A molecular census of arcuate hypothalamus and median eminence cell types

John N Campbell[1], Evan Z Macosko[2–4], Henning Fenselau[1], Tune H Pers[5,6], Anna Lyubetskaya[1], Danielle Tenen[1], Melissa Goldman[2,3], Anne M J Verstegen[1], Jon M Resch[1], Steven A McCarroll[2–4,7,8], Evan D Rosen[1,8], Bradford B Lowell[1,7] & Linus T Tsai[1]

The hypothalamic arcuate–median eminence complex (Arc-ME) controls energy balance, fertility and growth through molecularly distinct cell types, many of which remain unknown. To catalog cell types in an unbiased way, we profiled gene expression in 20,921 individual cells in and around the adult mouse Arc-ME using Drop-seq. We identify 50 transcriptionally distinct Arc-ME cell populations, including a rare tanycyte population at the Arc-ME diffusion barrier, a new leptin-sensing neuron population, multiple agouti-related peptide (AgRP) and pro-opiomelanocortin (POMC) subtypes, and an orexigenic somatostatin neuron population. We extended Drop-seq to detect dynamic expression changes across relevant physiological perturbations, revealing cell type–specific responses to energy status, including distinct responses in AgRP and POMC neuron subtypes. Finally, integrating our data with human genome-wide association study data implicates two previously unknown neuron populations in the genetic control of obesity. This resource will accelerate biological discovery by providing insights into molecular and cell type diversity from which function can be inferred.

The Arc is an evolutionarily conserved brain region with diverse roles in mammalian physiology, including energy homeostasis, reproduction and neuroendocrine control of growth hormone and prolactin release. Aside from its functional diversity, Arc is known for its unique anatomical relationship with the blood-brain barrier (BBB)[1], which protects the cell bodies and dendrites of arcuate neuroendocrine neurons while allowing their axons to enter BBB-free areas of the adjoining ME; these axons release signals into fenestrated capillaries that carry blood to the pituitary. Blood-borne signals can also diffuse from ME to Arc, giving Arc privileged access to peripheral hormones, nutrients and other metabolic signals[1]. This access is dynamically regulated by tanycytes[2]. The tanycyte is a specialized type of ependymal cell lining the third ventricle that extends processes throughout the Arc and ME. Together the Arc and ME form an anatomically unique and functionally important complex, the Arc-ME.

The varied functions of Arc-ME are supported by molecularly specialized neuron subtypes. For instance, orexigenic AgRP neurons and anorexigenic POMC neurons control feeding[3,4], while kisspeptin (KISS1) neurons regulate serum luteinizing hormone levels[5]. Despite decades of research on Arc-ME, a complete census of its cell types is not available, with several functional Arc neuron populations having no known markers[6,7]. Studies relying on immunohistochemistry or *in situ* hybridization have been limited by the number of proteins or transcripts that can be simultaneously visualized in individual cells and by a strong bias toward known markers. With recent advances in transcriptomic technology, however, thousands of cells can be profiled individually, enabling discovery of cell types with reduced bias (for example, refs. 8–10).

Using Drop-seq we systematically catalogued cell types from in and around mouse Arc-ME, identifying 34 distinct neuronal populations (24 from Arc-ME) and 36 non-neuronal populations (26 from Arc-ME) from 20,921 individual cell profiles. We determined specific markers that can be used both to identify cell types and to infer their function. Among our findings, we uncover several distinct subsets of AgRP and POMC neurons, a new group of leptin-responsive neurons and an undescribed orexigenic role for Arc somatostatin neurons. We further used Drop-seq to assess cell type–specific responses to fasting and high-fat diet, revealing energy status–sensitive populations and reinforcing the functional heterogeneity of AgRP and POMC subtypes. Lastly, we show how the increased detail provided by such profiles improves the ability to connect genome-wide association study (GWAS) genes to relevant cell types. Together our results demonstrate how such a molecular census can be used to transform our understanding of a complex tissue and the biological processes it regulates.

[1]Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, USA. [2]Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA. [3]Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA. [4]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA. [5]The Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark. [6]Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark. [7]Program in Neuroscience, Harvard Medical School, Boston, Massachusetts, USA. [8]Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA. Correspondence should be addressed to B.B.L. (blowell@bidmc.harvard.edu), E.D.R. (erosen@bidmc.harvard.edu) or L.T.T. (ltsai@bidmc.harvard.edu).

Received 19 July 2016; accepted 4 January 2017; published online 6 February 2017; doi:10.1038/nn.4495

RESULTS

Unbiased transcriptomics identifies 50 distinct Arc-ME cell types

Using Drop-seq[8] we profiled 20,921 transcriptomes from acutely dissociated Arc-ME cells of adult mice under various feeding conditions: *ad libitum* access to standard mouse chow, low-fat diet or high-fat diet, as well as overnight fasting, with or without subsequent refeeding (**Fig. 1a** and **Supplementary Fig. 1a**). After correcting for batch effects, we performed principal component (PC) analysis, dimensionality reduction with spectral *t*-distributed stochastic neighbor embedding (tSNE), and density-based clustering (**Fig. 1a** and Online Methods). Our initial analysis identified 20 distinct clusters (**Fig. 1b**,**c**). Each cluster contained cells from each feeding condition and sample batch, indicating the transcriptional identities of these cell clusters are stable across those experimental conditions (**Supplementary Fig. 1a** and **Supplementary Table 1**). Using expression patterns of cell type–specific marker genes, we assigned a single identity to each cluster: neurons (*Tubb3*$^+$), ependymocytes (*Ccdc153*$^+$), tanycytes (*Rax*$^+$), oligodendrocyte lineage cells (*Mag*$^+$), oligodendrocyte precursor cells (also known as NG2 cells), *Cspg4*$^+$), macrophages (*Aif1*$^+$), endothelial cells (*Slco1c1*$^+$), mural cells (*Mustn1*$^+$) and astrocytes (*Gfap*$^+$; **Fig. 1d** and **Supplementary Table 2**).

We also identified clusters of non-neural cells from tissues adjoining the Arc-ME—namely, vascular and leptomeningeal cells (VLMCs; *Lum*$^+$)[9] and pituitary cells from pars tuberalis (*Tshb*$^+$). Exemplifying our detection of new markers, we found that pars tuberalis cells the gene encoding cholecystokinin (*Cck*) at a high level (**Fig. 1d** and **Supplementary Fig. 1b**,**c**). Of note, CCK has been shown to control pituitary release of prolactin[11], a well-known function of the pars tuberalis[12].

In many cases, a previously described cell type was represented by more than one cluster. We detected oligodendrocyte markers (for example, *Olig1*, *Cd9*, *Pllp*) in four clusters, each corresponding to a distinct stage of oligodendrocyte differentiation[9] (**Supplementary Fig. 1d**). Neurons and tanycytes also formed multiple clusters, consistent with their known heterogeneity (see below). We subclustered each original non-neuronal cluster (**Supplementary Fig. 1e** and **Supplementary Table 3**) to yield a total of 36 subclusters, revealing additional heterogeneity: for example, the mural cell cluster (**Fig. 1b**) comprises pericytes and vascular smooth muscle cells (**Supplementary Fig. 1f**) and a macrophage cluster (**Fig. 1b**) consists of transcriptionally related microglia and perivascular macrophages (**Supplementary Fig. 1g**).

Our subclustering revealed eight clusters of ependymal cells (*Vim*$^+$ *Sox2*$^+$; **Supplementary Fig. 2a**,**b**). Ependymal cells line the ventricular surface of the hypothalamus and are broadly categorized as either ependymocytes, multiciliated cells that secrete and move cerebrospinal fluid, or tanycytes, monociliated cells with basal processes extending throughout the mediobasal hypothalamus and median eminence[1]. Tanycytes have a variety of functions, from glucose sensing and neurogenesis to controlling the chemical exchanges between brain parenchyma, cerebrospinal fluid and bloodstream[1,13]. Studies in rodents have described four tanycyte subtypes (α1, α2, β1 and β2) occupying distinct regions along the third ventricle[13] (**Fig. 2a**). While markers for these subtypes have been suggested[13], it is unclear whether those markers truly distinguish between molecularly distinct cellular subtypes.

To trace the anatomical origin of our ependymal cell clusters, we cross-referenced the markers of each with *in situ* hybridization data from the Allen Mouse Brain Atlas (http://mouse.brain-map.org/)[14]. We found many to be expressed in well-defined regions along the third ventricle (**Fig. 2b**), allowing us to assign each cluster to the ependymal cell subtype occupying those regions (**Fig. 2c**). Our results confirm and extend functional categorization of ependymal cell subtypes by anatomical position: we validated several ependymal subtype markers (**Supplementary Fig. 2b**) and discovered new markers for each subtype (**Fig. 2d** and **Supplementary Fig. 2c**,**d**). Our data nearly double the number of ependymal cell subtypes thought to exist (**Supplementary Fig. 2c**,**d**) and provide insight into each's possible functions.

While many genes were expressed in gradients along the third ventricle, some showed very restricted patterns of expression. For instance, *Sprr1a* was found only at the border between Arc and ME (**Fig. 2e**), where tanycytes are thought to form a diffusion barrier[1,2]. Using an intravenous injection of Evan's blue to mark this diffusion barrier, we found that SPRR1A immunoreactive tanycytes were located precisely at the Arc-ME barrier (**Fig. 2e**). Small proline-rich (SPRR) proteins, including SPRR1A, are crucial constituents of the cornified envelope, the diffusion barrier in the skin[15]. By identifying *Sprr1a* as a specific marker for these tanycytes, our results provide a genetic means by which to develop tools for investigating the role of SPRR1A and *Sprr1a*$^+$ tanycytes in the Arc-ME diffusion barrier.

Neuron-specific clustering reveals new types and subtypes of known Arc-ME populations

At least six distinct types of neurons have been identified in the Arc-ME. Among these are two neuroendocrine populations, growth hormone releasing hormone (GHRH) and tuberoinfundibular dopaminergic (TIDA) neurons, that control pituitary release of growth hormone and prolactin, respectively. In addition, there are at least four types of centrally projecting neurons, including AgRP neurons and POMC neurons, which play vital opposing roles in energy balance, and kisspeptin/neurokinin B/dynorphin ('KNDy') neurons, which regulate fertility. Somatostatin (SST) neurons have also been described as distinct[16], but their function is not well understood. Other functional types of neurons exist in the Arc-ME as well, including a thermogenic population known for its expression of the rat insulin II promoter (RIP)-Cre transgene as well as a leptin-sensing GABAergic population distinct from AgRP neurons that is critical for weight regulation[6,7,17,18].

To further characterize neuronal diversity in the Arc-ME, we performed additional rounds of clustering on the 13,079 neurons in our data set, identifying 34 clusters (**Fig. 3a**,**b** and **Supplementary Fig. 3a**,**b**) which we then annotated with a neuron cluster number (e.g., "n01") and candidate markers. Many of these clusters are defined by unique candidate markers (**Fig. 3c** and **Supplementary Table 4**), though some lacked a unique marker and are annotated instead with a combination of markers. Two clusters, containing a combined 14% of the neurons in our data set, appeared to be heterogenous, lacking any strong and specific markers, and so are annotated as "unassigned." The reason(s) for lack of assignment are not clear but could include technical issues with sample quality or purity (suggested by the lower unique molecular identifier (UMI) and gene counts in this population), or could reflect true biological ambiguity. However, owing to a lack of specific markers or specific combination of markers for these neurons, we could not investigate them further.

Notably, we detected markers of all six previously described types of Arc-ME neurons (*Agrp*, *Pomc*, *Sst*, *Kiss1*/*Tac2*, *Ghrh* and *Th*). In total, 14 clusters expressed those markers (2 *Agrp*$^+$, 3 *Pomc*$^+$, 5 *Sst*$^+$ and 6 *Th*$^+$ clusters), revealing molecular heterogeneity even within well-studied arcuate populations. We observe clusters expressing both *Th* and *Sst* and both *Agrp* and *Sst*; these represent overlaps between Arc-ME neuron populations previously thought to be distinct.

Also, while six of the neuron clusters are potentially dopaminergic (*Th*, little to no *Dbh*), only two of these expressed prolactin receptor (*Prlr*) and dopamine transporter (*Slc6a3*), representing true TIDA populations (**Fig. 3c**; clusters n03.Th/Sst and n08.Th/Slc6a3)[19,20].

To validate our clustering, we checked expression of genes previously shown to be enriched in AgRP neurons and POMC neurons, two well-studied Arc-ME neuron populations. While we detected leptin receptor transcript (*Lepr*) in both AgRP neurons and POMC neurons,

