMOREGULATION IN MAMMALS. *Biological Reviews* **41**, 317-365, doi:10.1111/j.1469-185x.1966.tb01496.x (1966).
17 Gao, H.-L. *et al.* Astaxanthin Ameliorates Blood Pressure in Salt-Induced Prehypertensive Rats Through ROS/MAPK/NF-κB Pathways in the Hypothalamic Paraventricular Nucleus. *Cardiovascular Toxicology* **21**, 1045-1057, doi:10.1007/s12012-021-09695-6 (2021).
18 Horio, N. & Liberles, S. D. Hunger enhances food-odour attraction through a neuropeptide Y spotlight. *Nature* **592**, 262-266, doi:10.1038/s41586-021-03299-4 (2021).
19 Siemian, J. N., Arenivar, M. A., Sarsfield, S. & Aponte, Y. Hypothalamic control of interoceptive hunger. *Curr Biol* **31**, 3797-3809.e3795, doi:10.1016/j.cub.2021.06.048 (2021).
20 Tan, C. L. & Knight, Z. A. Regulation of Body Temperature by the Nervous System. *Neuron* **98**, 31-48, doi:10.1016/j.neuron.2018.02.022 (2018).
21 Zhang, S. X. *et al.* Hypothalamic dopamine neurons motivate mating through persistent cAMP signalling. *Nature* **597**, 245-249, doi:10.1038/s41586-021-03845-0 (2021).
22 Atasoy, D., Betley, J. N., Su, H. H. & Sternson, S. M. Deconstruction of a neural circuit for hunger. *Nature* **488**, 172-177, doi:10.1038/nature11270 (2012).
23 Heck, A. L. & Handa, R. J. Sex differences in the hypothalamic-pituitary-adrenal axis' response to stress: an important role for gonadal hormones. *Neuropsychopharmacology* **44**, 45-58, doi:10.1038/s41386-018-0167-9 (2019).
24 Seale, J. V., Wood, S. A., Atkinson, H. C., Harbuz, M. S. & Lightman, S. L. Gonadal Steroid Replacement Reverses Gonadectomy-Induced Changes in the Corticosterone Pulse Profile and Stress-Induced Hypothalamic-Pituitary-Adrenal Axis Activity of Male and Female Rats. *Journal of Neuroendocrinology* **16**, 989-998, doi:10.1111/j.1365-2826.2004.01258.x (2004).
25 Seale, J. V., Wood, S. A., Atkinson, H. C., Lightman, S. L. & Harbuz, M. S. Organizational Role for Testosterone and Estrogen on Adult Hypothalamic-Pituitary-Adrenal Axis Activity in the Male Rat. *Endocrinology* **146**, 1973-1982, doi:10.1210/en.2004-1201 (2005).
26 Allen, L. S., Hines, M., Shryne, J. E. & Gorski, R. A. Two sexually dimorphic cell groups in the human brain. *J Neurosci* **9**, 497-506, doi:10.1523/JNEUROSCI.09-02-00497.1989 (1989).
27 Bailey, M. & Silver, R. Sex differences in circadian timing systems: implications for disease. *Front Neuroendocrinol* **35**, 111-139, doi:10.1016/j.yfrne.2013.11.003 (2014).
28 Stamatiades, G. A. & Kaiser, U. B. Gonadotropin regulation by pulsatile GnRH: Signaling and gene expression. *Molecular and cellular endocrinology* **463**, 131-141, doi:10.1016/j.mce.2017.10.015 (2018).
29 Bai, L. *et al.* Genetic Identification of Vagal Sensory Neurons That Control Feeding. *Cell* **179**, 1129-1143.e1123, doi:10.1016/j.cell.2019.10.031 (2019).
30 Chen, R., Wu, X., Jiang, L. & Zhang, Y. Single-Cell RNA-Seq Reveals Hypothalamic Cell Diversity. *Cell reports* **18**, 3227-3241, doi:10.1016/j.celrep.2017.03.004 (2017).

31 Huang, W.-K. *et al.* Generation of hypothalamic arcuate organoids from human induced pluripotent stem cells. *Cell stem cell* **28**, 1657-1670.e1610, doi:10.1016/j.stem.2021.04.006 (2021).
32 Khrameeva, E. *et al.* Single-cell-resolution transcriptome map of human, chimpanzee, bonobo, and macaque brains. *Genome Res* **30**, 776-789, doi:10.1101/gr.256958.119 (2020).
33 Kim, D. W. *et al.* The cellular and molecular landscape of hypothalamic patterning and differentiation from embryonic to late postnatal development. *Nature communications* **11**, 4360-4360, doi:10.1038/s41467-020-18231-z (2020).
34 Mickelsen, L. E. *et al.* Single-cell transcriptomic analysis of the lateral hypothalamic area reveals molecularly distinct populations of inhibitory and excitatory neurons. *Nature neuroscience* **22**, 642-656, doi:10.1038/s41593-019-0349-8 (2019).
35 Moffitt, J. R. *et al.* Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science (New York, N.Y.)* **362**, eaau5324, doi:10.1126/science.aau5324 (2018).
36 Ren, J. *et al.* Single-cell transcriptomes and whole-brain projections of serotonin neurons in the mouse dorsal and median raphe nuclei. *Elife* **8**, e49424, doi:10.7554/eLife.49424 (2019).
37 Romanov, R. A. *et al.* Molecular design of hypothalamus development. *Nature* **582**, 246-252, doi:10.1038/s41586-020-2266-0 (2020).
38 Wen, S. a. *et al.* Spatiotemporal single-cell analysis of gene expression in the mouse suprachiasmatic nucleus. *Nature Neuroscience* **23**, 456-467, doi:10.1038/s41593-020-0586-x (2020).
39 Yu, H., Rubinstein, M. & Low, M. J. Developmental single-cell transcriptomics of hypothalamic POMC neurons reveal the genetic trajectories of multiple neuropeptidergic phenotypes. *Elife* **11**, e72883, doi:10.7554/eLife.72883 (2022).
40 Zhou, X. *et al.* Deciphering the spatial-temporal transcriptional landscape of human hypothalamus development. *Cell Stem Cell* **29**, 328-343.e325, doi:10.1016/j.stem.2021.11.009 (2022).
41 Steuernagel, L. *et al.* HypoMap-a unified single-cell gene expression atlas of the murine hypothalamus. *Nat Metab* **4**, 1402-1419, doi:10.1038/s42255-022-00657-y (2022).
42 Wu, J., Zhao, Q., Kurosaki, T. & Weiss, A. The Vav binding site (Y315) in ZAP-70 is critical for antigen receptor-mediated signal transduction. *J Exp Med* **185**, 1877-1882, doi:10.1084/jem.185.10.1877 (1997).
43 Jiao, H. *et al.* Genome-Wide Interaction and Pathway Association Studies for Body Mass Index. *Front Genet* **10**, 404, doi:10.3389/fgene.2019.00404 (2019).
44 Wang, Y. *et al.* Melanocortin 4 receptor signals at the neuronal primary cilium to control food intake and body weight. *J Clin Invest* **131**, doi:10.1172/JCI142064 (2021).
45 Govaerts, C. *et al.* Obesity-associated mutations in the melanocortin 4 receptor provide novel insights into its function. *Peptides* **26**, 1909-1919, doi:10.1016/j.peptides.2004.11.042 (2005).

