MOLECULAR
METABOLISM

Heterogeneity of hypothalamic pro-opiomelanocortin-expressing neurons revealed by single-cell RNA sequencing

CrossMark

Brian Y.H. Lam [1], Irene Cimino [1], Joseph Polex-Wolf [1], Sara Nicole Kohnke [1], Debra Rimmington [1], Valentine Iyemere [1], Nicholas Heeley [1], Chiara Cossetti [2], Reiner Schulte [2], Luis R. Saraiva [3], Darren W. Logan [3], Clemence Blouet [1], Stephen O'Rahilly [1], Anthony P. Coll [1], Giles S.H. Yeo [1],*

ABSTRACT

Objective: Arcuate proopiomelanocortin (POMC) neurons are critical nodes in the control of body weight. Often characterized simply as direct targets for leptin, recent data suggest a more complex architecture.
Methods: Using single cell RNA sequencing, we have generated an atlas of gene expression in murine POMC neurons.
Results: Of 163 neurons, 118 expressed high levels of *Pomc* with little/no *Agrp* expression and were considered "canonical" POMC neurons (P^+). The other 45/163 expressed low levels of *Pomc* and high levels of *Agrp* (A^+P_+). Unbiased clustering analysis of P^+ neurons revealed four different classes, each with distinct cell surface receptor gene expression profiles. Further, only 12% (14/118) of P^+ neurons expressed the leptin receptor (*Lepr*) compared with 58% (26/45) of A^+P_+ neurons. In contrast, the insulin receptor (*Insr*) was expressed at similar frequency on P^+ and A^+P_+ neurons (64% and 55%, respectively).
Conclusion: These data reveal arcuate POMC neurons to be a highly heterogeneous population.
Accession Numbers: GSE92707.

1. INTRODUCTION

The leptin-melanocortin pathway plays a key role in the control of food intake and body weight, with genetic disruption of multiple components of the pathway resulting in severe, early-onset obesity in both humans and mice (reviewed by [1]).

In the brain, POMC expression is localized to neurons in the arcuate nucleus (ARC) of the hypothalamus and in the nucleus of the solitary tract (NTS) [2]. POMC is a pro-peptide that is processed to a range of bioactive peptide products including α and β-MSH [3,4], both of which act on melanocortin receptors 3 and 4, which are widely expressed in the central nervous system. The role of MC4R in mediating the effects of POMC on energy balance is well established, whereas the role of MC3R is less clear [2]. In the ARC, anorexigenic POMC neurons exist alongside orexigenic neurons that express the endogenous melanocortin antagonist agouti-related peptide (AgRP) [5,6]. The canonical view of the melanocortin pathway is one where leptin, which is produced from white adipose tissue to reflect fat mass, signals to leptin receptors on the POMC and AgRP neurons, increasing the activity of

POMC neurons and reducing the activity of AgRP expressing neurons [1,2].

However, there is a body of evidence that indicates this view of the leptin melanocortin pathway to be an over-simplistic representation of the relevant circuitry. Woods and Stock, using FOS immunoreactivity of the immediate early gene *c-fos* as a marker of neuronal activation, reported that while neurons in the arcuate nucleus of leptin deficient *ob/ob* mice were activated by peripheral leptin administration, the area of the hypothalamus most strongly activated was the paraventricular nucleus [7]. Later, it became clear that not all POMC neurons express the receptors for leptin or insulin, with whole-cell patch-clamp electrophysiology indicating that a subpopulation is depolarized by leptin, while a separate and distinct population is hyperpolarized by insulin [8]. POMC neurons also vary in their expression of the neurotransmitters glutamate, γ-aminobutyric acid (GABA) [9,10], and acetylcholine [11] and their downstream projections. For instance, 40% of POMC neurons appear to be GABAergic and hence inhibitory [9]. In addition, only about 40% of POMC neurons express the serotonin 2C receptor ($5HT_{2C}$-R), the target for the weight-loss drug Lorcaserin,

through which it is thought to mediate much of its actions on satiety [12,13]. Finally, mice with a targeted deletion of the leptin receptor specifically from POMC neurons display a surprisingly mild phenotype [14], as compared to complete *Pomc* null mice [15,16] and *db/db* mice homozygous for a *Lepr* mutation [17].