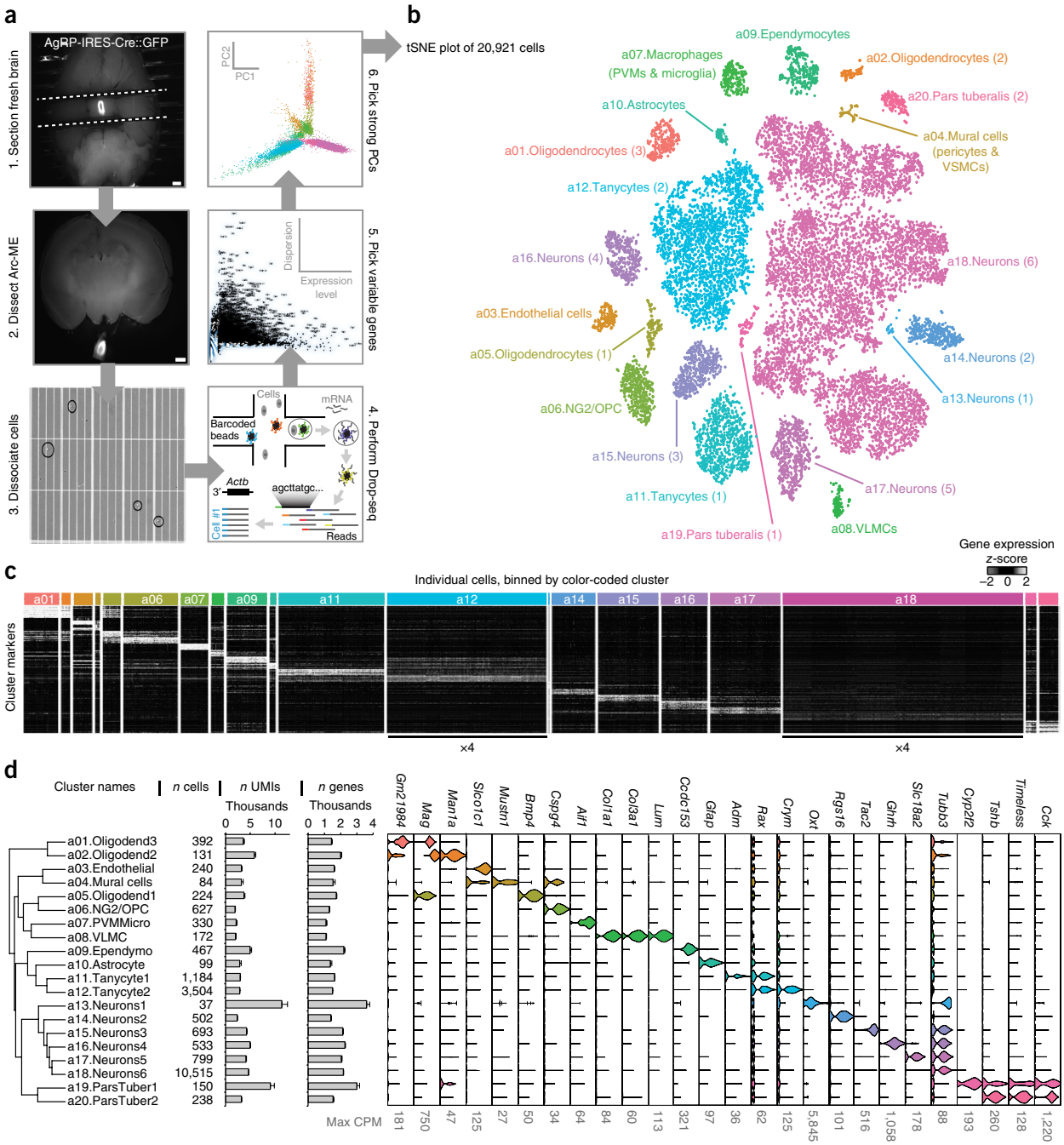

Figure 1 Overview of all cell types. (**a**) Schematic of Arc-ME single-cell transcriptomics. (**b**) Spectral tSNE plot of 20,921 cells, colored per density clustering and annotated according to known cell types. (**c**) Heat map of top marker genes for each cluster. The two largest clusters, a12 and a18, were reduced to one-quarter size to better visualize the smaller clusters (**d**) Dendrogram showing relatedness of cell clusters, followed by (from left to right) cluster identification numbers, cells per cluster, mean ± s.e.m. UMIs per cluster, mean ± s.e.m. genes detected per cluster and violin plots showing expression of cell type marker genes. Oligodend, oligodendrocyte; NG2/OPC, oligodendrocyte precursor cell; ependymo, ependymocyte; PVMMicro, peripheral vascular macrophage and microglia; VLMC, vascular and leptomeningeal cell; ParsTuber, pars tuberalis.

Npy, *Ghsr* and *Acvr1c* were highly enriched in AgRP neurons, as were *Cartpt*, *Htr2c* and *Calb1* in POMC neurons (**Fig. 4a**). Markers from our unbiased clustering therefore concur with previous studies of these neuron populations (for example, ref. 21).

POMC neurons have been shown to have functional heterogeneity[22], but it is unclear whether these subtypes are transcriptionally distinct. Our analysis found three distinct subtypes of POMC neurons, each defined by a set of enriched transcripts (**Supplementary Fig. 4**).

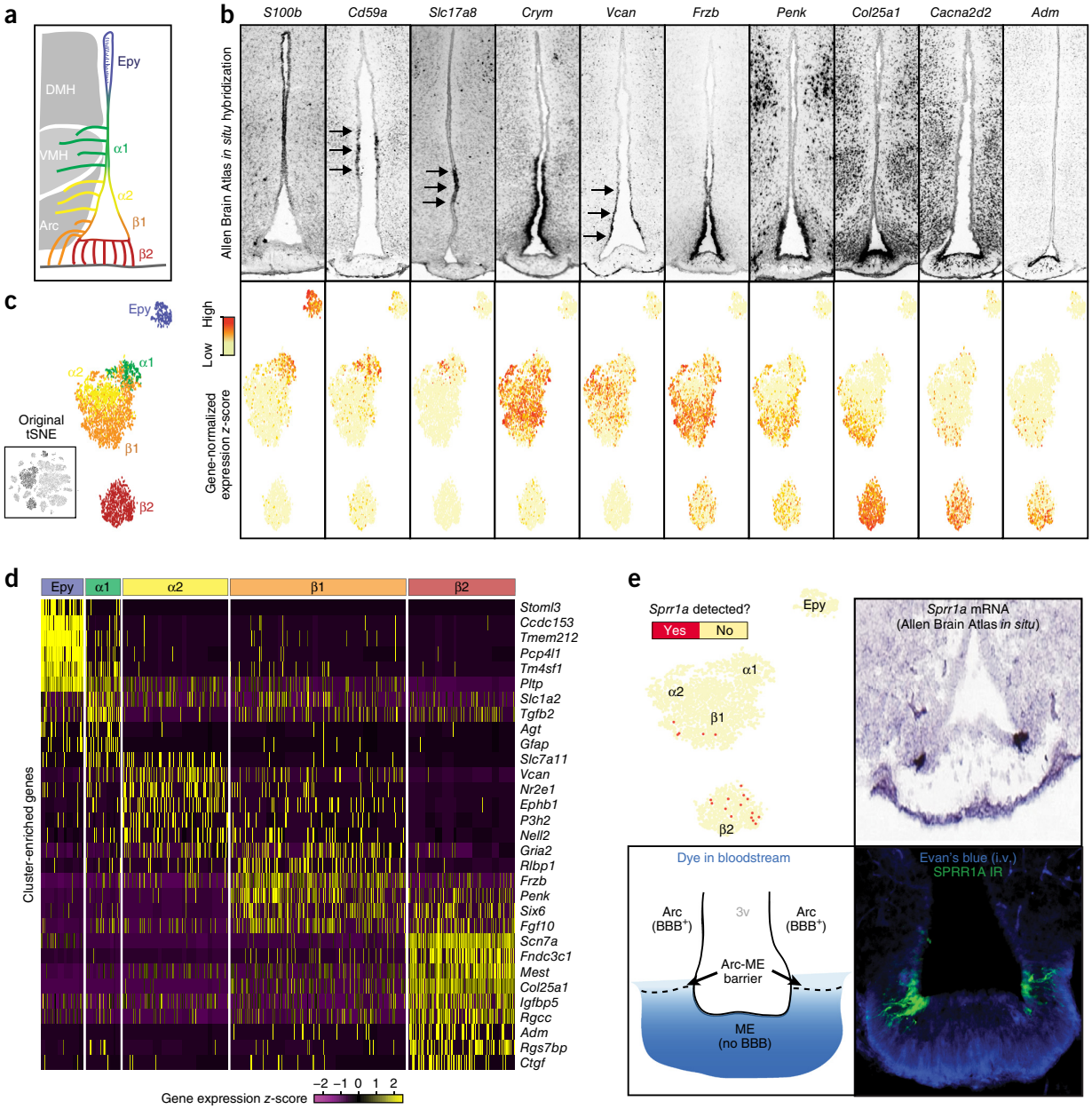

Figure 2 Ependymal cell types. (**a**) Illustration of known subtypes of hypothalamic ependymal (Epy) cells, their approximate anatomical locations and the orientations of their processes. Ependymocytes have cilia in the ventricle and tanycytes have basal processes in the brain parenchyma and median eminence. DMH, dorsomedial hypothalamus; VMH, ventromedial hypothalamus. (**b**) Marker gene expression shown by *in situ* hybridization of coronal brain sections (Allen Mouse Brain Atlas; top) and ependymal cell feature plot (bottom) derived from tSNE plot (**Fig. 1b**). Genes were selected from those differentially expressed among ependymal cell clusters. (**c**) Annotation of ependymal cell clusters on the basis of anatomical localization of marker genes. Figure was derived from tSNE plot (thumbnail; **Fig. 1b**). (**d**) Heat map of single-cell expression of cluster-enriched transcripts. (**e**) Top left, ependymal cell feature plot recolored to indicate cells with any amount of *Sprr1a* transcript. Top right, *Sprr1a in situ* hybridization in a coronal brain section (Allen Mouse Brain Atlas). Bottom left, schematic of an experiment to define the diffusion barrier between Arc and ME; 3v, third ventricle. Bottom right, confocal micrograph comparing SPRR1A immunoreactivity (IR) to the location of the Arc-ME diffusion barrier, visualized by extravasation of intravascular (i.v.) Evan's blue; micrograph is representative of 2 mice.

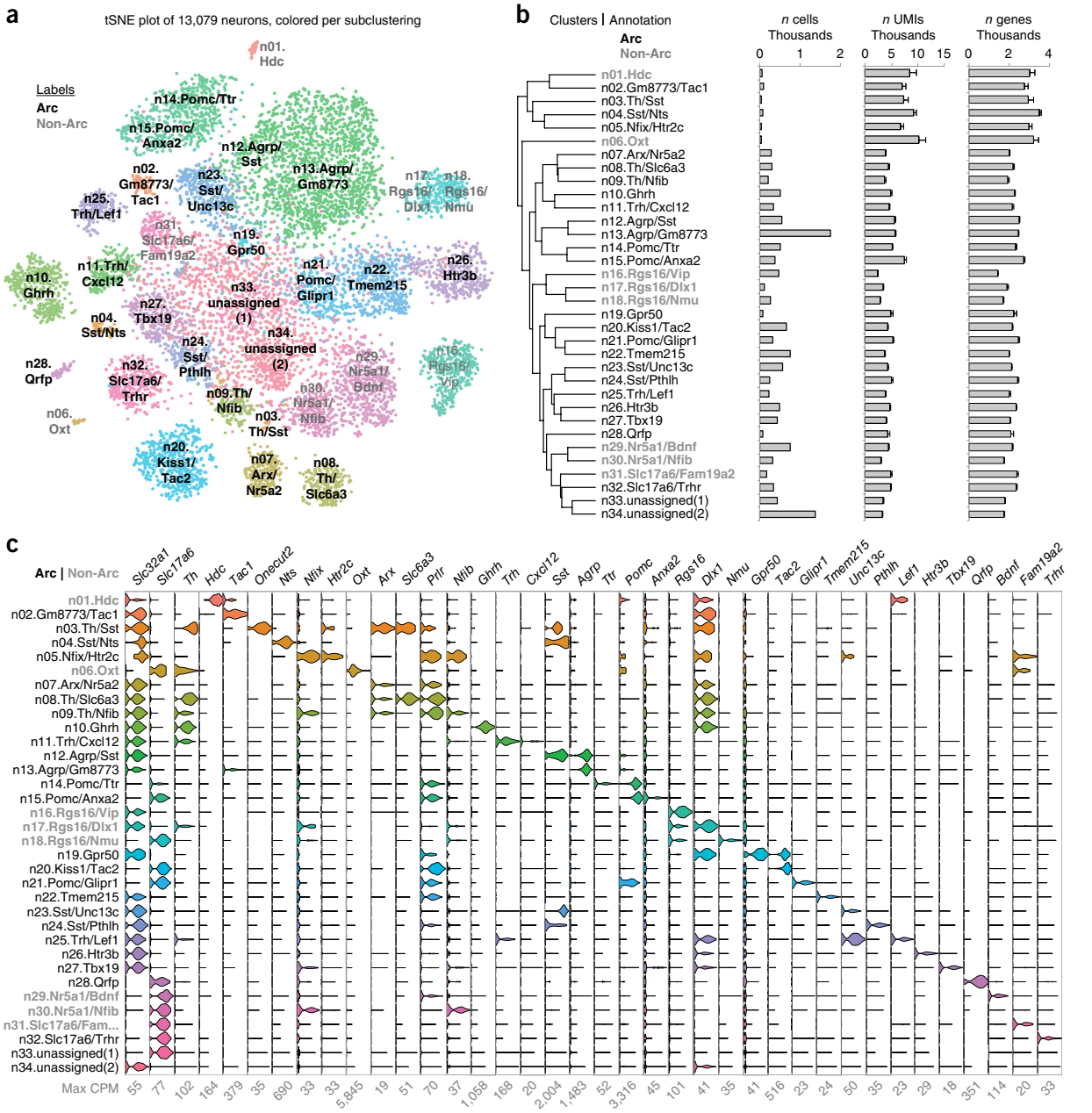

Figure 3 Neuronal cell types. (**a**) Spectral tSNE plot of 13,079 neurons, colored according to the results of iterative subclustering (**Supplementary Fig. 3b**) and labeled according to expression of either a specific marker gene or a specific combination of marker genes. Clusters with gray labels most likely originated from regions surrounding the Arc-ME (see **Supplementary Fig. 5a–d**). (**b**) Dendrogram showing relatedness of neuron clusters, followed by (from left to right) cluster identification numbers, histograms of neurons per cluster, mean ± s.e.m. UMIs per cluster and mean ± s.e.m. unique genes detected per cluster. (**c**) Violin plots of known and novel markers of neuron types in and around the Arc-ME, with maximum counts per million (max CPM) below.

Some of these transcripts encoded secreted peptides, receptors and transcriptional regulators that may underlie their transcriptional identity (**Fig. 4b**). Notably, among POMC subtypes, *Lepr* is largely restricted to the n15.Pomc/Anxa2 subtype, whereas *Htr2c* is predominantly expressed by the other POMC subtypes (**Fig. 4b**). Thus, these transcriptionally distinct subtypes of POMC neurons likely correspond to the leptin-sensing and serotonin-sensing (HTR2C$^+$) subtypes defined by previous studies[22].