46 Lubrano-Berthelier, C. *et al.* Molecular Genetics of Human Obesity-Associated MC4R Mutations. *Annals of the New York Academy of Sciences* **994**, 49-57, doi:10.1111/j.1749-6632.2003.tb03161.x (2003).
47 Lin, S., Storlien, L. H. & Huang, X.-F. Leptin receptor, NPY, POMC mRNA expression in the diet-induced obese mouse brain. *Brain Research* **875**, 89-95, doi:10.1016/s0006-8993(00)02580-4 (2000).
48 Wauman, J. Leptin receptor signaling: pathways to leptin resistance. *Frontiers in Bioscience* **16**, 2771, doi:10.2741/3885 (2011).
49 Magon, N. & Kalra, S. The orgasmic history of oxytocin: Love, lust, and labor. *Indian J Endocrinol Metab* **15 Suppl 3**, S156-161, doi:10.4103/2230-8210.84851 (2011).
50 Diniz, G. B. & Bittencourt, J. C. The Melanin-Concentrating Hormone as an Integrative Peptide Driving Motivated Behaviors. *Front Syst Neurosci* **11**, 32, doi:10.3389/fnsys.2017.00032 (2017).
51 Ganjavi, H. & Shapiro, C. M. Hypocretin/Orexin: A Molecular Link Between Sleep, Energy Regulation, and Pleasure. *The Journal of Neuropsychiatry and Clinical Neurosciences* **19**, 413-419, doi:10.1176/jnp.2007.19.4.413 (2007).
52 Mignot, E. Sleep, sleep disorders and hypocretin (orexin). *Sleep Medicine* **5**, S2-S8, doi:10.1016/s1389-9457(04)90001-9 (2004).
53 Ashley Grossman, M. O. & Besser, S. G. M. 8 Growth hormone releasing hormone. *Clinics in Endocrinology and Metabolism* **15**, 607-627, doi:10.1016/s0300-595x(86)80012-3 (1986).
54 García-Tornadu, I. *et al.* in *Frontiers of Hormone Research* 59-69 (S. Karger AG, 2010).
55 Bartter, F. C. Vasopressin and Blood Pressure. *New England Journal of Medicine* **304**, 1097-1098, doi:10.1056/nejm198104303041809 (1981).
56 Shiba, K., Kageyama, H., Takenoya, F. & Shioda, S. Galanin-like peptide and the regulation of feeding behavior and energy metabolism. *FEBS Journal* **277**, 5006-5013, doi:10.1111/j.1742-4658.2010.07933.x (2010).
57 Ressler, K. J. *et al.* Post-traumatic stress disorder is associated with PACAP and the PAC1 receptor. *Nature* **470**, 492-497, doi:10.1038/nature09856 (2011).
58 Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* **16**, 1289-1296, doi:10.1038/s41592-019-0619-0 (2019).
59 de Souza, F. S. *et al.* Identification of neuronal enhancers of the proopiomelanocortin gene by transgenic mouse analysis and phylogenetic footprinting. *Mol Cell Biol* **25**, 3076-3086, doi:10.1128/MCB.25.8.3076-3086.2005 (2005).
60 Baldini, G. & Phelan, K. D. The melanocortin pathway and control of appetite-progress and therapeutic implications. *J Endocrinol* **241**, R1-R33, doi:10.1530/JOE-18-0596 (2019).
61 Biglari, N. *et al.* Functionally distinct POMC-expressing neuron subpopulations in hypothalamus revealed by intersectional targeting. *Nat Neurosci* **24**, 913-929, doi:10.1038/s41593-021-00854-0 (2021).

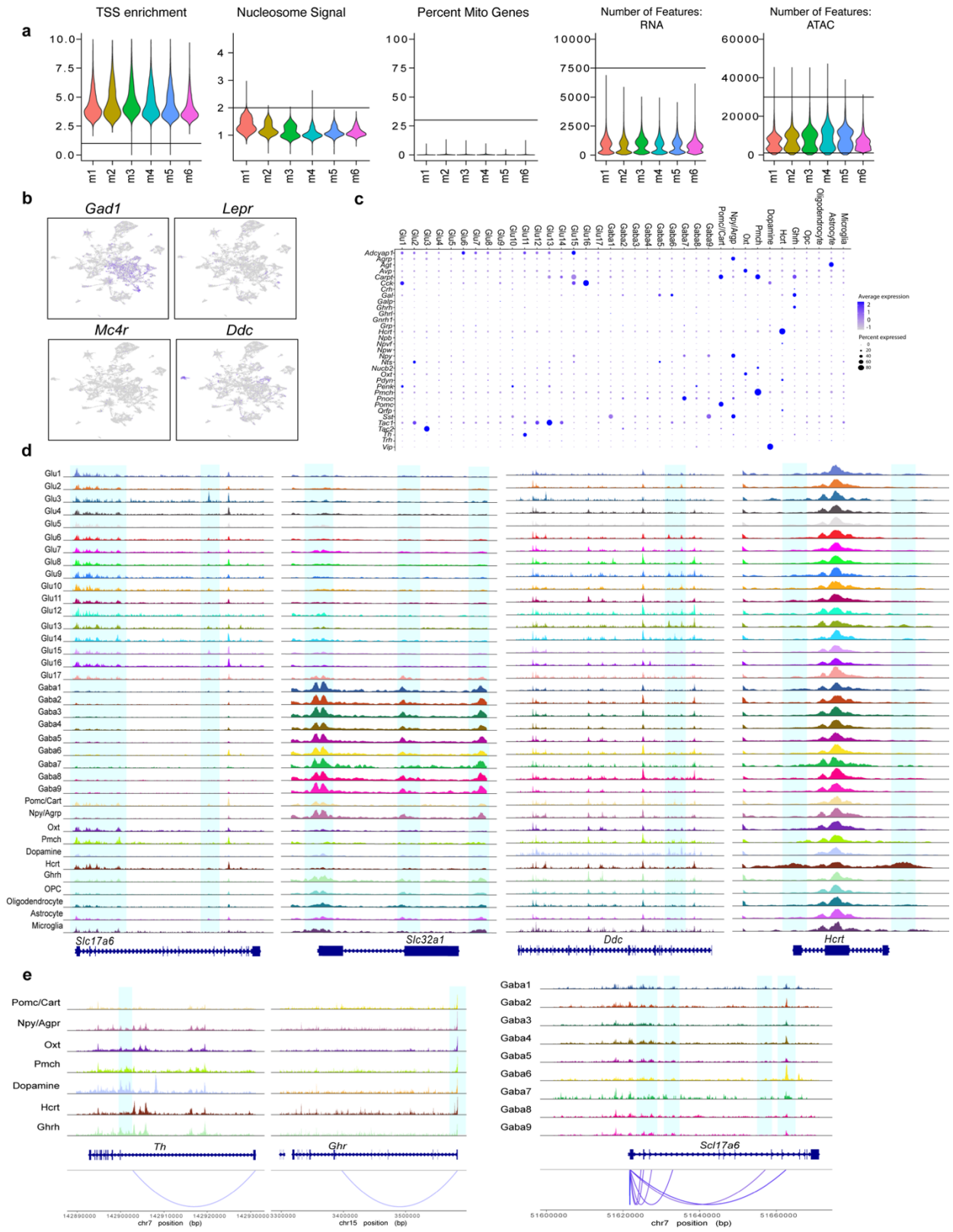
62 Young, M. D. & Behjati, S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *Gigascience* **9**, doi:10.1093/gigascience/giaa151 (2020).
63 Brase, L. *et al.* A landscape of the genetic and cellular heterogeneity in Alzheimer disease. *medRxiv*, 2021.2011.2030.21267072, doi:10.1101/2021.11.30.21267072 (2021).
64 Adams, L., Song, M. K., Tanaka, Y. & Kim, Y.-s. Single-nuclei paired multiomic analysis of young, aged, and Parkinson's disease human midbrain reveals age-associated glial changes and their contribution to Parkinson's disease. *medRxiv*, 2022.2001.2018.22269350, doi:10.1101/2022.01.18.22269350 (2022).
65 Moeller, S. J. *et al.* Effects of an opioid (proenkephalin) polymorphism on neural response to errors in health and cocaine use disorder. *Behav Brain Res* **293**, 18-26, doi:10.1016/j.bbr.2015.07.004 (2015).
66 Trivedi, C. *et al.* Tachykinin-1 in the central nervous system regulates adiposity in rodents. *Endocrinology* **156**, 1714-1723, doi:10.1210/en.2014-1781 (2015).
67 Daimon, C. M., Chirdon, P., Maudsley, S. & Martin, B. The role of Thyrotropin Releasing Hormone in aging and neurodegenerative diseases. *Am J Alzheimers Dis (Columbia)* **1**, doi:10.7726/ajad.2013.1003 (2013).
68 Chen, R., Wu, X., Jiang, L. & Zhang, Y. Single-Cell RNA-Seq Reveals Hypothalamic Cell Diversity. *Cell Rep* **18**, 3227-3241, doi:10.1016/j.celrep.2017.03.004 (2017).
69 Grieve, D. A., Chen, S. C., Chapuis, P. H. & Bradbury, R. Streptococcus bovis bacteraemia: its significance for the colorectal surgeon. *Aust N Z J Surg* **60**, 550-552, doi:10.1111/j.1445-2197.1990.tb07425.x (1990).
70 Jurga, A. M., Paleczna, M. & Kuter, K. Z. Overview of General and Discriminating Markers of Differential Microglia Phenotypes. *Front Cell Neurosci* **14**, 198, doi:10.3389/fncel.2020.00198 (2020).
71 McArthur, E. & Capra, J. A. Topologically associating domain boundaries that are stable across diverse cell types are evolutionarily constrained and enriched for heritability. *Am J Hum Genet* **108**, 269-283, doi:10.1016/j.ajhg.2021.01.001 (2021).
72 Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods* **14**, 975-978, doi:10.1038/nmeth.4401 (2017).
73 Vella, K. R., Burnside, A. S., Brennan, K. M. & Good, D. J. Expression of the hypothalamic transcription factor Nhlh2 is dependent on energy availability. *J Neuroendocrinol* **19**, 499-510, doi:10.1111/j.1365-2826.2007.01556.x (2007).
74 Leon, S. *et al.* Sex-specific pubertal and metabolic regulation of Kiss1 neurons via Nhlh2. *Elife* **10**, doi:10.7554/eLife.69765 (2021).
75 Hisaoka, T., Nakamura, Y., Senba, E. & Morikawa, Y. The forkhead transcription factors, Foxp1 and Foxp2, identify different subpopulations of projection neurons in the mouse cerebral cortex. *Neuroscience* **166**, 551-563, doi:10.1016/j.neuroscience.2009.12.055 (2010).
76 Tutukova, S., Tarabykin, V. & Hernandez-Miranda, L. R. The Role of Neurod Genes in Brain Development, Function, and Disease. *Front Mol Neurosci* **14**, 662774, doi:10.3389/fnmol.2021.662774 (2021).