A recent paper by Henry et al. reported the gene expression profiles of pooled AGRP and POMC neurons from fed and food deprived mice and found that the global pattern of gene expression in response to an acute energy deficit are far more dramatic in AGRP expressing than POMC expressing neurons [18]. However, because of its use of pooled cells, the transcriptomic profiles of individual neurons, and hence their variability in gene expression, was not examined.

We now report the results of single cell RNA sequencing in isolated POMC expressing neurons of the arcuate nucleus, which has revealed a previously unappreciated degree of heterogeneity within this anatomically highly localised population.

2. RESULTS

2.1. RNA-sequencing of POMC neurons

Using a POMC-eGFP mouse [19], we successfully sequenced the transcriptome of 163 GFP positive (MBH Pos) and 7 GFP negative (MBH Neg) from hypothalamic tissue encompassing the arcuate nucleus (Figure 1A), and 24 cells, as positive controls, from the pituitary (Pit) known to express high levels of POMC (GEO Acc# GSE92707). When we visualized the transcriptomic data using t-SNE dimension reduction analysis [20], the pituitary and hypothalamic cells

segregated into two discrete clusters, with the hypothalamic neurons further segregating into GFP positive and GFP negative cells (Figure 1B).

All cells sequenced were of similar size, as measured by degree of forward scatter on the cell-sorter and expressed both *Actb* and *Gapdh* (Figure 1C). All the hypothalamic cells expressed the neuronal marker *Ncam*, but, because the pituitary cells contained both neuronal derived posterior and the ectodermally derived anterior, there was a larger heterogeneity of *Ncam* expression in the pituitary cells (Figure 1C). In contrast, the GFP signal, which reflects the degree of *Pomc* expression, was consistently high with very little heterogeneity in the pituitary POMC cells, whereas hypothalamic GFP neurons show a much larger variation in signal. *Cartpt*, which is canonically expressed together with *Pomc* in the same neurons, also showed variable expression in hypothalamic GFP neurons and was not detected in the pituitary (Figure 1C). POMC is processed by the proprotein convertases, PC1 (PCSK1) and PC2 (PCSK2). Using the detection cut-off $\geq$5 linear RPM, most cells (156/163) expressed either or both *Pcsk1* and *Pcsk2* (Figure 1D). On average, 6882 different transcripts were detected from each individual neuron ($\geq$5 linear reads per million or RPM). However, across all 163 POMC neurons, 12,187 different transcripts were expressed (Figure 1E), indicating significant heterogeneity in gene expression repertoire among the neurons.