Roughly half the neuron clusters did not express markers of known Arc-ME neurons. These unknown clusters could represent neurons from outside Arc-ME, inadvertently included during dissection (**Supplementary Fig. 5a**,**b**). Indeed, we found transcripts enriched in eight neuron clusters, or 17% of all neurons, that the Allen Mouse Brain Atlas shows are expressed in regions neighboring Arc-ME, suggesting that these eight neuron clusters originated from neighboring regions: ventromedial hypothalamus ('SF-1' *Nr5a1*$^+$ *Fezf1*$^+$;

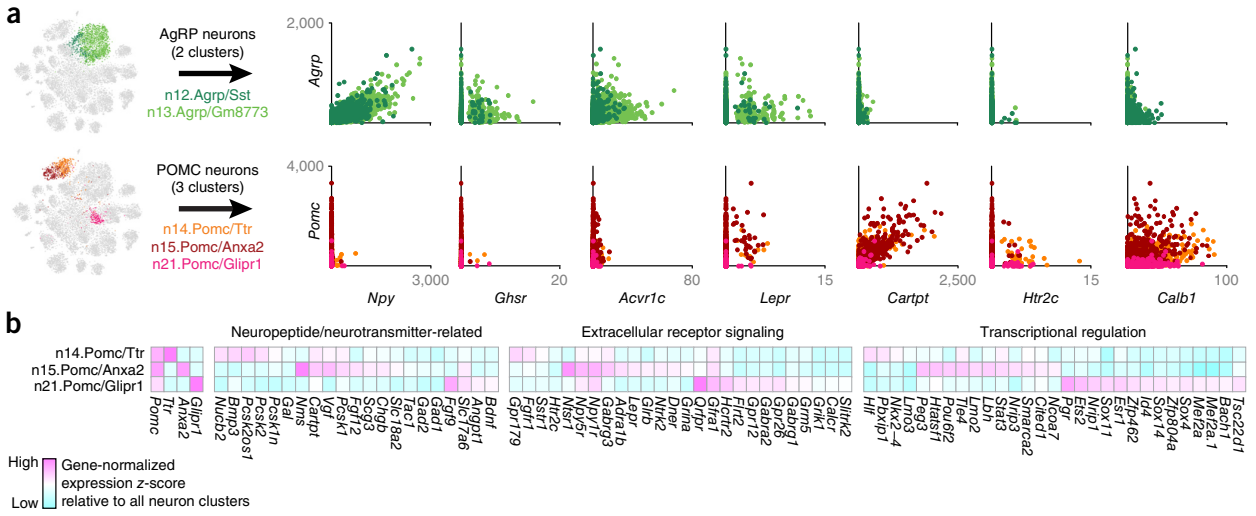

Figure 4 AgRP neurons and POMC neurons. (**a**) Left, selection of AgRP neurons and POMC neurons for analysis of gene coexpression. Right, coexpression of *Agrp* or *Pomc* with genes known to be enriched in AgRP neurons and/or POMC neurons. Values are in counts per million (CPM). (**b**) Differentially expressed genes related to neuropeptide or neurotransmitter signaling and transcriptional regulation in three subtypes of POMC neurons.

Supplementary Fig. 5c; three clusters), suprachiasmatic nucleus (*Rgs16*$^+$, **Supplementary Fig. 5c**; three clusters), tuberomammillary nucleus (*Hdc*$^+$, **Fig. 3c**; one cluster) and retrochiasmatic area (*Oxt*$^+$, **Fig. 3c**; one cluster). Conversely, another 24 neuron clusters (70% of neurons) expressed transcripts enriched in Arc-ME relative to neighboring regions, indicating these clusters likely came from Arc-ME (**Supplementary Fig. 4d**). The two unassigned clusters had heterogeneous expression of regional marker genes and could not be confidently attributed to Arc-ME or surrounding regions.

While inclusion of non-Arc neurons represents a dissection artifact, it is noteworthy for two reasons. First, the presence of cells from beyond the rostral, caudal and lateral boundaries of the Arc suggests that our samples included the full extent of the Arc. Second, the fact that non-Arc neurons clustered apart from Arc neurons indicates that regional identity genes strongly influence clustering and can be used *post hoc* to determine origin.

After excluding non-Arc neuron clusters and those expressing markers of previously described Arc-ME neuron populations, 12 neuronal clusters remained, representing newly identified types of Arc-ME neurons. Some of the markers identified for these types were previously detected in rodent Arc-ME but not known to mark distinct types of neurons (*Qrfp*, *Nts*, *Trh*, *Tbx19*; **Fig. 3c**)[23–26]. We highlight four of the neuron types here: n11.Trh/Cxcl12, n19.Gpr50, n26.Htr3b and n27.Tbx19 (**Fig. 5a**). Allen Mouse Brain Atlas data confirm their marker genes to be transcribed in Arc (**Fig. 5b**). Focusing on neuropeptides and receptor expression profiles, we found that one type, n19.Gpr50 neurons, expressed several receptor genes linked to energy balance, including ghrelin receptor (*Ghsr*), melanocortin 4 receptor (*Mc4r*), cannabinoid receptor (*Cnr1*) and calcitonin receptor (*Calcr*) (**Fig. 5c**). In addition, n27.Tbx19 neurons were enriched for corticotrophin-releasing hormone 2 receptor (*Crhr2*) (**Fig. 5c**), potentially linking these neurons with known cardiovascular effects of CRH2 receptor agonism in Arc-ME[27,28].

The n11.Trh/Cxcl12 cluster had the highest expression of *Lepr* among all Arc-ME clusters (**Fig. 5c**). To confirm functional LepR expression *in vivo*, we administered leptin to fasted mice and compared fluorescent *Trh* (thyrotropin releasing hormone) reporter (Trh-IRES-Cre::AAV-DIO-mCherry) expression with phosphorylated STAT3 (pSTAT3) immunofluorescence, which marks leptin-sensing neurons[29]. Most mCherry$^+$ neurons in the Arc-ME were also pSTAT3 immunofluorescent after leptin treatment, confirming their sensitivity to leptin (**Fig. 5d**; 62 of 106 mCherry$^+$ neurons (59%) were pSTAT3$^+$; $n = 4$ mice). Our results thus identify n11.Trh/Cxcl12 neurons as a GABAergic (*Slc32a1*$^+$; see above) leptin-sensing Arc-ME population. Of note, we previously found that leptin controls body weight largely through its action on uncharacterized populations of non-AgRP GABAergic neurons that control the activity of POMC neurons[6] and AgRP neurons[18]. Our data strongly suggest that n11.Trh/Cxcl12 neurons may be one of these missing leptin-sensing populations, exemplifying how this transcriptomic resource can fill known gaps in Arc-ME knowledge. Future studies will address whether these LepR$^+$ TRH neurons provide synaptic input to POMC or AgRP neurons, as this circuit could be key to leptin-mediated control of body weight.

Arc-ME contains neurons marked by the RIP-Cre transgene that are distinct from AgRP neurons and POMC neurons[17]. These so-called RIP-Cre neurons sense leptin and drive energy expenditure but do not affect feeding[7]. However, other than their lack of *Agrp* and *Pomc* expression, these RIP-Cre neurons are transcriptionally uncharacterized. We performed single-cell RNA-seq on RIP-Cre neurons manually isolated from acutely dissociated Arc-ME (**Fig. 5e**). As expected[7], nearly all of the 25 RIP-Cre neurons we profiled appeared to be GABAergic, containing GABAergic marker transcript *Slc32a1* (21 of 25 neurons), and roughly half had *Lepr* transcript (12 of 25 neurons; **Supplementary Fig. 5e**).

To determine which Arc-ME neuron cluster(s) each RIP-Cre neuron was most similar to, we examined RIP-Cre neuron expression of Arc-ME neuron cluster markers and matched each RIP-Cre neuron to a Drop-seq neuron cluster on the basis of its expression of cluster markers. RIP-Cre neurons are quite heterogeneous, with individual neurons localized to ten different clusters, though 23 of 25 RIP-Cre neurons mapped to three distinct transcriptional clades (**Fig. 5f**,**g**). The most common clusters represented were n07.Arx/Nr5a2, n27.Tbx19 and n05.Nfix/Ht2rc. Of these, the only GABAergic cluster that expressed *Lepr* at a significant level was the n27.Tbx19 neurons (**Fig. 5c**),

making it the likeliest candidate to mediate the previously characterized function of RIP-Cre neurons in energy expenditure. Consistent with this hypothesis, n27.Tbx19 neurons showed enriched expression of *Cartpt* and *Nmu*, which encode neuropeptides that induce thermogenesis when administered centrally, as well as *Vipr2*, knockouts of which are lean and have increased metabolic rates[30–32]. These results

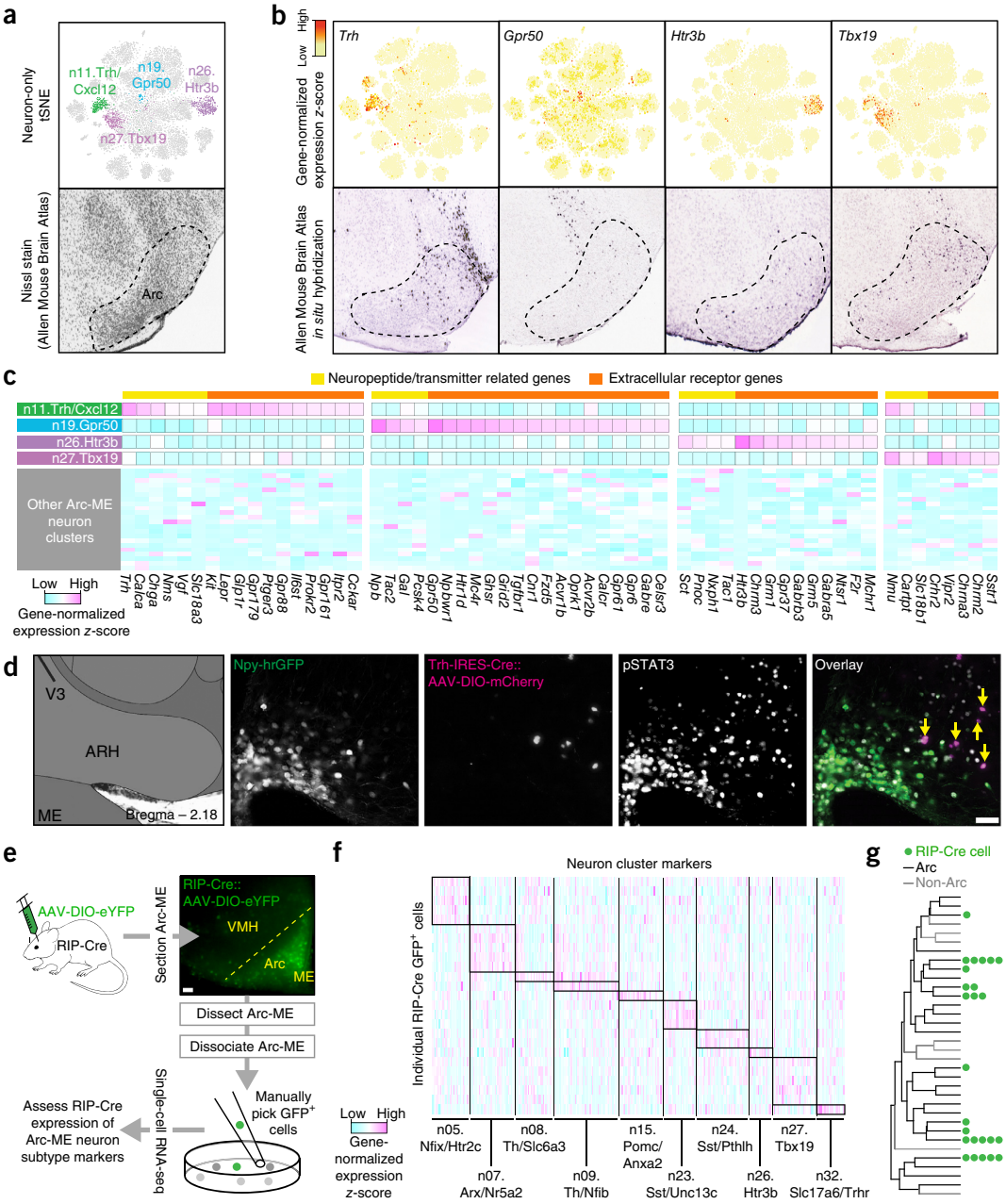

Figure 5 New subtypes of Arc-ME neurons. (**a**) Top, neuron-only tSNE plot recolored to indicate four neuron types selected for further analysis: n11.Trh/Cxcl12 neurons, n19.Gpr50 neurons, n26.Htr3b neurons and n27.Tbx19 neurons; bottom, Nissl stain of sagittal Arc-ME sections (Allen Mouse Brain Atlas). (**b**) Expression of marker genes shown by recoloring of neuron-only tSNE plot (top) and by *in situ* hybridization of sagittal brain sections (Allen Mouse Brain Atlas; bottom). (**c**) Heat map of neuropeptide and receptor genes enriched in the four neuron types. Other Arc-ME neuron types included for comparison are in numerical order (by cluster identifier). (**d**) Leptin-induced pSTAT3 immunofluorescence in the caudal Arc-ME of a fasted Trh-IRES-Cre mouse in which cells were labeled by Cre-dependent AAV-mCherry (micrograph representative of 4 mice). Yellow arrows indicate pSTAT3$^+$ mCherry$^+$ cells. Left-most panel adapted from Allen Mouse Brain Reference Atlas. V3, third ventricle; ARH, arcuate nucleus of the hypothalamus. Scale bar, 50 µm. (**e**) Single-cell RNA-seq of eYFP-labeled RIP-Cre$^+$ neurons that were acutely dissociated and manually isolated from Arc-ME of 2 adult male mice. Micrograph representative of 2 mice. VMH, ventromedial hypothalamus. Scale bar, 50 µm. (**f**) RIP-Cre neuron expression of Arc-ME neuron type markers. (**g**) Dendrogram of Arc-ME neuron types with green dots indicating the subtypes most similar to RIP-Cre neurons as based on marker expression.

demonstrate how this transcriptional database can be used to classify and characterize individual Arc-ME neurons of interest.

Transcriptional relationship predicts anatomical and functional similarities between AgRP and SST neurons

We observed that GABAergic (*Slc32a1*$^+$) neuron clusters n13.Agrp/Gm8773 and n23.Sst/Unc13c were coenriched in several transcripts (**Fig. 6a**), including those encoding a transcription factor (*Otp*) and the calcitonin receptor (*Calcr*) (**Fig. 6a**,**b**). In addition, we found a cluster of neurons (n12.Agrp/Sst) expressing both *Agrp* and *Sst* at high levels (**Fig. 6c**) with a transcriptional profile that included both n23.Sst/Unc13c and n13.Agrp/Gm8773 cluster markers (**Supplementary Fig. 6a**). Overall, 9% of *Agrp*$^+$ neurons also expressed *Sst* and 19% of all *Sst*$^+$ neurons also expressed *Agrp*, above gene-specific expression thresholds (calculated as in ref. 10). Of note, transcripts distinguishing n12.Agrp/Sst neurons from other AgRP neurons (n13.Agrp/Gm8773) were largely enriched in other *Sst*$^+$ neuron clusters as well, suggesting a common somatostatinergic gene program (**Supplementary Fig. 6b**,**c**). To confirm that neurons coexpressing AgRP and SST exist *in vivo*, we compared fluorescent *in situ* hybridization (FISH) for *Sst* mRNA to Agrp-IRES-Cre::loxSTOPlox-GFP immunofluorescence, which marks AgRP neurons[33]. The results showed that 5% of AgRP neurons were *Sst*$^+$ (38 of 779 neurons) and 9% of *Sst*$^+$ neurons were AgRP neurons (38 of 442 neurons; **Fig. 6d**), confirming that neurons coexpressing AgRP and SST exist *in vivo*. While these neurons represent a small minority of all Arc-ME AgRP neurons, our findings provide the first evidence for transcriptional subtypes of AgRP neurons. Of note, a recent study found that deleting *Crhr1* from AgRP neurons alters hepatic glucose production and thermogenesis but not feeding or body weight[34]. Our data show *Crhr1* to be expressed predominantly by the n12.Agrp/Sst subtype (**Supplementary Fig. 6b**), predicting that this subtype mediates AgRP neuron control of glucoregulation and metabolism specifically.