77 Suzuki, T. *et al.* CUX2 deficiency causes facilitation of excitatory synaptic transmission onto hippocampus and increased seizure susceptibility to kainate. *Sci Rep* **12**, 6505, doi:10.1038/s41598-022-10715-w (2022).
78 Pristera, A. *et al.* Transcription factors FOXA1 and FOXA2 maintain dopaminergic neuronal properties and control feeding behavior in adult mice. *Proc Natl Acad Sci U S A* **112**, E4929-4938, doi:10.1073/pnas.1503911112 (2015).
79 Stolt, C. C. *et al.* The Sox9 transcription factor determines glial fate choice in the developing spinal cord. *Genes & development* **17**, 1677-1689, doi:10.1101/gad.259003 (2003).
80 Cakir, B. *et al.* Expression of the transcription factor PU.1 induces the generation of microglia-like cells in human cortical organoids. *Nat Commun* **13**, 430, doi:10.1038/s41467-022-28043-y (2022).
81 Kim, S. *et al.* Tcf3 function is required for the inhibition of oligodendroglial fate specification in the spinal cord of zebrafish embryos. *Mol Cells* **32**, 383-388, doi:10.1007/s10059-011-0152-1 (2011).
82 Tchieu, J. *et al.* NFIA is a gliogenic switch enabling rapid derivation of functional human astrocytes from pluripotent stem cells. *Nat Biotechnol* **37**, 267-275, doi:10.1038/s41587-019-0035-0 (2019).
83 de Lombares, C. *et al.* Dlx5 and Dlx6 expression in GABAergic neurons controls behavior, metabolism, healthy aging and lifespan. *Aging (Albany NY)* **11**, 6638-6656, doi:10.18632/aging.102141 (2019).
84 Oriel, J. D. The global pattern of sexually transmitted diseases. *S Afr Med J* **61**, 993-998 (1982).
85 Wu, J. P. H. *et al.* The Transcription Factor Pou3f1 Sheds Light on the Development and Molecular Diversity of Glutamatergic Cerebellar Nuclear Neurons in the Mouse. *Front Mol Neurosci* **15**, 921901, doi:10.3389/fnmol.2022.921901 (2022).
86 Earley, A. M., Burbulla, L. F., Krainc, D. & Awatramani, R. Identification of ASCL1 as a determinant for human iPSC-derived dopaminergic neurons. *Sci Rep* **11**, 22257, doi:10.1038/s41598-021-01366-4 (2021).
87 Quintana, A. M. *et al.* Mutations in THAP11 cause an inborn error of cobalamin metabolism and developmental abnormalities. *Hum Mol Genet* **26**, 2838-2849, doi:10.1093/hmg/ddx157 (2017).
88 Wang, J. *et al.* The transcription repressor, ZEB1, cooperates with CtBP2 and HDAC1 to suppress IL-2 gene activation in T cells. *Int Immunol* **21**, 227-235, doi:10.1093/intimm/dxn143 (2009).
89 Koşar, Z. & Erbaş, A. Can the Concentration of a Transcription Factor Affect Gene Expression? *Frontiers in Soft Matter* **2**, doi:10.3389/frsfm.2022.914494 (2022).
90 Juricek, L. & Coumoul, X. The Aryl Hydrocarbon Receptor and the Nervous System. *Int J Mol Sci* **19**, doi:10.3390/ijms19092504 (2018).
91 Ma, K., Zheng, S. & Zuo, Z. The transcription factor regulatory factor X1 increases the expression of neuronal glutamate transporter type 3. *J Biol Chem* **281**, 21250-21255, doi:10.1074/jbc.M600521200 (2006).