2.2. AgRP/NPY expressing POMC neurons

Unexpectedly, we found 27% (45/163) of *Pomc*-expressing neurons also expressed high levels of the endogenous melanocortin antagonist
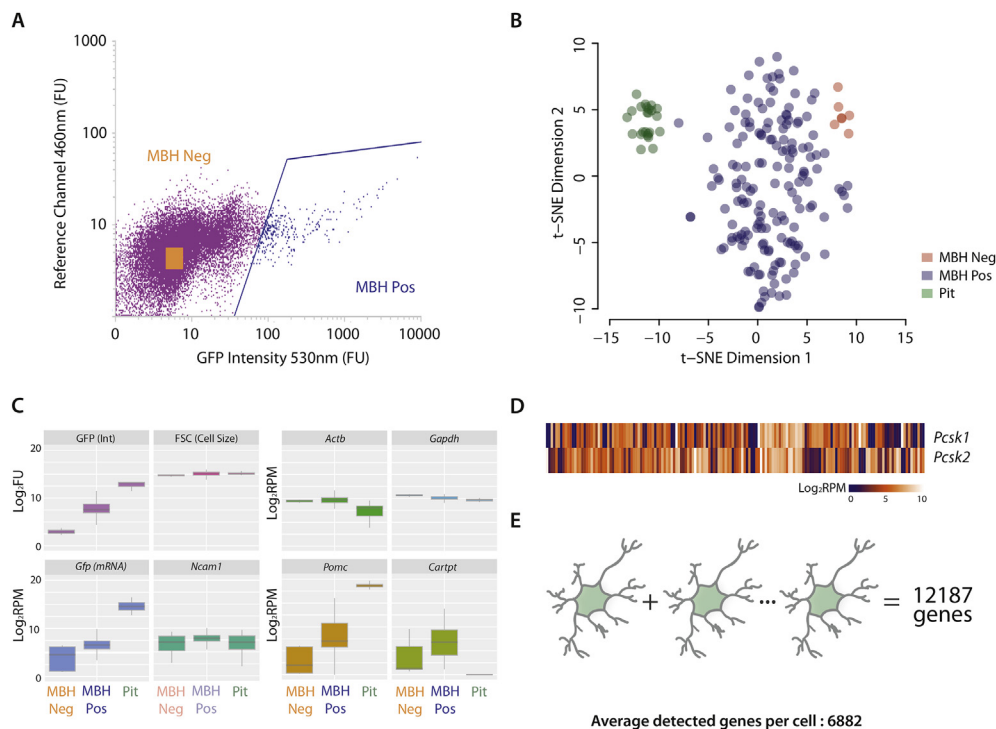
Figure 1: FACS sorting and single-cell RNA sequencing of POMC-eGFP neurons. **(A)** The FACS gating for GFP (530 nm) and reference channel (460 nm). The GFP positive cells from the mediobasal hypothalamus (MBH Pos) are labelled blue, and negative cells (MBH Neg) are collected from the orange square within the negative population (purple). The cells were further gated for nuclear stain (DraQ5), cell size (FSC), and granularity (SSC), data not shown. **(B)** t-SNE was used to show the segregation of cells (MBH Pos and Neg, as well as GFP positive cells collected from the pituitary) based on their transcriptomic profiles. **(C)** Quality checks for FACS showing the GFP intensity and cell size measurement among the 3 cell types. We also examined for RNA expression of *Gfp* mRNA, neuronal marker *Ncam1*, the housekeeping genes *Actb* and *Gapdh*, *Pomc*, and *Cartpt*. **(D)** *Pcsk1* and *Pcsk2* are widely expressed in the cells sequenced. **(E)** We detected an average of 6882 genes per cell ($\geq$5 RPM), adding all the genes from all the 163 cells together resulted in a total of 12,187 detected genes, indicating heterogeneity in gene expression.

384 **MOLECULAR METABOLISM 6 (2017) 383—392**
www.molecularmetabolism.com

Agrp (Figure 2A, C). In addition, these neurons expressed a high level of *Npy*, while *Pomc* expression itself was a magnitude (>5 log$_2$FC) lower than that seen in POMC neurons not expressing or expressing low levels of *Agrp* (Figure 2B). Comparing our data with a published transcriptome of pooled AgRP neurons [18], we found a significant
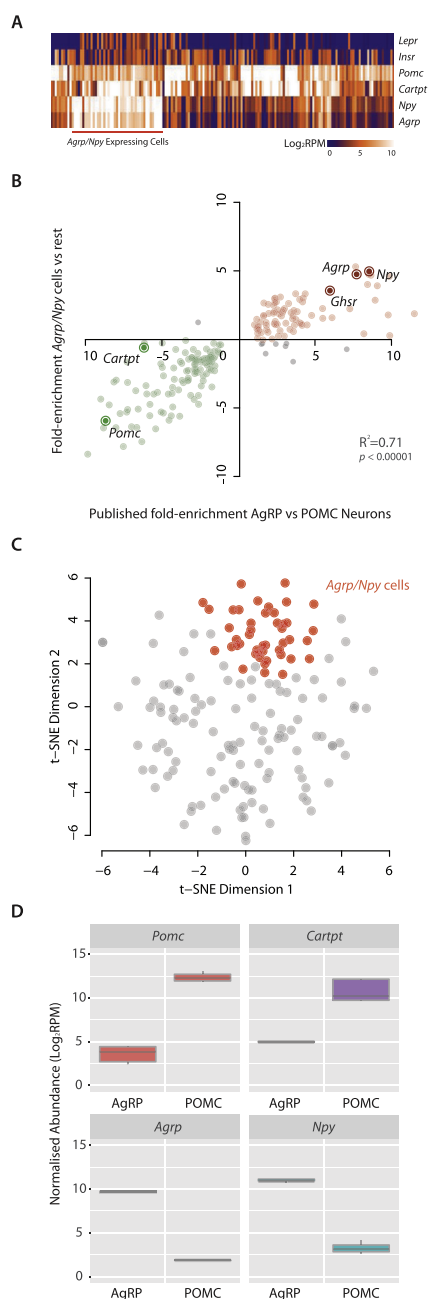