The transcriptional similarities between AgRP neurons and SST neurons led us to consider other potential similarities, such as in their synaptic circuitry and function. Previous studies in rat showed that Arc SST (ARC^{SST}) neurons and AgRP neurons both innervate the paraventricular hypothalamus (PVH)[35]. To confirm this in mice and further compare AgRP neuron and ARC^{SST} neuron circuitry, we injected Arc-ME of Sst-IRES-Cre mice with Cre-dependent AAV-channelrhodopsin2 (ChR2)-mCherry and examined AgRP immunofluorescence, thus visualizing ARC^{SST} axons and AgRP axons, respectively. We found that ARC^{SST} neurons and AgRP neurons had similar projection patterns, both innervating regions where AgRP neurons control feeding, such as the PVH, paraventricular thalamus and bed nucleus of the stria terminalis (**Fig. 6e** and **Supplementary Fig. 6d**)[36].

The PVH is a major synaptic target of AgRP neurons in their control of feeding behavior[37,38]. Since we observed ARC^{SST} axons in PVH, we tested whether they form functional synapses on PVH neurons. Specifically, we performed channelrhodopsin-assisted circuit mapping[39] with mice in which ChR2 was selectively expressed in ARC_{SST} neurons (see above; $n = 2$ mice). Photostimulating ARC^{SST} axons in the PVH evoked time-locked postsynaptic currents in ~12% of PVH neurons (3 of 26), indicating that ARC^{SST} neurons release fast-acting neurotransmitters onto a subset of PVH neurons (**Fig. 6f**). Bath application of a GABA receptor antagonist, bicuculline, abolished the light-evoked postsynaptic currents, revealing the GABAergic nature of these synapses (**Fig. 6f**). Thus, ARC^{SST} neurons, like AgRP neurons, directly inhibit a subset of PVH neurons.

We next determined whether chemogenetic activation of ARC^{SST} neurons could affect feeding, as has been robustly shown for AgRP neurons[4]. After injecting Arc-ME of Sst-IRES-Cre mice with a Cre-dependent AAV expressing hM3Dq, a mutant G_q-coupled muscarinic receptor[40], we activated ARC^{SST} neurons using the hM3Dq ligand, clozapine-*N*-oxide, (CNO; i.p.; $n = 5$ mice). Administering CNO to these mice acutely and significantly increased feeding (**Fig. 6g**), thus establishing ARC^{SST} neurons as a new orexigenic population whose activation is sufficient to drive feeding behavior. Notably, previous studies have shown that adult mice can recover 3 weeks after total ablation of AgRP neurons[41]. Whether this recovery results from an increased role of ARC^{SST} neurons to compensate for AgRP neuron loss is a question for further investigation. Further experiments are also needed to determine which of the five *Sst*$^+$ Arc-ME neuron subtypes is or are responsible for the feeding phenotype we observed. Our gene expression data provide subtype-specific markers from which to design genetic tools to specifically manipulate each *Sst*$^+$ subtype (**Fig. 2** and **Supplementary Fig. 6e**,**f**).

In vivo metabolic stress triggers generalized as well as cell type– and subtype-specific changes in gene expression

Studies of how *in vivo* perturbations affect gene expression are typically performed at the tissue level, or at best on samples of cells pooled on the basis of transgenic reporter expression. Such studies may be limited by their dependence on the availability or fidelity of transgenic reporter animals and, more importantly, their presumption of homogeneity within a target cell population. Unbiased single-cell approaches such as Drop-seq have the potential to overcome these limitations but so far have been used mostly to parse cell types from complex tissues, where transcriptional differences are large and static. To find out whether dynamic transcriptional changes can be assessed, we performed Drop-seq profiling on Arc-ME from mice across different feeding conditions and energy states: *ad libitum* fed or fasted overnight, or 1 week of low-fat diet (10% calories from fat; LFD) or high-fat diet (60% calories from fat; HFD). We identified thousands of genes significantly up- or downregulated in response to fasting and HFD across individual non-neuronal (**Supplementary Fig. 7a** and **Supplementary Tables 5** and **6**) and neuronal populations (**Fig. 7a** and **Supplementary Tables 7** and **8**). Generally, fasting induced a stronger transcriptional response than HFD (that is, greater number and amplitude of changes; **Fig. 7a** and **Supplementary Fig. 7b**), and the neuron types most responsive to fasting also tended to be most responsive to HFD (**Fig. 7a**), suggesting energy state–responsive neuron types.

To globally validate gene expression changes, we compared them to a recently published RNA-seq study on pooled neurons marked by AgRP or POMC transgenic reporter expression[21]. We found strong correlations between AgRP neuron subtypes and pooled AgRP neurons, and between POMC neuron subtypes and pooled POMC neurons (**Supplementary Fig. 7c–f**), demonstrating that Drop-seq detects gene expression changes found by more conventional methods.

To further test whether the AgRP and POMC neuron subtypes we found are functionally distinct, we compared their transcriptional responses to fasting. While most genes changed similarly across the AgRP and POMC neuron populations (for example, *Agrp*, *Pomc*, *Vgf*), many were differentially regulated between their subtypes (**Fig. 7b**,**c**). Some of these were fasting-sensitive genes identified in pooled AgRP and POMC neuron samples[21], but which we found to be fasting-sensitive only in one subtype (in AgRP neurons: *Bhlhe40*, *Gria3*, in POMC neurons: *Tmtc4*, *Gm22426*; **Fig. 7c**). We also identified gene regulatory events that were discordant in different subtypes and had been missed in a bulk-tissue analysis, a benefit of the single-cell resolution of Drop-seq. For example, fasting did not significantly alter the

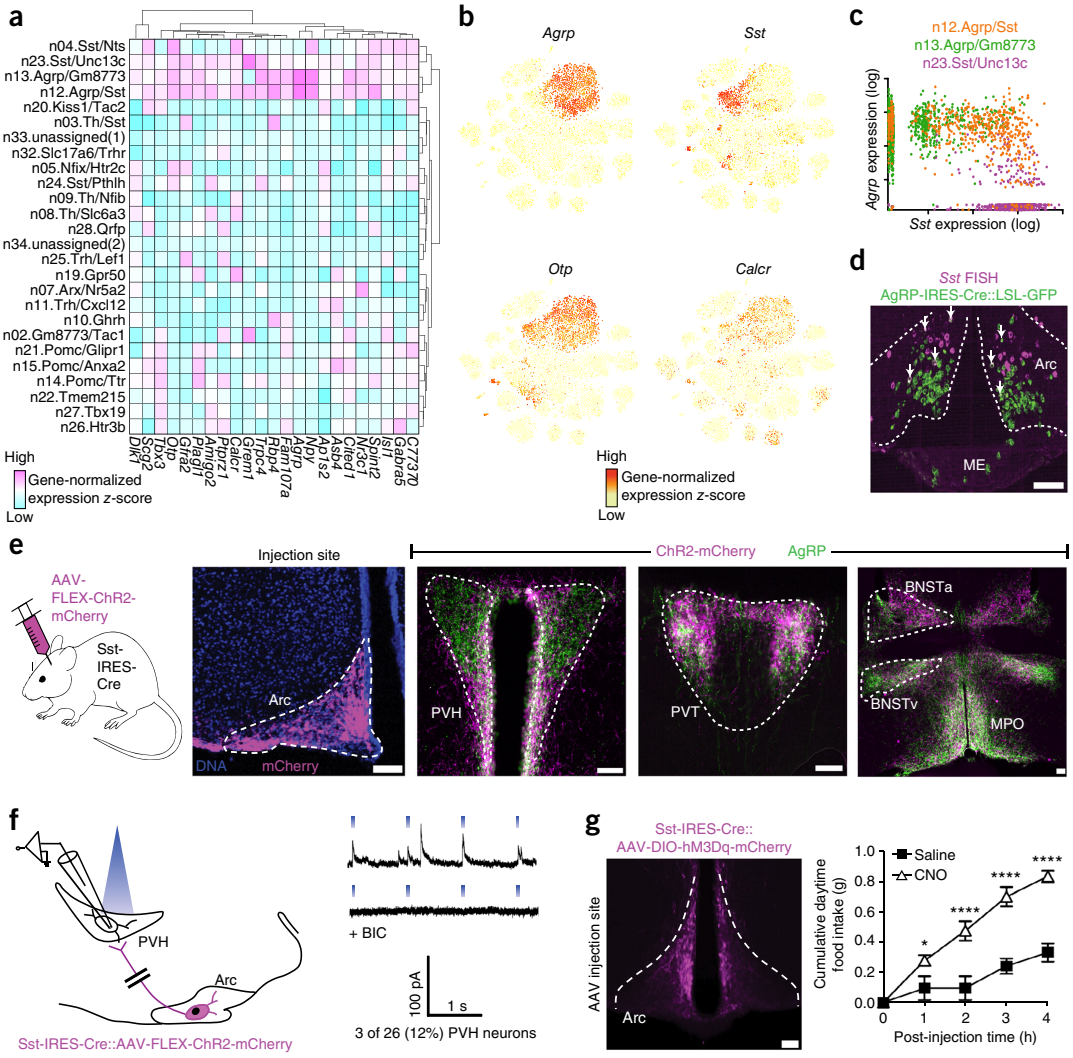

Figure 6 Similarities between AgRP neurons and SST neurons. (**a**) Heat map of gene coenriched in AgRP neurons and SST neurons. (**b**) Expression of two coenriched genes, *Otp* and *Calcr*, as well as *Agrp* and *Sst*, shown by recoloring of neuron-only tSNE plots. (**c**) Coexpression of *Agrp* and *Sst* by individual cells in three neuron clusters: n12.Agrp/Sst, n13.Agrp/Gm8773 and n23.Sst/Unc13c. (**d**) Representative micrograph comparing Agrp-IRES-Cre::loxSTOPlox-GFP immunofluorescence to Sst mRNA *in situ* hybridization. White arrows indicate colabeled cells. (**e**) Axon projection patterns of AgRP neurons and Arc-ME SST neurons. From left to right, micrographs of mCherry immunofluorescence at the site where Cre-dependent AAV-ChR2-mCherry was injected in the Arc-ME of a Sst-IRES-Cre mouse and comparisons of mCherry and AgRP immunofluorescence in paraventricular hypothalamus (PVH); paraventricular thalamus (PVT); bed nucleus of the stria terminalis, anterior and ventral parts (BNSTa and BNSTv, respectively); and medial preoptic nucleus (MPO). Micrographs are representative of 3 mice. Scale bars, 100 µm. (**f**) Left, schematic of channelrhodopsin-assisted circuit mapping from ChR2$^+$ ARC^{SST} neurons to unidentified PVH neurons (n = 2 mice). Right, representative patch-clamp recordings of PVH neurons during photostimulation of ARC^{SST} neuron axons in PVH, in the absence (top) or presence (bottom) of the bath-applied $GABA_A$ receptor antagonist bicuculline (BIC); blue dash indicates light pulse. (**g**) Effect of chemogenetic stimulation of ARC^{SST} neurons on daytime food intake. Left, micrograph of mCherry immunofluorescence at the site where Cre-dependent AAV-hM3Dq-mCherry was injected bilaterally in the Arc-ME of an Sst-IRES-Cre mouse; representative of 5 mice; scale bar, 100 µm. Right, cumulative food intake in 4-h period after intraperitoneal injection of either the hM3Dq ligand (CNO) or vehicle (saline). N = 5 mice; data shown as mean ± s.e.m. *P = 0.0315, ****P < 0.0001; two-way ANOVA followed by Sidak's multiple comparisons test.

activity-dependent synaptogenic marker gene *Gap43* in published pooled data for AgRP neurons (FDR > 0.25)[21], but Drop-seq analysis revealed that *Gap43* was significantly upregulated by fasting in the n13.Agrp/Gm8773 neurons (FDR = 0.08) and significantly downregulated in n12.Agrp/Sst neurons (FDR = 0.04); **Fig. 7c**; see also *Btg2*, *Rin2*). Overall, 647 genes had a highly significant (FDR < 0.05) change in expression in one AgRP neuron subtype but not in the AgRP neuron population as a whole. We found similar examples of subtype-specific transcriptional responses in POMC neurons, including the neuropeptide Y receptor gene *Npy2r* (**Fig. 7c**; see also *Dcdc2a*, *Tmem237*, *Coro7*). In total, fasting altered expression of 42 genes to a highly significant degree (FDR < 0.05) in one or two POMC neuron subtypes that were not significantly altered in the whole POMC population. The observed differences in how individual AgRP and