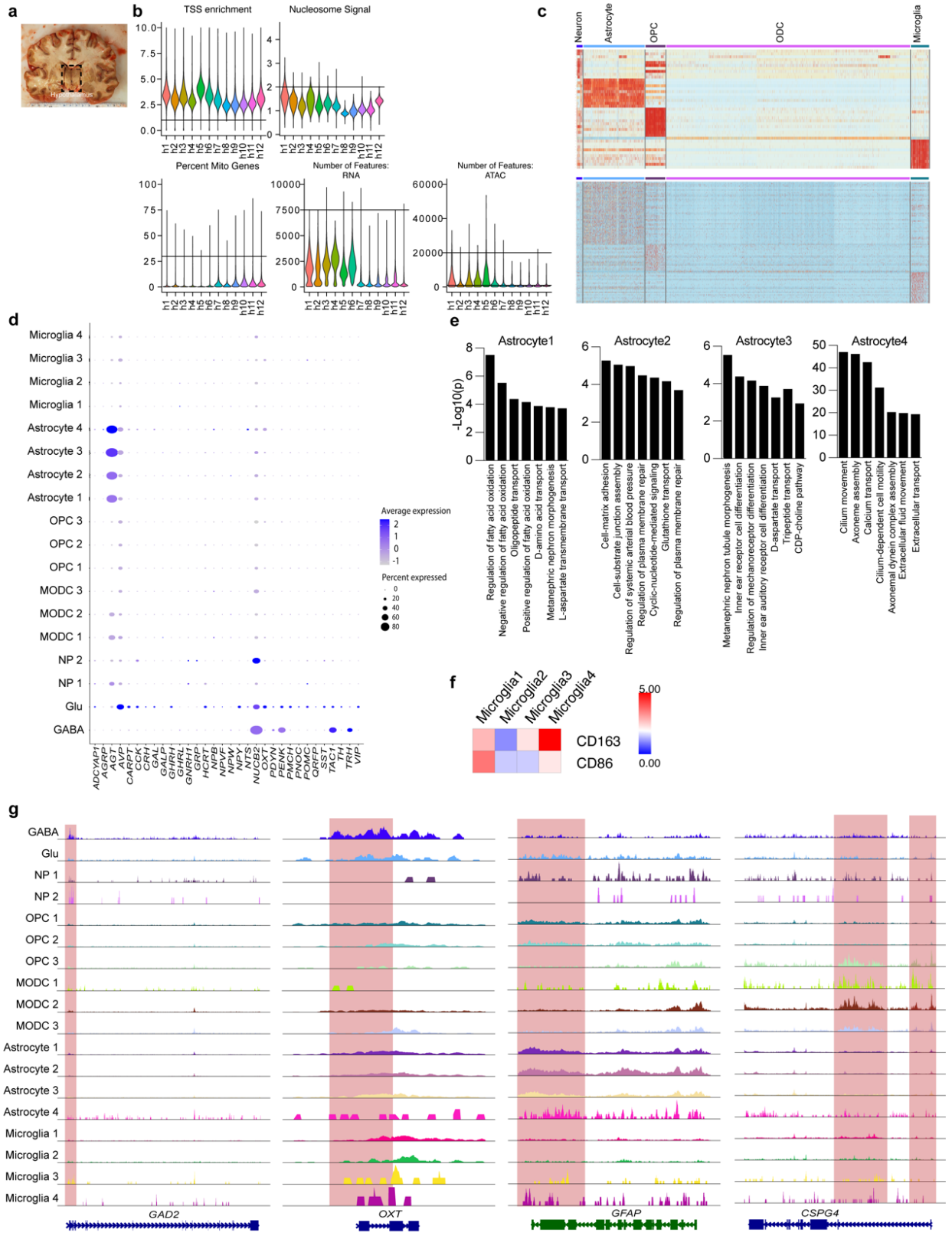
92 Feng, C., Li, J. & Zuo, Z. Expression of the transcription factor regulatory factor X1 in the mouse brain. *Folia Histochem Cytobiol* **49**, 344-351, doi:10.5603/fhc.2011.0047 (2011).
93 Galvan, L. *et al.* The striatal kinase DCLK3 produces neuroprotection against mutant huntingtin. *Brain* **141**, 1434-1454, doi:10.1093/brain/awy057 (2018).
94 Gene Ontology, C. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res* **49**, D325-D334, doi:10.1093/nar/gkaa1113 (2021).
95 de Souza, F. S. *et al.* The estrogen receptor alpha colocalizes with proopiomelanocortin in hypothalamic neurons and binds to a conserved motif present in the neuron-specific enhancer nPE2. *Eur J Pharmacol* **660**, 181-187, doi:10.1016/j.ejphar.2010.10.114 (2011).
96 Clarke, M. V., Russell, P. K., Zajac, J. D. & Davey, R. A. The androgen receptor in the hypothalamus positively regulates hind-limb muscle mass and voluntary physical activity in adult male mice. *J Steroid Biochem Mol Biol* **189**, 187-194, doi:10.1016/j.jsbmb.2019.02.018 (2019).
97 Aulino, E. A. & Caldwell, H. K. Subtle sex differences in vasopressin mRNA expression in the embryonic mouse brain. *J Neuroendocrinol* **32**, e12835, doi:10.1111/jne.12835 (2020).
98 Tong, W. H., Abdulai-Saiku, S. & Vyas, A. Arginine vasopressin in the medial amygdala causes greater post-stress recruitment of hypothalamic vasopressin neurons. *Mol Brain* **14**, 141, doi:10.1186/s13041-021-00850-2 (2021).
99 Falsetti, L. *et al.* Vasoactive intestinal polypeptide (VIP) selectively stimulates prolactin release in healthy women. *Gynecol Endocrinol* **2**, 11-18, doi:10.3109/09513598809029335 (1988).
100 Kim, Y. *et al.* Brain-wide Maps Reveal Stereotyped Cell-Type-Based Cortical Architecture and Subcortical Sexual Dimorphism. *Cell* **171**, 456-469 e422, doi:10.1016/j.cell.2017.09.020 (2017).
101 Goodwill, H. L. *et al.* Early Life Stress Drives Sex-Selective Impairment in Reversal Learning by Affecting Parvalbumin Interneurons in Orbitofrontal Cortex of Mice. *Cell Rep* **25**, 2299-2307 e2294, doi:10.1016/j.celrep.2018.11.010 (2018).
102 Jessup, S. K., Dimaraki, E. V., Symons, K. V. & Barkan, A. L. Sexual dimorphism of growth hormone (GH) regulation in humans: endogenous GH-releasing hormone maintains basal GH in women but not in men. *J Clin Endocrinol Metab* **88**, 4776-4780, doi:10.1210/jc.2003-030246 (2003).
103 Moncini, S. *et al.* Functional characterization of CDK5 and CDK5R1 mutations identified in patients with non-syndromic intellectual disability. *J Hum Genet* **61**, 283-293, doi:10.1038/jhg.2015.144 (2016).
104 Chowen, J. A. & Garcia-Segura, L. M. Role of glial cells in the generation of sex differences in neurodegenerative diseases and brain aging. *Mech Ageing Dev* **196**, 111473, doi:10.1016/j.mad.2021.111473 (2021).
105 Simerly, R. B. Prodynorphin and proenkephalin gene expression in the anteroventral periventricular nucleus of the rat: Sexual differentiation and hormonal regulation. *Mol Cell Neurosci* **2**, 473-484, doi:10.1016/1044-7431(91)90014-f (1991).