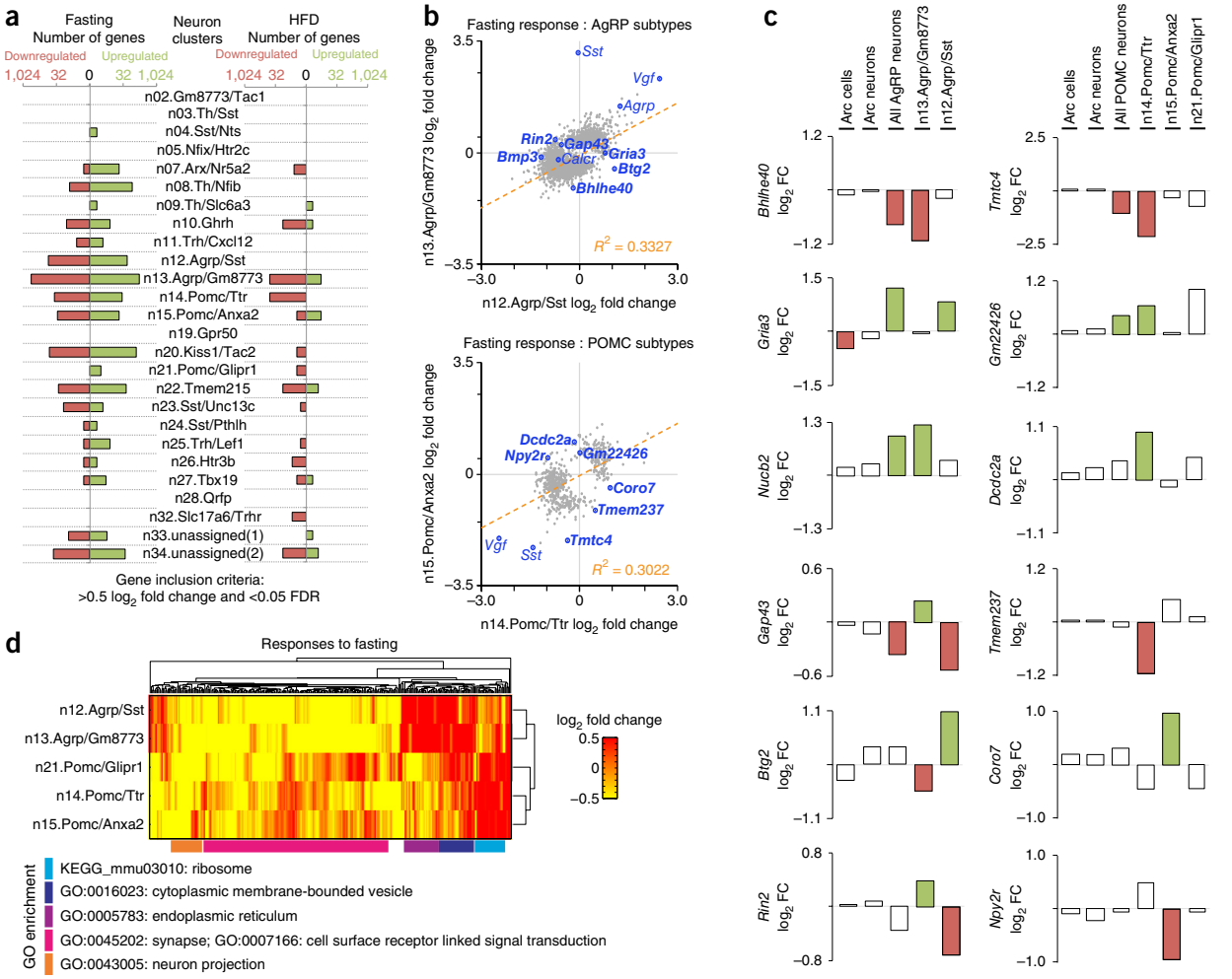

Figure 7 Transcriptional responses to energy imbalance. (**a**) Histograms showing the number of genes significantly up- or downregulated in response to fasting and HFD in each Arc-ME neuron type. (**b**) Comparison of fasting responses of AgRP neuron subtypes (top) and POMC neuron subtypes (bottom). Genes plotted were significantly affected by fasting (false-discovery rate (FDR) < 0.25 in at least one AgRP neuron subtype (top) or POMC neuron subtype (bottom). While subtypes generally show significant correlation (Pearson correlation = 0.58, $P = 2 \times 10^{-237}$ for AgRP neuron subtypes and Pearson correlation = 0.55, $P = 2 \times 10^{-47}$ using Student's *t*-distribution test for POMC neuron subtypes), there are many individual genes that are differentially affected by fasting (for example, some are in the top left and some in the bottom right quadrant). Boldface indicates genes that are also shown in **c**. (**c**) Examples of genes affected by fasting only in one subtype of AgRP neurons and POMC neurons, or affected oppositely between subtypes. For comparison, average $\log_2$ fold-change (FC) values are also shown for all Arc-ME cells, all Arc-ME neurons and all AgRP or POMC neurons. Bars are shaded to indicate genes differentially expressed at FDR < 0.25; green and red indicate up- and downregulation, respectively. (**d**) Heat map of gene expression fold-change values for genes significantly affected in at least one AgRP or POMC subtype. Genes are clustered by gene expression similarities across AgRP and POMC subtypes. GO terms with highest significance for each cluster are shown.

POMC subtypes respond to fasting underscores the necessity of a single-cell approach to identifying such heterogeneity and shows that our transcriptionally defined cell subtypes make distinct functional responses to the same stimulus.

While fasting affects a subset of genes in the two AgRP subtypes differently, the transcriptional responses of these subtypes were broadly similar to each other and different from responses of the POMC subtypes (**Fig. 7d**). The POMC subtypes showed greater functional heterogeneity in their fasting responses, with the n14.Pomc/Ttr and n15.Pomc/Anxa2 subtypes showing high correlation and being distinct from n21.Pomc/Glipr1 neurons (**Fig. 7d**). Extending these comparisons to all Arc-ME neurons (**Supplementary Fig. 7g**), we found that AgRP neurons and n08.Th/Slc6a3 neurons had similar transcriptional responses to fasting that were distinct from those of other Arc-ME neurons. Specifically, AgRP neurons and n08.Th/Slc6a3 (TIDA) neurons strongly upregulated endoplasmic reticulum genes but not the ribosomal and protein translation genes upregulated by other Arc-ME neuron types. Fasting is known to activate AgRP neurons and TIDA neurons[42,43], which may explain at least some of the similarity in their transcriptional response.

Also of note, many of the fasting-induced changes in AgRP neurons were inversely regulated by HFD (**Supplementary Fig. 7h**). A specific account of all responsive genes and gene sets is beyond the scope of this manuscript; although we focus on responses of a few neuron types to fasting and feeding, we provide differential expression data for all the Arc-ME cell types, many of which are known to be important in

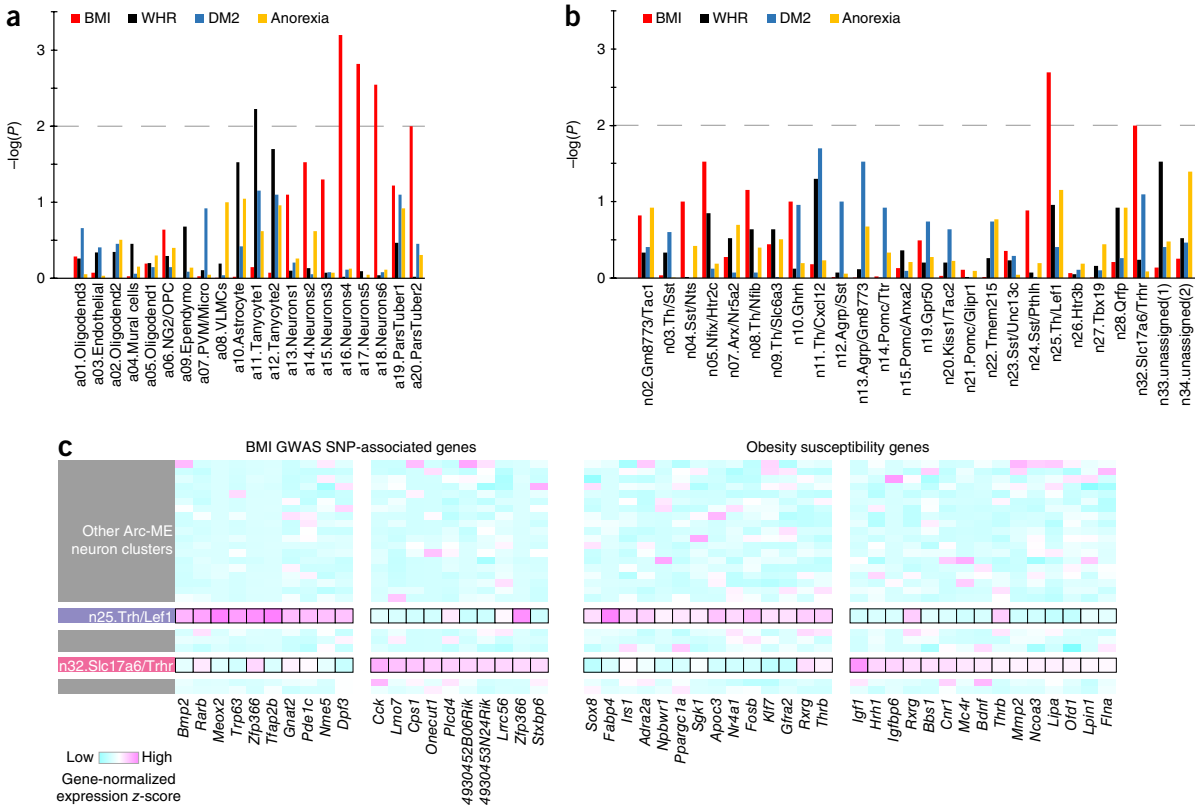

Figure 8 DEPICT predicts specific neuron types affecting BMI. (**a**) DEPICT predicts that transcripts from BMI-associated loci ($P < 3 \times 10^{-3}$), but not loci linked to waist-to-hip ratio (WHR), type 2 diabetes mellitus (DM2) or anorexia, are enriched in neurons. The dotted line shows statistically significant enrichment. Abbreviations as in **Figure 1**. (**b**) DEPICT predicts transcripts from BMI-associated loci ($P < 1 \times 10^{-2}$) are enriched in n25.Trh/Lef1 and n32.Slc17a6/Trhr neuron clusters. (**c**) Heat map of n25.Trh/Lef1 and n32.Slc17a6/Trhr neuron cluster expression of genes near BMI-linked loci or related to obesity susceptibility. Other Arc-ME neuron types included for comparison are in numerical order by cluster identifier.

regulating responses to fasting and HFD (**Supplementary Tables 5–8**). These data should provide a resource for identification of genes whose dynamic expression in specific cell types may be important for that cell type's functional response. Overall, these results show that single-cell profiling by Drop-seq can robustly and comprehensively detect changes in gene expression across perturbations, with unprecedented sensitivity, specificity and scale. Diet-sensitive cell types and genes yield an array of new translational targets for obesity treatment.

Cell-type specific expression profiles connect candidate GWAS genes to new cell types

GWAS have uncovered thousands of single-nucleotide polymorphisms (SNPs) conferring disease risk, but determining the gene(s) each SNP regulates and the cell type(s) in which that gene acts remains a challenge and is a major barrier to deriving biological insight from such data. DEPICT (Data-driven Expression Prioritized Integration for Complex Traits) is an analytical tool designed to systematically prioritize tissue type(s) on the basis of enriched expression of GWAS-associated genes[44]. Previously, body mass index (BMI) GWAS-linked loci genes were shown to have enriched expression in CNS tissues[45]. However, this analysis was limited by the tissue-level resolution of available expression data, and it remained unclear what specific CNS cell type(s) were involved. We therefore used DEPICT with our Drop-seq data to determine which, if any, of our identified cell types mediate GWAS signals for BMI.

Normalizing to the expression profiles of the other Arc-ME cell populations, we found significant enrichment of BMI-linked gene expression in neuronal, but not non-neuronal, populations (**Fig. 8a**). This neuronal enrichment is specific, as it was not observed for anorexia, type 2 diabetes mellitus or waist-to-hip ratio GWAS loci, the last of which were previously shown to be enriched in adipose tissue and adipocytes[46]. To pinpoint the signal, we focused DEPICT on the Arc-ME neuron types and found two types, n25.Trh/Lef1 and n32.Slc17a6/Trhr, with significant BMI GWAS enrichment (**Fig. 8b**). In addition to the BMI-linked genes, we also observe that these neuron types were enriched in expression of many previously described obesity susceptibility genes (**Fig. 8c** and **Supplementary Table 9**)[47]. These analyses thus identify two types of Arc-ME neurons, newly described in this study, as clear candidates for future study in body weight regulation, and prioritize a set of obesity-associated genes whose regulation in these neurons may affect phenotype.

A parallel study from our laboratory also implicated the n32.Slc17a6/Trhr subpopulation in energy balance. Specifically, Fenselau and colleagues[48] identified a population of *Slc17a6*$^{+}$ *Oxtr*$^{+}$ Arc-ME neurons that are both necessary and sufficient for satiety. Chemogenetic activation of Arc-ME neurons expressing Slc17a6-IRES-Cre and Oxtr-Cre rapidly suppressed feeding, a timescale not seen with similar activation of POMC neurons[48]. Our Drop-seq data show that, among non-POMC neurons in the Arc-ME, *Slc17a6* and *Oxtr* are predominantly coexpressed by a single population, the n32.

Slc17a6/Trhr neurons. These results validate our DEPICT prioritization analysis, and together they suggest that, as in mice, Slc17a6/Trhr neurons control body weight in humans.

We also analyzed GWAS data for other phenotypes potentially related to Arc-ME function: growth (height), fertility (menarche, menopause) and glucose regulation (type 2 diabetes mellitus). We found enriched expression of menarche-associated genes in neurons, specifically the n04.Sst/Nts neuron type (**Supplementary Fig. 8a**,**b**). Notably, *Nts* encodes neurotensin, a neuropeptide that can directly stimulate gonadotropin releasing hormone neurons, contributing to the luteinizing hormone surge[49]. We observed enrichment for menopause-associated genes in pars tuberalis, an endocrine population that includes cells expressing the reproductive hormones luteinizing hormone and follicle-stimulating hormone (**Supplementary Fig. 8a**). Lastly, we observed highly significant enrichment for height-associated gene expression in multiple cell types, including endothelial, mural and VLMC populations, as well as in pars tuberalis cells and many neuron types (AgRP subtypes, n14.Pomc/Ttr, n20.Kiss1/Tac2, n22.Tmem215) (**Supplementary Fig. 8a**,**b**). These results agree with and refine previous DEPICT height analyses showing enrichment in cardiovascular, endocrine, skeletal, respiratory and urogenital tissues[50].

DISCUSSION

We have extended unbiased single-cell profiling to a complex brain region with myriad important physiological functions. This approach reveals many Arc-ME cell types and subtypes not previously described, despite intensive study of this region. It is important to keep in mind that our cell-type census may not be complete. Our study may have missed rare cell types, cells that do not survive dissociation or particularly large cells that were filtered out of the cell suspension (see Online Methods). Nonetheless, we establish a resource that provides molecular profiles of many Arc-ME cell types in exceptional detail.