106 Prilutsky, D. *et al.* Sexual dimorphism of complement-dependent microglial synaptic pruning and other immune pathways in the developing brain. *bioRxiv*, 204412, doi:10.1101/204412 (2017).
107 Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic acids research* **47**, D1005-D1012, doi:10.1093/nar/gky1120 (2019).
108 Poveda, A., Ibáñez, M. E. & Rebato, E. Common variants in BDNF, FAIM2, FTO, MC4R, NEGR1, and SH2B1 show association with obesity-related variables in <scp>S</scp> panish <scp>R</scp> oma population. *American Journal of Human Biology* **26**, 660-669, doi:10.1002/ajhb.22576 (2014).
109 Yohn, C. N., Leithead, A. B., Ford, J., Gill, A. & Becker, E. A. Paternal Care Impacts Oxytocin Expression in California Mouse Offspring and Basal Testosterone in Female, but Not Male Pups. *Front Behav Neurosci* **12**, 181-181, doi:10.3389/fnbeh.2018.00181 (2018).
110 Matys, V. *et al.* TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res* **34**, D108-110, doi:10.1093/nar/gkj143 (2006).
111 Messeguer, X. *et al.* PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics* **18**, 333-334, doi:10.1093/bioinformatics/18.2.333 (2002).
112 Castro-Mondragon, J. A. *et al.* JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* **50**, D165-D173, doi:10.1093/nar/gkab1113 (2022).
113 Shishay, G. *et al.* Variation in the Promoter Region of the MC4R Gene Elucidates the Association of Body Measurement Traits in Hu Sheep. *Int J Mol Sci* **20**, doi:10.3390/ijms20020240 (2019).
114 Wang, C. *et al.* Scalable Production of iPSC-Derived Human Neurons to Identify Tau-Lowering Compounds by High-Content Screening. *Stem Cell Reports* **9**, 1221-1233, doi:10.1016/j.stemcr.2017.08.019 (2017).
115 Smemo, S. *et al.* Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* **507**, 371-375, doi:10.1038/nature13138 (2014).
116 Claussnitzer, M. *et al.* FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *The New England journal of medicine* **373**, 895-907, doi:10.1056/NEJMoa1502214 (2015).
117 Jung, D. O., Jasurda, J. S., Egashira, N. & Ellsworth, B. S. The forkhead transcription factor, FOXP3, is required for normal pituitary gonadotropin expression in mice. *Biol Reprod* **86**, 144, 141-149, doi:10.1095/biolreprod.111.094904 (2012).
118 Cao, Y. *et al.* Dynamic effects of Fto in regulating the proliferation and differentiation of adult neural stem cells of mice. *Hum Mol Genet* **29**, 727-735, doi:10.1093/hmg/ddz274 (2020).
119 Siletti, K. *et al.* Transcriptomic diversity of cell types across the adult human brain. *bioRxiv*, 2022.2010.2012.511898, doi:10.1101/2022.10.12.511898 (2022).

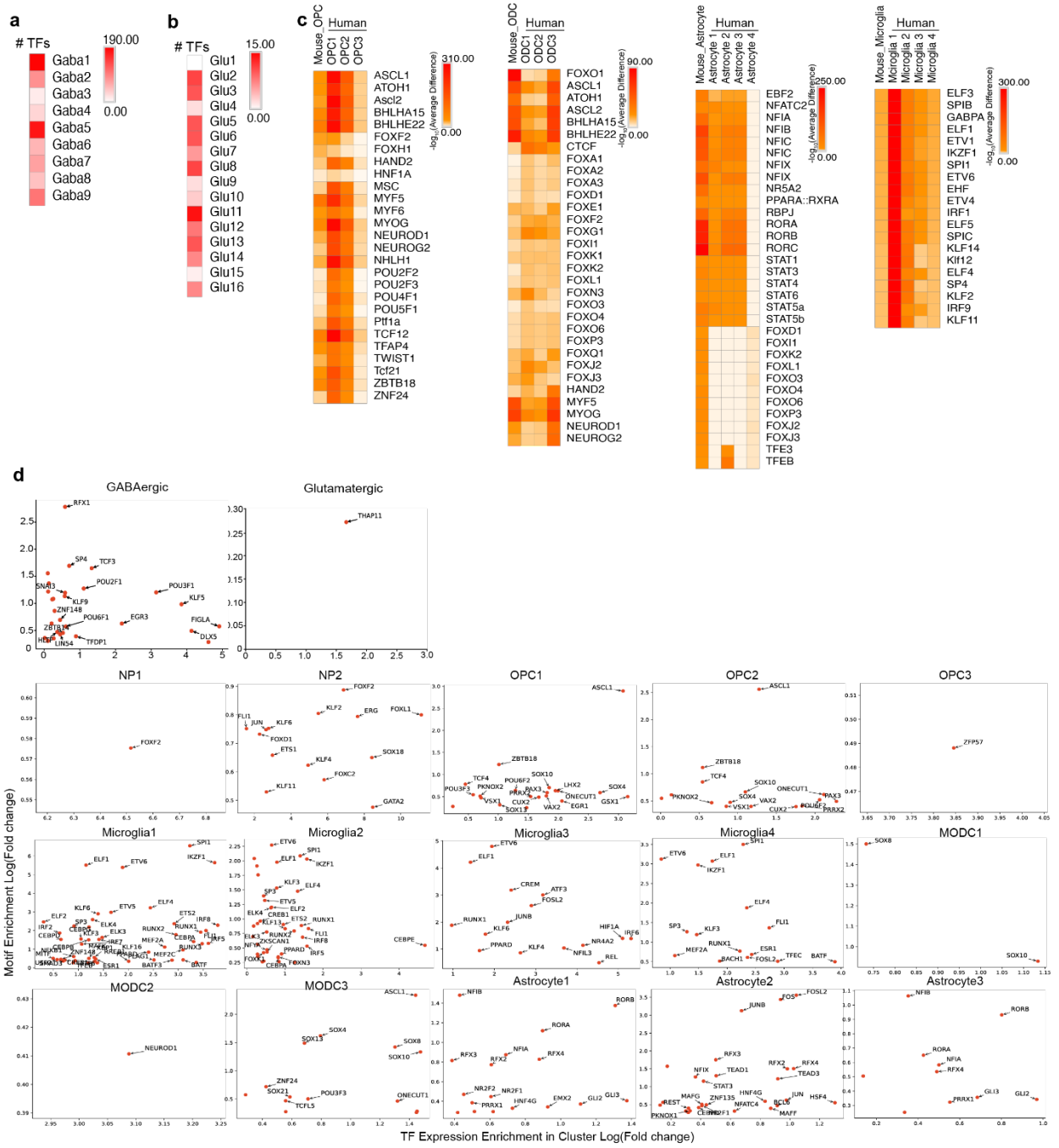
120 Yoshida, S., Sasa, M. & Takaori, S. Different sensitivity to hypoxia in neuronal activities of lateral vestibular and spinal trigeminal nuclei. *Stroke* **19**, 357-364, doi:10.1161/01.str.19.3.357 (1988).
121 Bharadwaj, R. *et al.* Conserved chromosome 2q31 conformations are associated with transcriptional regulation of GAD1 GABA synthesis enzyme and altered in prefrontal cortex of subjects with schizophrenia. *J Neurosci* **33**, 11839-11851, doi:10.1523/JNEUROSCI.1252-13.2013 (2013).
122 Forbes, S., Herzog, H. & Cox, H. M. A role for neuropeptide Y in the gender-specific gastrointestinal, corticosterone and feeding responses to stress. *Br J Pharmacol* **166**, 2307-2316, doi:10.1111/j.1476-5381.2012.01939.x (2012).
123 Nahvi, R. J. & Sabban, E. L. Sex Differences in the Neuropeptide Y System and Implications for Stress Related Disorders. *Biomolecules* **10**, doi:10.3390/biom10091248 (2020).
124 Ramanathan, L. & Siegel, J. M. Gender differences between hypocretin/orexin knockout and wild type mice: age, body weight, body composition, metabolic markers, leptin and insulin resistance. *J Neurochem* **131**, 615-624, doi:10.1111/jnc.12840 (2014).
125 Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e1821, doi:10.1016/j.cell.2019.05.031 (2019).
126 Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nature methods* **16**, 1289-1296, doi:10.1038/s41592-019-0619-0 (2019).
127 Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin state analysis with Signac. *Nature methods* **18**, 1333-1341, doi:10.1038/s41592-021-01282-5 (2021).
128 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).
129 Schep, A. (2022).
130 Machiela, M. J. & Chanock, S. J. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* **31**, 3555-3557, doi:10.1093/bioinformatics/btv402 (2015).
131 Coetzee, S. G., Coetzee, G. A. & Hazelett, D. J. motifbreakR: an R/Bioconductor package for predicting variant effects at transcription factor binding sites. *Bioinformatics* **31**, 3847-3849, doi:10.1093/bioinformatics/btv470 (2015).
132 Chandler, D. B., Jackson, R. M., Briggs, A. D., Fuller, W. C. & Fulmer, J. D. The effect of bleomycin on lung metabolism of prostaglandin E2 in hamster. *Prostaglandins, Leukotrienes and Medicine* **19**, 139-151, doi:10.1016/0262-1746(85)90080-0 (1985).
133 Kim, H. K. *et al.* Deep learning improves prediction of CRISPR-Cpf1 guide RNA activity. *Nat Biotechnol* **36**, 239-241, doi:10.1038/nbt.4061 (2018).