These data will enable rational design and discovery of tools to manipulate Arc-ME cell types with unprecedented specificity. We show for the first time, to our knowledge, that Drop-seq can assess dynamic transcriptional responses comprehensively across cell types *in vivo* and that our *ab initio*-inferred cell types respond differently to *in vivo* perturbations, affirming the close relationship between a cell's transcriptional and functional identities. Finally, we show how detailed cell type transcriptional profiles can unlock the full power of GWAS studies. Overall these studies suggest a multitude of testable hypotheses and provide the means to address them. We look forward to seeing how this resource is used to guide future studies.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We gratefully acknowledge Z. Yang, J. Madara and C. Wu for technical assistance, A. Garfield for editorial advice and P. Kharchenko for advice on SCDE. Quantitative PCR and confocal imaging were done at BIDMC's Molecular Medicine Core and Confocal Imaging Core, respectively. Funding was provided by US National Institutes of Health grants to B.B.L. (R01 DK096010, R01 DK089044, R01 DK071051, R01 DK075632, R37 DK053477, BNORC Transgenic Core P30 DK046200, BADERC Transgenic Core P30 DK057521), E.D.R. (R01 DK102170, R01 DK085171, R01 DK102173), E.D.R. and L.T.T. (BNORC Functional Genomics Core P30 DK046200), L.T.T. (BADERC Pilot and Feasibility grant NIH 2P30DK057521-16) and J.M.R. (F32 DK103387); a Department of Defense grant to L.T.T. (Discovery Award W81XWH-15-1-0251); an American Heart Association Postdoctoral Fellowship to J.N.C. (14POST20100011); the Lundbeck Foundation and the Benzon Foundation (T.H.P.); the Stanley Center for Psychiatric Research (S.A.M.); and the Stanley-MGH Fellowship in Psychiatric Neuroscience (E.Z.M.).

AUTHOR CONTRIBUTIONS

J.N.C., L.T.T., E.Z.M., S.A.M., E.D.R. and B.B.L. conceived the study. J.N.C., L.T.T., E.Z.M., E.D.R. & B.B.L. designed the study. J.N.C., A.M.J.V. and L.T.T. prepared samples for Drop-seq. L.T.T., D.T., J.N.C., E.Z.M. and M.G. did Drop-seq. D.T., M.G. and L.T.T. made Drop-seq libraries. J.N.C. did single-cell RNA-seq. L.T.T., J.N.C., A.L. and E.Z.M. analyzed transcriptomic data. J.M.R. did *in situ* hybridization. J.N.C. and H.F. did histology and imaging, with advice from B.B.L. H.F. did electrophysiology. H.F. and A.M.J.V. did stereotaxic injections and feeding studies. T.H.P. performed DEPICT analyses. J.N.C., L.T.T. and A.L. prepared figures. J.N.C., L.T.T., S.A.M., E.D.R. and B.B.L. wrote the manuscript with input from all authors.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Rodríguez, E.M., Blázquez, J.L. & Guerra, M. The design of barriers in the hypothalamus allows the median eminence and the arcuate nucleus to enjoy private milieus: the former opens to the portal blood and the latter to the cerebrospinal fluid. *Peptides* **31**, 757–776 (2010).
2. Langlet, F. *et al.* Tanycytic VEGF-A boosts blood-hypothalamus barrier plasticity and access of metabolic signals to the arcuate nucleus in response to fasting. *Cell Metab.* **17**, 607–617 (2013).
3. Aponte, Y., Atasoy, D. & Sternson, S.M. AGRP neurons are sufficient to orchestrate feeding behavior rapidly and without training. *Nat. Neurosci.* **14**, 351–355 (2011).
4. Krashes, M.J. *et al.* Rapid, reversible activation of AgRP neurons drives feeding behavior in mice. *J. Clin. Invest.* **121**, 1424–1428 (2011).
5. Han, S.Y., McLennan, T., Czieselsky, K. & Herbison, A.E. Selective optogenetic activation of arcuate kisspeptin neurons generates pulsatile luteinizing hormone secretion. *Proc. Natl. Acad. Sci. USA* **112**, 13109–13114 (2015).
6. Vong, L. *et al.* Leptin action on GABAergic neurons prevents obesity and reduces inhibitory tone to POMC neurons. *Neuron* **71**, 142–154 (2011).
7. Kong, D. *et al.* GABAergic RIP-Cre neurons in the arcuate nucleus selectively regulate energy expenditure. *Cell* **151**, 645–657 (2012).
8. Macosko, E.Z. *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
9. Marques, S. *et al.* Oligodendrocyte heterogeneity in the mouse juvenile and adult central nervous system. *Science* **352**, 1326–1329 (2016).
10. Usoskin, D. *et al.* Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. *Nat. Neurosci.* **18**, 145–153 (2015).
11. Vijayan, E., Samson, W.K. & McCann, S.M. In vivo and in vitro effects of cholecystokinin on gonadotropin, prolactin, growth hormone and thyrotropin release in the rat. *Brain Res.* **172**, 295–302 (1979).
12. Morgan, P.J. & Williams, L.M. The pars tuberalis of the pituitary: a gateway for neuroendocrine output. *Rev. Reprod.* **1**, 153–161 (1996).
13. Goodman, T. & Hajihosseini, M.K. Hypothalamic tanycytes-masters and servants of metabolic, neuroendocrine, and neurogenic functions. *Front. Neurosci.* **9**, 387 (2015).
14. Lein, E.S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).
15. Tesfaigzi, J. & Carlson, D.M. Expression, regulation, and function of the SPR family of proteins. A review. *Cell Biochem. Biophys.* **30**, 243–265 (1999).
16. Everitt, B.J. *et al.* The hypothalamic arcuate nucleus-median eminence complex: immunohistochemistry of transmitters, peptides and DARPP-32 with special reference to coexistence in dopamine neurons. *Brain Res.* **396**, 97–155 (1986).
17. Choudhury, A.I. *et al.* The role of insulin receptor substrate 2 in hypothalamic and beta cell function. *J. Clin. Invest.* **115**, 940–950 (2005).
18. Garfield, A.S. *et al.* Dynamic GABAergic afferent modulation of AgRP neurons. *Nat. Neurosci.* **19**, 1628–1635 (2016).
19. Lyons, D.J., Hellysaz, A. & Broberger, C. Prolactin regulates tuberoinfundibular dopamine neuron discharge pattern: novel feedback control mechanisms in the lactotrophic axis. *J. Neurosci.* **32**, 8074–8083 (2012).
20. Demaria, J.E. *et al.* Dopamine transporters participate in the physiological regulation of prolactin. *Endocrinology* **141**, 366–374 (2000).
21. Henry, F.E., Sugino, K., Tozer, A., Branco, T. & Sternson, S.M. Cell type-specific transcriptomics of hypothalamic energy-sensing neuron responses to weight-loss. *Elife* **4**, 09800 (2015).
22. Sohn, J.W. *et al.* Serotonin 2C receptor activates a distinct population of arcuate pro-opiomelanocortin neurons via TRPC channels. *Neuron* **71**, 488–497 (2011).
23. Kahn, D., Abrams, G.M., Zimmerman, E.A., Carraway, R. & Leeman, S.E. Neurotensin neurons in the rat hypothalamus: an immunocytochemical study. *Endocrinology* **107**, 47–54 (1980).
24. Nishiyama, T. *et al.* Hypothalamic thyrotropin-releasing hormone (TRH)-containing neurons involved in the hypothalamic-hypophysial-thyroid axis. Light microscopic immunohistochemistry. *Brain Res.* **345**, 205–218 (1985).

25. Liu, J. *et al.* Tbx19, a tissue-selective regulator of POMC gene expression. *Proc. Natl. Acad. Sci. USA* **98**, 8674–8679 (2001).
26. Takayasu, S. *et al.* A neuropeptide ligand of the G protein-coupled receptor GPR103 regulates feeding, behavioral arousal, and blood pressure in mice. *Proc. Natl. Acad. Sci. USA* **103**, 7438–7443 (2006).
27. Chitravanshi, V.C., Kawabe, K. & Sapru, H.N. Mechanisms of cardiovascular actions of urocortins in the hypothalamic arcuate nucleus of the rat. *Am. J. Physiol. Heart Circ. Physiol.* **305**, H182–H191 (2013).
28. Reyes, T.M. *et al.* Urocortin II: a member of the corticotropin-releasing factor (CRF) neuropeptide family that is selectively bound by type 2 CRF receptors. *Proc. Natl. Acad. Sci. USA* **98**, 2843–2848 (2001).
29. Vaisse, C. *et al.* Leptin activation of Stat3 in the hypothalamus of wild-type and ob/ob mice but not db/db mice. *Nat. Genet.* **14**, 95–97 (1996).
30. Kong, W. *et al.* A role for arcuate cocaine and amphetamine-regulated transcript in hyperphagia, thermogenesis, and cold adaptation. *FASEB J.* **17**, 1688–1690 (2003).
31. Asnicar, M.A. *et al.* Vasoactive intestinal polypeptide/pituitary adenylate cyclase-activating peptide receptor 2 deficiency in mice results in growth retardation and increased basal metabolic rate. *Endocrinology* **143**, 3994–4006 (2002).
32. Novak, C.M., Zhang, M. & Levine, J.A. Neuromedin U in the paraventricular and arcuate hypothalamic nuclei increases non-exercise activity thermogenesis. *J. Neuroendocrinol.* **18**, 594–601 (2006).
33. Tong, Q., Ye, C.P., Jones, J.E., Elmquist, J.K. & Lowell, B.B. Synaptic release of GABA by AgRP neurons is required for normal regulation of energy balance. *Nat. Neurosci.* **11**, 998–1000 (2008).
34. Kuperman, Y. *et al.* CRFR1 in AgRP neurons modulates sympathetic nervous system activity to adapt to cold stress and fasting. *Cell Metab.* **23**, 1185–1199 (2016).
35. Moga, M.M. & Saper, C.B. Neuropeptide-immunoreactive neurons projecting to the paraventricular hypothalamic nucleus in the rat. *J. Comp. Neurol.* **346**, 137–150 (1994).
36. Betley, J.N., Cao, Z.F., Ritola, K.D. & Sternson, S.M. Parallel, redundant circuit organization for homeostatic control of feeding behavior. *Cell* **155**, 1337–1350 (2013).
37. Atasoy, D., Betley, J.N., Su, H.H. & Sternson, S.M. Deconstruction of a neural circuit for hunger. *Nature* **488**, 172–177 (2012).
38. Garfield, A.S. *et al.* A neural basis for melanocortin-4 receptor-regulated appetite. *Nat. Neurosci.* **18**, 863–871 (2015).
39. Petreanu, L., Huber, D., Sobczyk, A. & Svoboda, K. Channelrhodopsin-2-assisted circuit mapping of long-range callosal projections. *Nat. Neurosci.* **10**, 663–668 (2007).
40. Armbruster, B.N., Li, X., Pausch, M.H., Herlitze, S. & Roth, B.L. Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proc. Natl. Acad. Sci. USA* **104**, 5163–5168 (2007).
41. Wu, Q., Boyle, M.P. & Palmiter, R.D. Loss of GABAergic signaling by AgRP neurons to the parabrachial nucleus leads to starvation. *Cell* **137**, 1225–1234 (2009).
42. Hahn, T.M., Breininger, J.F., Baskin, D.G. & Schwartz, M.W. Coexpression of Agrp and NPY in fasting-activated hypothalamic neurons. *Nat. Neurosci.* **1**, 271–272 (1998).
43. Hsueh, Y.C., Cheng, S.M. & Pan, J.T. Fasting stimulates tuberoinfundibular dopaminergic neuronal activity and inhibits prolactin secretion in oestrogen-primed ovariectomized rats: involvement of orexin A and neuropeptide Y. *J. Neuroendocrinol.* **14**, 745–752 (2002).
44. Pers, T.H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6**, 5890 (2015).
45. Locke, A.E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197–206 (2015).
46. Shungin, D. *et al.* New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518**, 187–196 (2015).
47. Vimaleswaran, K.S. *et al.* Candidate genes for obesity-susceptibility show enriched association within a large genome-wide association study for BMI. *Hum. Mol. Genet.* **21**, 4537–4542 (2012).
48. Fenselau, H. *et al.* A rapidly acting glutamatergic ARC→PVH satiety circuit postsynaptically regulated by α-MSH. *Nat. Neurosci.* **20**, 42–51 (2017).
49. Boules, M., Li, Z., Smith, K., Fredrickson, P. & Richelson, E. Diverse roles of neurotensin agonists in the central nervous system. *Front. Endocrinol. (Lausanne)* **4**, 36 (2013).
50. Wood, A.R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46**, 1173–1186 (2014).

ONLINE METHODS

Mice. All animal care and experimental procedures were approved in advance by the National Institute of Health and Beth Israel Deaconess Medical Center Institutional Animal Care and Use Committee.

The Drop-seq experiments used a total of 53 adult (4–12 weeks old) virgin male and female mice randomly assigned to experimental groups and processed in 5 sample batches (see **Supplementary Fig. 1a**). Pubertal stage and day of estrous cycle were not determined. Mice were generated by crossing Agrp-IRES-Cre (ref. 33) mice or Pomc-Cre mice (ref. 51) with transgenic loxSTOPlox-GFP mice, all of which had been backcrossed to C57BL6/J background for at least nine generations. The CCK and ARC^{SST} immunofluorescence experiments used Cck-IRES-Cre::loxSTOPlox-GFP, Sst-IRES-Cre::loxSTOPlox-GFP and Sst-IRES-Cre heterozygous adult (6–12 weeks old) male mice derived from crossing Cck-IRES-Cre (ref. 52) and Sst-IRES-Cre (ref. 52) homozygous mice (Jackson Laboratories, JAX; strains 012706 and 013044, respectively) with loxSTOPlox-GFP homozygous mice or wild-type mice (C57BL6/J and mixed background, respectively). The leptin and pSTAT3 immunofluorescence experiments used adult male Trh-IRES-Cre (ref. 53) and Trh-IRES-Cre::Npy-hrGFP (ref. 54) mice (mixed background). The *in situ* hybridization experiment used an Agrp-IRES-Cre::loxSTOPlox-GFP adult male mouse bred from Agrp-IRES-Cre and loxSTOPlox-GFP mice (both of C57BL/6J background). The electrophysiological studies used adult male and female Sst-IRES-Cre mice bred from Sst-IRES-Cre homozygous mice (JAX) and wild-type mice (mixed background). The food intake studies used Sst-IRES-Cre adult males bred from Sst-IRES-Cre homozygous mice (JAX strain 013044) and wild-type mice (mixed background). The Arc-ME RIP-Cre studies used adult male RIP-Cre heterozygous mice[55] (JAX strain 003573).

Housing and diets. Mice were housed at 22–24 °C with a 12 h light:12 h dark cycle and *ad libitum* access to standard mouse chow (Teklad F6 Rodent Diet 8664; 4.05 kcal per g, 3.3 kcal per g metabolizable energy, 12.5% kcal from fat; Harlan Teklad) and water. For high-fat diet (HFD) experiments, mice were switched from standard chow to high-fat chow (Research Diets, Inc. D12492, 60 kcal% fat) or to a control, low-fat chow (Research Diets, Inc. D12450B, 10 kcal% fat) for 1 week before sampling. All diets were provided as pellets.