a
TSS enrichment
Nucleosome Signal
Percent Mito Genes
Number of Features: RNA
Number of Features: ATAC
m1
m2
m3
m4
m5
m6
b
Gad1
Lepr
Mc4r
Ddc
c
Average expression
Percent expressed
d
Slc17a6
Slc32a1
Ddc
Hcrt
e
Th
Ghr
Scl17a6
chr7 position (bp)
chr15 position (bp)

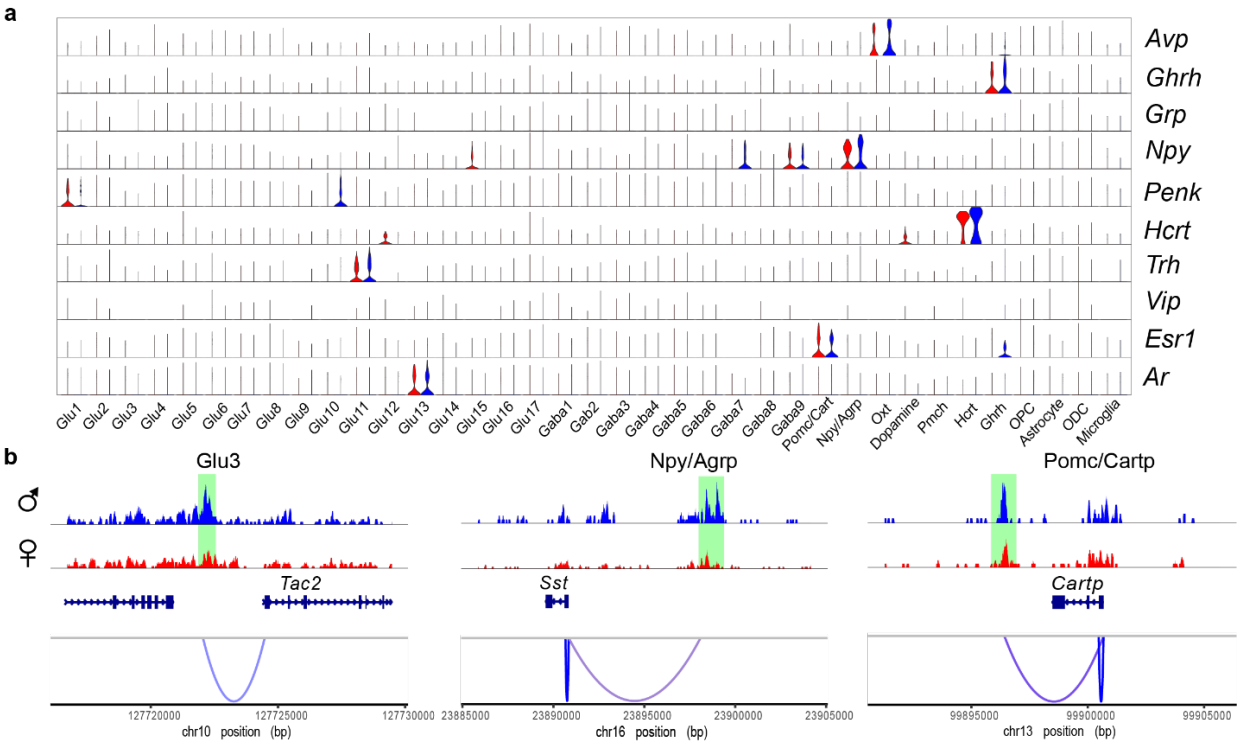
Extended Data Fig.1 Combined scRNA and ATAC profiling of the mouse hypothalamus. a, Violin plot of quality control metrics: TSS enrichment scores, nucleosome signal, percentage of mitochondrial genes, number of features in RNA data, and number of features in ATAC data. **b,** Uniform Manifold Approximation and Projection (UMAP) of scRNA of *Gad1*, *Lepr*, *Mc4r*, and *Ddc*. **c,** Dot plot of neuropeptide gene expression across all cell types in mouse hypothalamus. **d,** Coverage plot of genes, *Slc17a6*, *Slc32a1*, *Ddc*, *Hcrt* showing scATAC peaks highlighted in blue around the gene body. **f,** scATAC peaks to gene linkage plots for *Th* and *Ghr* in specialized neurons, and *Slc17a6* in the GABAergic neurons with the genome tracks. The linked scATAC peaks to gene promoters are highlighted in blue.


a
Hypothalamus
b
TSS enrichment
Nucleosome Signal
Percent Mito Genes
Number of Features: RNA
Number of Features: ATAC
c
Neuron
Astrocyte
OPC
ODC
Microglia
d
Microglia 4
Microglia 3
Microglia 2
Microglia 1
Astrocyte 4
Astrocyte 3
Astrocyte 2
Astrocyte 1
OPC 3
OPC 2
OPC 1
MODC 3
MODC 2
MODC 1
NP 2
NP 1
Glu
GABA
ADCYAP1
AGRP
AGT
AVP
CARPT
CCK
CRH
GAL
GALP
GHRH
GHRL
GNRH1
GRP
HCRT
NPB
NPVF
NPW
NPY
NTS
NUCB2
OXT
PDYN
PENK
PMCH
PNOC
POMC
QRFP
SST
TAC1
TH
TRH
VIP
Average expression
Percent expressed
e
Astrocyte1
Astrocyte2
Astrocyte3
Astrocyte4
-Log10(p)
Regulation of fatty acid oxidation
Negative regulation of fatty acid oxidation
Oligopeptide transport
Positive regulation of fatty acid oxidation
D-amino acid transport
Metanephric nephron morphogenesis
L-aspartate transmembrane transport
Cell-matrix adhesion
Cell-substrate junction assembly
Regulation of systemic arterial blood pressure
Regulation of plasma membrane repair
Cyclic-nucleotide-mediated signaling
Glutathione transport
Regulation of plasma membrane repair
Metanephric nephron tubule morphogenesis
Inner ear receptor cell differentiation
Regulation of mechanoreceptor differentiation
Inner ear auditory receptor cell differentiation
D-aspartate transport
Tripeptide transport
CDP-choline pathway
Cilium movement
Axoneme assembly
Calcium transport
Cilium-dependent cell motility
Axonemal dynein complex assembly
Extracellular fluid movement
Extracellular transport
f
Microglia1
Microglia2
Microglia3
Microglia4
CD163
CD86
5.00
0.00
g
GABA
Glu
NP 1
NP 2
OPC 1
OPC 2
OPC 3
MODC 1
MODC 2
MODC 3
Astrocyte 1
Astrocyte 2
Astrocyte 3
Astrocyte 4
Microglia 1
Microglia 2
Microglia 3
Microglia 4
GAD2
OXT
GFAP
CSPG4