Generation of Arc-ME single-cell suspensions. Brains were rapidly extracted, cooled in ice-cold DMEM/F12 medium for 5 min and then placed, ventral surface up, into a chilled stainless steel brain matrix (catalog no. SA-2165, Roboz Surgical Instrument Co., Gaithersburg, MD). Using GFP fluorescence to demarcate the arcuate's rostral and caudal boundaries, brains were blocked to obtain a single coronal section containing the entire GFP^+ arcuate, ~2 mm thick. The Arc-ME was microdissected by knife cuts at its visually approximated dorsolateral borders and pooled by experimental condition. For example, in batch 5, we pooled tissue from all fasted males into a single sample before digestion and dissociation.

Pooled tissue samples were digested in papain solution with DNase for 1 h at 37 °C with gentle agitation, according to a published protocol[56] using an entire vial (≥100 units) of papain in a 1.6-mL final reaction volume. Digestion was stopped by dilution with EBSS #2 (ref. 56), with 10% FBS used in place of the ovomucoid protease inhibitor. Tissue was centrifuged at 300*g* for 5 min, resuspended in the modified EBSS #2 and then gently triturated using a series of five Pasteur pipettes with tips fire-polished to incrementally smaller openings. The resulting cell suspension was divided between two microfuge tubes and diluted with Drop-seq buffer[8] to a volume of 1.5 mL per tube. Cells were then washed twice by centrifugation at 300*g* for 5 min, followed by resuspension in Drop-seq buffer. Drop-seq buffer consisted of 5% (wt/vol) trehalose, Hank's buffered salt solution (magnesium- and calcium-free), 2.13 mM $MgCl_2$, 2 mM $MgSO_4$, 1.26 mM $CaCl_2$, 1 mM glucose and 0.01% bovine serum albumin (molecular biology grade). Cells were filtered through a 20-µm mesh, diluted to 220 cells/µL using cell concentration estimates from a hemocytometer and then kept on ice until use.

Drop-seq generation of single-cell sequencing libraries. Drop-seq was performed as in ref. 8 but with the following modifications: first, flow rates of 2.1 mL/h were used for each aqueous suspension and 12 mL/h for the oil[8]. Second, libraries were sequenced on the Illumina NextSeq500, using 1.7 pM in a volume of 1.2 mL HT1 and 3 mL of 0.3 µM Read1CustSeqB (GCCTGTCCGCGGAAGC AGTGGTATCAACGCAGAGTAC). Read 1 was 20 bp (bases 1–12 cell barcode, bases 13–20 UMI), read 2 (paired end) was 60 bp, and the index primer was 8 bp (on multiplexed samples).

Drop-seq read alignment and generation of digital expression data. Raw sequence data was first filtered to remove all read pairs with a barcode base quality of less than 10. The second read (60 bp) was then trimmed at the 5′ end to remove any TSO adaptor sequence and at the 3′ end to remove poly(A) tails of length 6 or greater, then aligned to the mouse (mm10) genome using STAR v2.4.0 a with default setting. Uniquely mapped reads were grouped by cell barcode. To digitally count gene transcripts, a list of UMIs in each gene, within each cell, was assembled, and UMIs within ED = 1 were merged together. The total number of unique UMI sequences was counted, and this number was reported as the number of transcripts of that gene for a given cell. All cell barcodes in which 800 or more genes were detected were used in downstream analysis, while the remaining cell barcodes were discarded, resulting in a 21,241-cell barcode data set.

Unsupervised dimensionality reduction and clustering. Mouse Arc-ME suspensions were processed through Drop-Seq in 11 separate groups over six separate batches (one batch per day), and each sequenced separately. Raw digital expression matrices were generated for the seven sequencing runs. The full 21,241 cells were merged together in a single matrix. Gene expression was normalized to library size and genes showing expression in >50 cells were retained for clustering. Before clustering, batch effect correction was performed using the removeBatchEffect function of edgeR[57]. All calculations and data analysis were then performed in log space (that is, ln(CPM + 1)).

We used Seurat software to perform clustering as per ref. 58. We identified genes that were most variable across the entire data set, controlling for the known relationship between mean expression and variance. We calculated the mean and dispersion (variance/mean) for each gene across all cells and placed genes into 300 bins on the basis of their average expression. Within each bin, we then identified outlier genes whose dispersion was greater than the median dispersion value plus the difference between the median and minimum dispersion. This process yielded a total of 2,251 significantly variable genes.

Principal component analysis was performed using the prcomp function in R, after scaling and centering the data along each variable gene. To distinguish PCs for further analysis, we plotted the cumulative variance accounted for by each successive PC. Such data display a 'knee' at a PC number after which successive PCs explain diminishing degrees of variance. Empirically, we found that downstream clustering analyses were optimized when using this PC cutoff. 25 PCs were chosen for further all-cell clustering analyses and used as input for *t*-distributed stochastic neighbor embedding (tSNE)[59], implemented by the Seurat software package with the perplexity parameter set to the default, 30, as per ref. 8. The tSNE procedure returns a two-dimensional embedding of single cells, with cells that have similar expression signatures of genes within our variable set located near each other in the embedding. To identify cell types, a density clustering we used a density clustering approach implemented in the DBSCAN R package[60], setting the reachability distance parameter (eps) to 3.62. Clusters with fewer than ten cells and those containing expression markers for more than one canonical cell type (for example, neuron, oligodendrocyte, tanycyte), representing cell doublets (two cells in a single droplet), were removed. As a result of these steps, we were able to assign 20,921 cells (98.5% of our data) into 20 cell type clusters.

Average gene expression for each cluster was determined and Euclidean distances between all pairs was calculated, using this data as input for complete-linkage hierarchical clustering and dendrogram assembly. To identify marker genes, we compared each of the clusters using pairwise differential expression analysis using Bioconductor package edgeR with settings recommended for data with batch effect as described in section 4.2 of the manual. For heat maps, cluster markers were considered if average FC expression compared to other included clusters was >2.

For subclustering, each cluster was iteratively clustered as above until *post hoc* testing of subclusters showed <10 genes showing average expression difference greater than 1 natural log value between clusters with a Bonferroni corrected $P < 0.01$), thresholds based on those of ref. 8. For neuron subclustering, all cells from the six neuron groups from the initial clustering (clusters a13-a18) were first combined and then iteratively subclustered as above. Perplexity was set to the default (30) for all subclustering.

Statistical analysis. For differential expression between cell type clusters as well as for fasting versus fed comparison, we performed pairwise differential expression analysis using Bioconductor package edgeR with settings recommended for data with batch effect as described in section 4.2 of the manual. For 10% diet versus HFD comparison, edgeR was run with standard settings as these data were from a single batch. For heat maps, clusters showing at least one significant change (FDR < 0.25) in both directions were included. As edgeR does not account for gene dropout rate, we used the SCDE package, which adopts a Bayesian approach fitting individual error models to account for stochastic detection of low-expressing genes[61]. We performed SCDE on a subset of our data to compare detection of differentially expressed genes marking clusters (pair-wise comparisons of AgRP (n12 and n13) and POMC (n14, n15 and n21) subclusters, for which there are well-validated cluster markers). To reduce complexity to a manageable computation time, we used setting to model error using the nearest 100 cells (knn.error.models()) and modified batch correction methodology from Fisher's exact test to a chi-squared test to account for large number of cells.

Our results showed that fold changes are highly correlated across the statistical methods and that there is large overlap in significantly differentially expressed genes. When averaging across the ten comparisons, edgeR calls 69 ± 6% (mean ± s.d.) of the DEgenes that SCDE calls, and SCDE calls 49 ± 10% of the DE genes that edgeR does (mean ± s.d.) (FDR < 0.25 for SCDE and FDR < 0.001 for edgeR). For our benchmark set of marker genes, SCDE showed 92% sensitivity while edgeR performed at 100% sensitivity. We conclude from these results that edgeR performs adequately for Drop-seq differential expression analysis and recommend using a threshold of FDR <0.001 from the edgeR results we provide in **Supplementary Tables 2–8** for determination of robust differential expression that are likely to meet statistical significance using the more stringent SCDE.

Food intake was analyzed with two-way ANOVA followed by Sidak's multiple comparisons test.

No statistical methods were used to predetermine sample sizes, but sample size in the feeding study is comparable to previous studies[4]. Data distribution was assumed to be normal but this was not formally tested. Other than the random assignment of mice to experimental groups, no randomization was performed. Data collection and analysis were not performed blind to the conditions of the experiments.

A **Supplementary Methods Checklist** is available.

Sex-of-origin prediction for cells from mixed-sex samples. For each of the batch 5 experimental conditions, cells were pooled before Drop-seq. To predict whether each cell came from female or male, we used the MLSeq package with random forest algorithm to determine a model for predicting sex of origin. Using 310 genes showing differential expression (edgeR, with FDR < 0.25, FC > 0.5) between sexes in both fasted and fed conditions in batch 6 (where sex is known), we created a model on 80% of batch 6 cells (randomly selected) using standard recommended options. The resulting model on batches 1, 2, 3, 4 and the remaining 20% of the batch 6 showed 90% accuracy, with equal rates of misassignment for male and female cells.

Gene ontology analyses. Gene set enrichment analysis was carried out using Bioconductor package RDAVIDWebService (https://www.bioconductor.org/packages/devel/bioc/html/RDAVIDWebService.html). Hierarchical clustering of fold changes across clusters was performed, and all nodes including >10% of genes and <50% of genes were tested for enrichment in DAVID. Non-overlapping nodes showing the highest significance for a gene set were selected. The background set was defined as genes having at least one UMI across all samples[62]. Top scoring gene sets were chosen to include in figures.

Single-cell RNA-seq. Mice were anesthetized with isoflurane and then decapitated. Brains were quickly extracted and immediately chilled in ice-cold, carbogen-saturated (95% O_2, 5% CO_2) high sucrose solution (238 mM sucrose, 26 mM $NaHCO_3$, 2.5 mM KCl, 1.0 mM NaH_2PO_4, 5.0 mM $MgCl_2$, 10.0 mM $CaCl_2$, 11 mM glucose). Chilled brains were sectioned coronally at 300-µm intervals on a Leica VT1000S Vibratome (Leica, Wetzlar, Germany) and incubated in oxygenated artificial cerebrospinal fluid (aCSF; 126 mM NaCl, 21.4 mM $NaHCO_3$, 2.5 mM KCl, 1.2 mM NaH_2PO_4, 1.2 mM $MgCl_2$, 2.4 mM $CaCl_2$, 10 mM glucose) at 34 °C for 30 min. The Arc-ME was visualized by fluorescence stereoscope, microdissected and then enzymatically dissociated as described above, except that cells were resuspended in medium solution[56] instead of Drop-seq buffer. The resulting cell suspension was plated in a 35-mm plastic culture dish, and GFP-positive cells were individually collected and washed by glass micropipette according to a published protocol[63] except that medium solution was used in place of artificial cerebrospinal fluid. Each cell was transferred to a separate qPCR tube cap containing a 2-µL drop of PBS with 10% recombinant RNase inhibitor (Clontech). Droplets were confirmed by microscope to contain a single cell and excess buffer was aspirated away. To control for mRNA contamination during cell collection, an equivalent volume of cell-picking buffer was collected and processed along with the single-cell samples.

Cells and the buffer control sample were frozen at −80 °C until processed for cDNA synthesis and amplification[64]. After 20 cycles of amplification by PCR, cDNA was purified with AMPure XP (1:1 ratio; Beckman-Coulter, Brea, CA), eluted in 30 µL PCR-grade water and then analyzed by quantitative PCR (qPCR) for expression of a housekeeping gene, *Actb* or *Rplp0* (catalog no. Mm.PT.58.28904620.g and Mm.PT.58.43894205, respectively; Integrated DNA Technologies (IDT), Coralville, IA, USA). As quality control we excluded from sequencing any sample with a Ct value greater than 1 s.d. from the mean for their batch, and we excluded any sample batches showing contamination (defined as qPCR detection of housekeeping gene expression in the cell-picking buffer control sample). Samples passing quality control were diluted to 150 pg/µL and input to the Nextera XT library preparation kit (96 indices, Illumina), following the manufacturer's instructions except at one-quarter reaction volume. Library concentration was analyzed by fluorometry (Qubit 2.0; high-sensitivity dsDNA kit; ThermoFisher) and then pooled for equimolar representation. The pool concentration was measured by KAPA Library Quant qPCR according to the manufacturer's instructions (KAPA Biosciences, Wilmington, MA). The qPCR product and the 20-nM standard were analyzed for size (base pairs) by ethidium bromide agarose gel electrophoresis. The size of the 20-nM standard was divided by that of the library pool, and this value was multiplied by the library pool concentration to get a final, size-corrected concentration. The library pool was then diluted to 2 nM for sequencing according to manufacturer's instructions.

Viral injections. Stereotaxic injections were performed as previously described[7]. Briefly, mice were anesthetized with xylazine (5 mg per kg) and ketamine (75 mg per kg) diluted in saline and placed into a stereotaxic apparatus (model 963, David Kopf Instruments, Tujunga, CA). For postoperative care, mice were injected intraperitoneally with meloxicam (0.5 mg per kg). After exposing the skull via a small incision, a small hole was drilled for injection. A pulled-glass pipette was inserted into the brain and virus was injected by an air pressure system. A micromanipulator (model S48 stimulator, Grass Technologies, Rockland, MA) was used deliver the injection at 5 nL/min and the pipette was withdrawn 5 min after injection. For electrophysiology and tracing, AAV1-CBA-Flex-ChR2(H134R)-mCherry (University of Pennsylvania School of Medicine, Philadelphia) was injected unilaterally into the Arc (2–5 nL; from bregma, AP: −1.35 mm, DV: −6.00 mm, ML: ± 0.2 mm). For *in vivo* chemogenetic experiments, AAV8-hSyn-DIO-hM3Dq-mCherry (University of North Carolina Vector Core, Chapel Hill, NC) was bilaterally injected into the arcuate (2–5 nL, coordinates as above). For profiling of Arc-ME RIP-Cre neurons, AAV8-EF1a-DIO-eYFP (University of North Carolina Vector Core, Chapel Hill, NC) was injected into arcuate (100 nL, coordinates as above).

Evan's blue dye injection. After anesthetizing the mouse with 7% chloral hydrate (350 mg/kg), we injected 100 µL of Evan's blue dye solution (1% wt/vol in sterile saline) into the tail vein 20 min before sacrifice, as per ref. 65. Brains were quickly extracted, blocked into a 4-mm coronal section containing the entire arcuate hypothalamus (0 mm to −4 mm from bregma), and submerged in 10% neutral buffered formalin (PFA) for 1 week at room temperature, protected from light. Brain blocks were transferred to 20% sucrose overnight at room temperature and then cut by freezing microtome into 35-µm-thick coronal sections for immunohistochemistry.