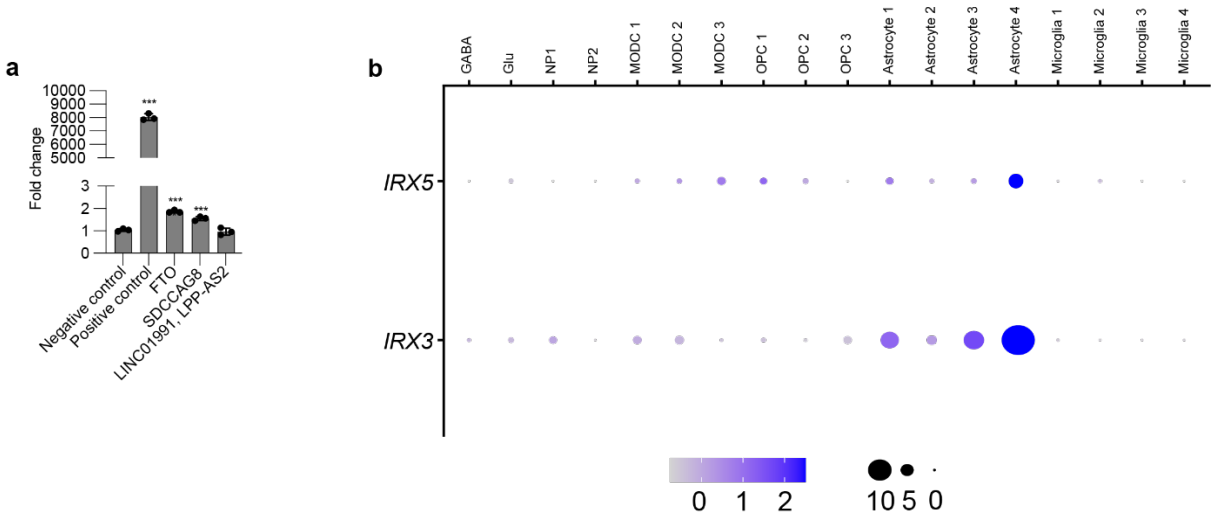
Extended Data Fig.2 Combined scRNA and ATAC profiling of the human hypothalamus. a, Representative image of human hypothalamus section. **b,** Violin plot of quality control metrics: TSS enrichment scores, nucleosome signal, percentage of mitochondrial genes, number of features in RNA data, and number of features in ATAC data. **c,** Heatmap of the top five differentially expressed genes and the top ten scATAC peaks across all cell populations in human hypothalamus prior to subclustering. **d,** Dot plot of neuropeptide gene expression across all cell types in the human hypothalamus. **e,** Gene ontology analyses of all astrocyte cell clusters. **f,** Heatmap of *CD163* and *CD86* expression in astrocyte populations. **g,** Coverage plot of genes, *GAD2*, *OXT*, *GFAP*, *CSPG4* showing scATAC peaks highlighted in blue around the gene body.


a
TFs
190.00
0.00
Gaba1
Gaba2
Gaba3
Gaba4
Gaba5
Gaba6
Gaba7
Gaba8
Gaba9
b
TFs
15.00
0.00
Glu1
Glu2
Glu3
Glu4
Glu5
Glu6
Glu7
Glu8
Glu9
Glu10
Glu11
Glu12
Glu13
Glu14
Glu15
Glu16
c
Mouse_OPC
Human
OPC1
OPC2
OPC3
310.00
0.00
ASCL1
ATOH1
Ascl2
BHLHA15
BHLHE22
FOXF2
FOXH1
HAND2
HNF1A
MSC
MYF5
MYF6
MYOG
NEUROD1
NEUROG2
NHLH1
POU2F2
POU2F3
POU4F1
POU5F1
Ptf1a
TCF12
TFAP4
TWIST1
Tcf21
ZBTB18
ZNF24
Mouse_ODC
ODC1
ODC2
ODC3
90.00
0.00
FOXO1
ASCL1
ATOH1
ASCL2
BHLHA15
BHLHE22
CTCF
FOXA1
FOXA2
FOXA3
FOXD1
FOXE1
FOXF2
FOXG1
FOXI1
FOXK1
FOXK2
FOXL1
FOXN3
FOXO3
FOXO4
FOXO6
FOXP3
FOXQ1
FOXJ2
FOXJ3
HAND2
MYF5
MYOG
NEUROD1
NEUROG2
Mouse_Astrocyte
Astrocyte 1
Astrocyte 2
Astrocyte 3
Astrocyte 4
250.00
0.00
EBF2
NFATC2
NFIA
NFIB
NFIC
NFIX
NR5A2
PPARA::RXRA
RBPJ
RORA
RORB
RORC
STAT1
STAT3
STAT4
STAT6
STAT5a
STAT5b
FOXD1
FOXI1
FOXK2
FOXL1
FOXO3
FOXO4
FOXO6
FOXP3
FOXJ2
FOXJ3
TFE3
TFEB
Mouse_Microglia
Mcirglia 1
Microglia 2
Microglia 3
Microglia 4
300.00
0.00
ELF3
SPIB
GABPA
ELF1
ETV1
IKZF1
SPI1
ETV6
EHF
ETV4
IRF1
ELF5
SPIC
KLF14
Klf12
ELF4
SP4
KLF2
IRF9
KLF11
d
GABAergic
Glutamatergic
NP1
NP2
OPC1
OPC2
OPC3
Microglia1
Microglia2
Microglia3
Microglia4
MODC1
MODC2
MODC3
Astrocyte1
Astrocyte2
Astrocyte3
Motif Enrichment Log(Fold change)
TF Expression Enrichment in Cluster Log(Fold change)
e
CD163
Expression Level
GABA
Glu
NP1
NP2
MODC 1
MODC 2
MODC 3
OPC 1
OPC 2
OPC 3
Astrocyte 1
Astrocyte 2
Astrocyte 3
Astrocyte 4
Microglia 1
Microglia 2
Microglia 3
Microglia 4
f
-Log10(P value)
Circulatory system development
Actin cytoskeleton organization
Actomyosin structure organization
Cytoskeleton organization
Actin filament-based process
Cell projection organization
Plasma membrane organization
Blood vessel development
g
-Log10(P value)
Regulation of cell-substrate adhesion
Extracellular matrix organization
Extracellular structure organization
Renal system development
External encapsulating structure organization
Kidney development
Regulation of cell-matrix adhesion
h
-Log10(P value)
Regulation of neuron projection development
Regulation of axonogenesis
Positive regulation of axon extension
Positive regulation of axonogenesis
Positive regulation of cell development
Regulation of actin filament organization
i
-Log10(P value)
Heparan sulfate proteoglycan biosynthetic process
G protein-coupled glutamate receptor signaling pathway
ERBB2 signaling pathway
Cellular response to forskolin
Response to forskolin
Nerve growth factor signaling pathway
Neurotransmitter receptor localization to membrane

Extended Data Fig.3 TFs and TFBS enrichment in mice and human hypothalamus cell populations. **a,** Heatmap showing the number of TFs in mouse GABAergic neurons shared with human GABAergic neurons. **b,** Heatmap showing the number of TFs in mouse glutamatergic neurons shared with human GABAergic neurons. **c,** Heatmap of TFBS in OPC, ODC, astrocyte, and microglia populations shared by mice and human. **d,** TFs that show both upregulation of gene expression and motif enrichment in all populations. **e,** Violin plot of *CD163* expression**. f-i,** Gene ontology of genes regulated by mouse-specific TFs in GABAergic (**f**), glutamatergic (**g**) neurons, and human- specific TFs in GABAergic (**h**) and glutamatergic (**i**) neurons.


a
Avp
Ghrh
Grp
Npy
Penk
Hcrt
Trh
Vip
Esr1
Ar
Glu1
Glu2
Glu3
Glu4
Glu5
Glu6
Glu7
Glu8
Glu9
Glu10
Glu11
Glu12
Glu13
Glu14
Glu15
Glu16
Glu17
Gaba1
Gab2
Gaba3
Gaba4
Gaba5
Gaba6
Gaba7
Gaba8
Gaba9
Pomc/Cart
Npy/Agrp
Oxt
Dopamine
Pmch
Hcrt
Ghrh
OPC
Astrocyte
ODC
Microglia
b
Glu3
Npy/Agrp
Pomc/Cartp
♂
♀
Tac2
Sst
Cartp
127720000
127725000
127730000
chr10 position (bp)
23885000
23890000
23895000
23900000
23905000
chr16 position (bp)
99895000
99900000
99905000
chr13 position (bp)

Extended Data Fig.4 Sex-differential expression and regulation. a, Violin plots of sex-differentially expressed neuropeptide genes in mouse hypothalamus. **b,** Linkage plots showing sex-specific scATAC peaks linked to sex-differentially expressed genes, including *Tac2* in Glu3 cluster, *Sst* in Npy/Agrp neurons, and *Cartpt* in Pomc/Cartp neurons.


a
Fold change
10000
9000
8000
7000
6000
5000
3
2
1
0
***
***
***
Negative control
Positive control
FTO
SDCCAG8
LINC01991, LPP-AS2
b
GABA
Glu
NP1
NP2
MODC 1
MODC 2
MODC 3
OPC 1
OPC 2
OPC 3
Astrocyte 1
Astrocyte 2
Astrocyte 3
Astrocyte 4
Microglia 1
Microglia 2
Microglia 3
Microglia 4
IRX5
IRX3
0 1 2
10 5 0

Extended Data Fig.5 a, Luciferase assays in human astrocytes for astrocyte-specific scATAC-seq peaks overlapping obesity-associated SNPs. **b,** Dot plots of *IRX3* and *IRX5* expression across cell populations in the human hypothalamus.

Supplementary Table 1. Human hypothalamus sample information.

ID	Age/Gender/PMI	Regions
Hypo 1	82/male/7hr	Hypothalamus (between optic chiasm and mammillary body)
Hypo 2	62/male/7hr	Hypothalamus (between optic chiasm and mammillary body)
Hypo 3	62/male/10hr	Hypothalamus (between optic chiasm and mammillary body)
Hypo 4	43/female/12hr	Hypothalamus (between optic chiasm and mammillary body)
Hypo 5	45/female/4hr	Hypothalamus (between optic chiasm and mammillary body)
Hypo 6	73/female/8hr	Hypothalamus (between optic chiasm and mammillary body)

Supplementary Table 2. Sequences of primers for RT-qPCR, cloning, genotyping and gRNAs.

Gene/Use/sgRNA	Primers
GAPDH	GTC TCC TCT GAC TTC AAC AGC G
	ACC ACC CTG TTG CTG TAG CCA A
MC4R	CCT GAG GTG TTT GTG ACT CTG G
	AAC GCT CAC CAG CAT ATC AGC C
FTO	CCA GAA CCT GAG GAG AGA ATG G
	CGA TGT CTG TGA GGT CAA ACG G
IRX3	CTC TCC CTG CTG GGC TCT
	CAA GGC ACT ACA GCG ATC TG
IRX5	GCA AAG ACT CTC CCT ATG A
	CTC TCT CGG TGA TGG AAA
1_THD	CTT TGT CCA GAG ACC ACA CC
	TGA TCC AAG GAA AAA TCA GGC G
2_THD	TGG GGA CCC GTT GAA GAA CA
	AGA AAG ACG CAG AGT GGA AGT T
GPR5B	CGA GAT AAC ACC TCT CGG TCC
	CAC AGT TTA AGC TGC TAG AGA GA
ADCY3	GCA GGC ACA CTT AGT CCT GT
	CCT CTG CGG AAA GCT CAG G
SULT1A1	AAG TGC GAT TTC TGG GTC TAA A
	GAG GCT CAA GGG TTC AAG ACT
SH2B1	GGG CCA CAG TAA ACT CCT CC
	TGC ATG CTC CAC CCT GTT AG
MLN	GAT GCA GAC GCT CAT GGG TA
	GAT GCC CTC CTC ACT CTC TG
THP12P9	CCT GTA CGA GAA CAC GCA GT
	GGA AGG CGG AGT AGA GCT AAG
GIPR	TCA CAG ATA CCA TTC ACT GGG G
	GAA GGC GGA GTA GAG CTA AGA GA
BCDIN3D	GTT TGC TTG AGG CAG CTG TG
	CTC ACT CAC TTC TCC ACC CG
FTO	GAA GGC AAT CTT TCT GCT TCC TG
	TTA GGC AAG ACA CAA ATC ACA AG
GAS8	TGT AGT TCC AGA GCG GCA G
	ATC GCT GAA TCT GGG CGA TAG
VPS45	GCT CAG TTG GAA GGA GCA CT
	GCA CGA CTA AAG GTG GGG AA
MC4R EKO gRNA	CGC TGG CAC ATT ATT GTT A
	ACA GTT TGT TGA CAG CCA GA
FTO enhancer g1	CAG AGA TGC CGA AGA CTC ATT
FTO enhancer g2	AGA CTC ATT TCG GGT ACA ATT
FTO enhancer g3	GGC TGA CAA GAC CAG AAC AAC
FTO enhancer g4	AGC ACA GTG AAC GAG CAA GGA

Supplementary Table 3. Predifferntiation media and maturation media for neuron differentiation from WTc11 iPSC cells.

Predifferentiation media				
Reagent	**Stock**	**Final conc.**	**Example (50mL)**	**Example (10mL)**
KnockOUT DMEM/F-12	--	100%	49mL	10mL Pre-differentiation media
N-2 Supplement (100X)	100X	1X	0.5mL	
NEAA (100X)	100X	1X	0.5mL	
BDNF	25 ug/mL	10 ng/mL	***Add these right before use***	4uL
NT-3	25 ug/mL	10 ng/mL		4uL
Laminin	*1.2 mg/mL*	*1.2 ug/mL*		10uL
Doxycycline	1 mg/mL	2 ug/mL		20uL
Maturation media				
Reagent	**Stock**	**Final conc.**	**Example (50mL)**	**Example (10mL)**
Neuralbasal-A Medium	--	50%	24.25mL	10 mL Maturation Medium
DMEM/F-12, HEPES	--	50%	24.25mL	
NEAA(100x)	100X	1X	0.5mL	
B-27 Supplement (50x)	50X	0.5X	0.5mL	
N-2 Supplement (100x)	100X	0.5X	0.25mL	
GlutaMAX Supplement (100x)	100X	0.5X	0.25mL	
BDNF	25 ug/mL	10 ng/mL	***Add these right before use***	4uL
NT-3	25 ug/mL	10 ng/mL		4uL
Laminin	*1.2 mg/mL*	*1.2 ug/mL*		10uL