***In situ* hybridization.** To generate riboprobe templates, cDNA from mouse hypothalamus was used to PCR amplify the mouse somatostatin (*Sst*) fragment corresponding to bases 89–488 (GenBank accession code NM_009215.1) with the T7 RNA polymerase recognition site incorporated into the product for subsequent transcription. Standard *in vitro* transcription methods using T7 polymerase

doi:10.1038/nn.4495

(Promega, Madison, WI) and digoxigenin-UTP RNA labeling mix (Roche, Basel, Switzerland) were used to produce both sense and antisense riboprobes for *Sst* mRNA. Prior to hybridization riboprobes were diluted in hybridization cocktail (Amresco, Solon, OH) with tRNA.

Brain tissue was cryosectioned coronally at 12 µm, thaw-mounted onto electrostatically clean slides and stored at −80 °C. Prior to hybridization, sections were postfixed in 4% paraformaldehyde, rinsed in 0.1 M PBS (pH 7.4), equilibrated in 0.1 M triethanolamine (pH 8.0), acetylated in triethanolamine containing 0.25% acetic anhydride and hybridized overnight at 56 °C with *Sst* riboprobe. The next day slides were treated with RNase A and stringently washed in 0.5× SSC at 60 °C for 30 min. Following washing, slides were incubated with an antibody against digoxigenin conjugated to horseradish peroxidase (Roche) overnight at 4 °C. Riboprobe signal was further enhanced using the TSA-plus biotin system (PerkinElmer) and streptavidin Alexa Fluor 647 conjugate (Invitrogen, Carlsbad, CA) was used for fluorescence visualization of *Sst* expression. Fluorescence image capture and analysis was done using an VS120 Slide Scanner (Olympus, Tokyo, Japan) and OlyVIA image analysis software.

Leptin treatment. Leptin was administered to induce STAT3 phosphorylation as previously described[6]. Briefly, mice were fasted overnight, injected the following morning with 4 mg/kg recombinant leptin (intraperitoneally, i.p.; A.F. Parlow, National Hormone and Peptide Program, NHPP, Harbor-UCLA Medical Center, Torrance, CA) and then sacrificed for immunohistochemistry 1 h later.

Immunohistochemistry. Mice were terminally anesthetized with 7% chloral hydrate (350 mg/kg) diluted in isotonic saline and transcardially perfused with phosphate-buffered saline (PBS) followed by 10% PFA. Brains were removed, stored in the same fixative for 2 h, transferred into 20% sucrose at 4 °C overnight and cut into 40-µm coronal sections on a freezing microtome. Brain sections were washed twice in 0.1 M phosphate-buffered saline (PBS) with Tween-20, pH 7.4 (PBST), blocked in 3% normal donkey serum and 0.25% Triton X-100 in PBS for 1 h at room temperature, and incubated overnight at room temperature in blocking solution containing primary antisera: rabbit anti-TSHb, 1:1,000 (catalog no. AFP-1274789, NHPP); goat anti-AgRP, 1:1,000 (catalog no. GT15023, Neuromics, Minneapolis, MN); rabbit anti-dsRed (mCherry), 1:1,000 (catalog no. 632496, Clontech, Mountain View, CA); chicken anti-GFP, 1:1,000 (catalog no. A10262, Invitrogen, Carlsbad, CA); rabbit anti-SPRR1A[66], 1:500, a kind gift from S. Strittmatter (Yale University, New Haven, CT); 1:100 rabbit anti-phospho-STAT3 (Tyr705) (D3A7) XP (catalog no. 9145S, Cell Signaling, Danvers, MA). Validation data for each antibody are available from its manufacturer. The morning after incubation in primary antibody, sections were washed three times in PBS and then incubated in one of the following donkey-derived secondary antisera for 2 h in the dark at room temperature: Alexa Fluor 488–conjugated donkey anti-chicken (diluted 1:1,000; catalog no. 703-545-155, Jackson ImmunoResearch, West Grove, PA, USA); Alexa Fluor 488–conjugated donkey anti-goat (diluted 1:1,000; catalog no. A-11055, ThermoFisher Scientific); Alexa Fluor 488–conjugated donkey anti-rabbit (diluted 1:1,000; catalog no. A-21206, ThermoFisher Scientific); Alexa Fluor 594–conjugated donkey anti-rabbit (for Sst-IRES-Cre experiments, diluted 1:1,000; catalog no. A-21207, ThermoFisher Scientific); Alexa Fluor 647–conjugated donkey anti-rabbit (for pSTAT3 experiments, diluted 1:100; catalog no. A-31573, ThermoFisher Scientific). Sections were washed three times in PBS, mounted onto gelatin-coated slides (Southern Biotech; Birmingham, AL), coverslipped with Vectashield Anti-fade Mounting Medium with DAPI (Vector Labs, Burlingame, CA) and sealed with nail polish. Fluorescence images were captured with Olympus VS120 slide scanner microscope and with a confocal microscope (Zeiss LSM510 Upright Confocal System).

Electrophysiology. Animals were deeply anesthetized with isoflurane, decapitated and brains quickly removed into ice-cold cutting solution consisting of (in mM) 93 NMDG, 2.5 KCl, 1.2 NaH_2PO_4, 30 $NaHCO_3$, 20 HEPES, 25 glucose, 5 sodium ascorbate, 2 thiourea, 3 sodium pyruvate, 10 $MgSO_4$, 0.5 $CaCl_2$ (pH 7.3 adjusted with HCl), oxygenated with 95% O_2, 5% CO_2, measured osmolarity 310–320 mOsm per L. Then 300-µm-thick coronal sections were cut with a Leica VT1000S vibratome and incubated in oxygenated cutting solution at 34 °C for 10 min. Next, slices were transferred to oxygenated aCSF consisting of (in mM) 126NaCl, 21.4 $NaHCO_3$, 2.5 KCl, 1.2 NaH_2PO_4, 1.2 $MgCl_2$, 2.4 $CaCl_2$, 10 glucose and recovered for 30 min at 34 °C. Slices were stored in oxygenated aCSF at room temperature (20–24 °C) for at least 60 min before recording. A single slice was placed in the recording chamber where it was continuously superfused at a rate of 3–4 mL per min with oxygenated aCSF. Neurons were visualized with a SliceScope microscope (Scientifica, Uckfield, UK) equipped with infrared differential interference contrast and fluorescence optics.

Recordings were obtained using borosilicate glass microelectrodes (5–7 MΩ) filled with a Cs^+-based low-Cl^- internal solution containing (in mM) 135 $CsMeSO_3$, 10 HEPES, 1 EGTA, 4 $MgCl_2$, 4 Na_2-ATP, 0.4 Na_2-GTP, 10 Na_2-phosphocreatine (pH 7.3 adjusted with CsOH; 295 mOsm per kg; $E_{Cl} = -70$mV) from unidentified neurons in the PVH. Photostimulation-evoked IPSCs were recorded in whole-cell voltage-clamp mode, with membrane potential clamped and 0 mV in presence of CNQX (10 µM) and D-AP5 (50 µM). To confirm whether postsynaptic currents were mediated by $GABA_A$ receptors, bicuculline was applied to the bath (10 µM) during recordings.

Recordings were made using a Multiclamp 700B amplifier, and data were filtered at 2 kHz and digitized at 10 kHz. To photostimulate channelrhodopsin2-positive fibers, a LED light source (473 nm; CoolLED, Andover, UK) was used, as previously described[7]. The light output was controlled by a programmable pulse stimulator, Master-8 (A.M.P.I., Jerusalem, Israel) and pClamp 10.2 software (Axon Instruments, Foster City, CA). The photostimulation-evoked IPSCs detection protocol consisted of four blue light pulses administered 1 s apart during the first 4 s of an 8-s sweep, repeated for a total of 30 sweeps. Evoked IPSCs with short latency (≤6 ms) upon light stimulation were considered as light-driven.

Food intake studies. All mice were singly housed for at least 2.5 weeks following surgery and handled for 10 consecutive days before the assay to reduce stress response. Feeding studies were performed in home cages with *ad libitum* food access. The day before the experiment, mice were provided with fresh cages to avoid leftover food spilling in the bedding.

Mice were intraperitoneally administered clozapine-*N*-oxide (CNO; diluted in isotonic saline) at 1 mg per kg of body weight, or an equivalent volume of isotonic saline. For light cycle measurements, animals were injected with either CNO or saline at 9:00 a.m. and food was weighed 1 h, 2 h, 3 h and 4 h later. A full trial consisted of assessing food intake from the study subjects after they received injections of saline on day 1 and CNO on day 2. For dark cycle measurements, animals were injected with either saline or CNO at 6:00 p.m. and food was weighed 1 h, 2 h, 3 h and 4 h later. A full trial consisted of assessing food intake from the study subjects after they received injections of saline on day 1 and CNO on day 2. Following feeding studies, mice were sacrificed for immunohistochemical assessment of AAV-Syn1-DIO-hM3Dq-mCherry expression in the Arc. Mice in which mCherry was expressed in cell bodies outside the Arc were excluded from analysis of behavioral data.

Classification of oligodendrocyte clusters and RIP-Cre neurons. To compare oligodendrocyte clusters to maturational stages, we generated a heat map showing single-cell expression of oligodendrocyte stage-specific marker genes identified by ref. 9. We then assigned each oligodendrocyte cluster to a maturational stage on the basis of how well it visually matched the stage-specific marker profile. We used a similar process to assign RIP-Cre neurons to neuron clusters. Specifically, we generated a heat map showing each RIP-Cre neuron's expression of neuron cluster marker genes and then visually matched each RIP-Cre neuron to a cluster.

DEPICT analysis. DEPICT version 1, release 194 (https://github.com/perslab/depict/) was used to prioritize Arc-ME cell clusters on the basis of GWAS summary statistics for body mass index[45], height[50], waist-to-hip ratio adjusted for body mass index[46], type 2 diabetes[67], menopause[68], menarche[69] and anorexia[70] (**Supplementary Table 8**). For body mass index, waist-to-hip ratio and height, we applied the same association *P*-value cutoffs as used in refs. 45,46,50. For the traits with no published DEPICT analyses (menarche, menopause and anorexia), we used a GWAS association *P* value cutoff of $<10^{-5}$. Default DEPICT settings were used to construct loci: (a) input SNPs were clumped using PLINK version 1.90 (ref. 71), enforcing a clumping distance of 500 kb and/or linkage disequilibrium r^2 threshold of <0.1, (b) loci were formed around the index SNP of each clump using $r^2 > 0.5$ as the locus definition, (c) genes were mapped to loci, and (d) loci containing the same gene(s) were merged into single loci (for details, see ref. 61). 1000 Genomes Project phase 1 genotype data[72] were used to perform clumping and compute locus boundaries using human genome build 19. Association

signals on sex chromosomes were omitted from the analyses. Normalized Arc-ME single-cell expression data were averaged across Seurat clusters, mapped from mouse gene symbols to mouse Ensembl[73] gene identifiers (using Ensembl version 83) and to the human Ensembl gene identifiers (using Ensembl version 82). Mouse gene identifiers mapping to several mouse Ensembl identifiers were discarded and the human gene with the highest degree of mouse homology was retained in instances where a mouse gene mapped to several human genes. Expression levels of mouse genes mapped to the same human gene were averaged. The resulting cell cluster Arc-ME gene expression levels were transformed to standard normal distributions before the DEPICT analyses. DEPICT was run using default settings.

Code availability. Code is available at https://github.com/perslab/campbell-2017.

Data availability. The data that support the findings of this study are available from the corresponding author upon reasonable request. Raw and fully processed single-cell RNA-seq data on RIP-Cre neurons are available at GEO accession code GSE90806. Raw Drop-seq data and processed DGE files are available at GEO accession code GSE93374.

51. Balthasar, N. *et al.* Leptin receptor signaling in POMC neurons is required for normal body weight homeostasis. *Neuron* **42**, 983–991 (2004).
52. Taniguchi, H. *et al.* A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. *Neuron* **71**, 995–1013 (2011).
53. Krashes, M.J. *et al.* An excitatory paraventricular nucleus to AgRP neuron circuit that drives hunger. *Nature* **507**, 238–242 (2014).
54. van den Pol, A.N. *et al.* Neuromedin B and gastrin-releasing peptide excite arcuate nucleus neuropeptide Y neurons in a novel transgenic mouse expressing strong *Renilla* green fluorescent protein in NPY neurons. *J. Neurosci.* **29**, 4622–4639 (2009).
55. Postic, C. *et al.* Dual roles for glucokinase in glucose homeostasis as determined by liver and pancreatic beta cell-specific gene knock-outs using Cre recombinase. *J. Biol. Chem.* **274**, 305–315 (1999).
56. Saxena, A. *et al.* Trehalose-enhanced isolation of neuronal sub-types from adult mouse brain. *Biotechniques* **52**, 381–385 (2012).
57. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
58. Satija, R., Farrell, J.A., Gennert, D., Schier, A.F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
59. van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
60. Esther, M., Kriegel, H.-P., Sander, J. & Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. in *2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)* (Institute for Computer Science, University of Munich, 1996).
61. Kharchenko, P.V., Silberstein, L. & Scadden, D.T. Bayesian approach to single-cell differential expression analysis. *Nat. Methods* **11**, 740–742 (2014).
62. Fresno, C. & Fernández, E.A. RDAVIDWebService: a versatile R interface to DAVID. *Bioinformatics* **29**, 2810–2811 (2013).
63. Hempel, C.M., Sugino, K. & Nelson, S.B. A manual method for the purification of fluorescently labeled neurons from the mammalian brain. *Nat. Protoc.* **2**, 2924–2929 (2007).
64. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
65. Mullier, A., Bouret, S.G., Prevot, V. & Dehouck, B. Differential distribution of tight junction proteins suggests a role for tanycytes in blood-hypothalamus barrier regulation in the adult mouse brain. *J. Comp. Neurol.* **518**, 943–962 (2010).
66. Bonilla, I.E., Tanabe, K. & Strittmatter, S.M. Small proline-rich repeat protein 1A is expressed by axotomized neurons and promotes axonal outgrowth. *J. Neurosci.* **22**, 1303–1315 (2002).
67. Morris, A.P. *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* **44**, 981–990 (2012).
68. Day, F.R. *et al.* Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair. *Nat. Genet.* **47**, 1294–1303 (2015).
69. Perry, J.R. *et al.* Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature* **514**, 92–97 (2014).
70. Boraska, V. *et al.* A genome-wide association study of anorexia nervosa. *Mol. Psychiatry* **19**, 1085–1094 (2014).
71. Chang, C.C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
72. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
73. Yates, A. *et al.* Ensembl 2016. *Nucleic Acids Res.* **44** D1: D710–D716 (2016).