Figure 2: **27.6% of POMC cells also express high levels of *Agrp* and *Npy*. (A)** We detected 45 cells that expressed high levels of *Agrp* and *Npy* (A$^+$P$_+$); interestingly, many of these cells also express *Lepr*. **(B)** Differential expression analysis of the A$^+$P$_+$ vs rest (P$^+$) cells shows similar fold enrichments compared to AgRP vs POMC comparison from [18]. The expression of the top 200 enriched genes (from the published data) are plotted in the graph. **(C)** The *Agrp/Npy* cells form a strong cluster in a t-SNE plot showing only the 163 MBH Pos cells. **(D)** Expression levels of *Pomc*, *Cartpt*, *Agrp*, and *Npy* in published AgRP and POMC neurons. The AgRP neurons express low but detectable levels of *Pomc* and *Cartpt*; on the other hand, *Npy* and *Agrp* are also detectable in POMC neurons.

similarity in gene expression profiles (R^2 = 0.71 $p < 0.00001$), with high expression, for instance, of the ghrelin receptor *Ghsr* (Figure 2B). Analysis of the raw data from Henry et al., 2015 also indicated that *Pomc* and *Cartpt* were expressed in the AgRP neurons that they had analysed [18], albeit at low levels (Figure 2D).

Taken together, these 45 high *Agrp* and low *Pomc* expressing neurons (hereafter referred to as A$^+$P$_+$ neurons) showed a striking similarity to canonical AgRP/NPY neurons, likely reflecting their developmental origins, with up to 25% of AgRP neurons sharing the same hypothalamic progenitors as POMC neurons [21]. We defined the rest of the 118 neurons, which expressed higher levels of *Pomc*, as canonical POMC expressing neurons (P$^+$) and conducted our further studies on this population.

2.3. Expression of hypothalamic neuroendocrine genes in POMC neurons

We produced a matrix showing the expression distribution of 23 well characterized neuroendocrine genes, all known to play a role in the regulation of energy homoeostasis, on the individual POMC (P$^+$) neurons (Figure 3A). We found, as has been previously reported, that a large percentage of POMC neurons expressed *Mc3r*, *Cnr1*, *Htr2c*, and the *Npy-y1*, *y2*, and *y5* receptors [1]. Of note, the percentage of *Htr2c* expressing neurons (34/118; 29%) was comparable to previous studies showing that 5HT$_{2C}$Rs (the protein product of *Htr2c*) are expressed on approximately 40% of POMC neurons [12]. We also found that a smaller percentage of neurons, 12/118, expressed *Mc4r* (co-localization validated as shown in Figure 3B), with 11/118 expressing *Glp1r*. ARC POMC neurons expressing GLP1R are thought to mediate, in part, the weight-loss effects of the GLP1 analogue liraglutide [22].

While we excluded the 45 A$^+$P$_+$ neurons from this matrix, a large number of the remaining P$^+$ neurons still expressed low levels of *Agrp* (66/118) and *Npy* (98/118). Our own data (Figure 3A) and re-analysis of the raw data from previously published work indicated that the converse was also likely to be true, with a percentage 'bona fide' AgRP neurons also expressing low levels of *Pomc* (Figure 2D).

2.4. Unbiased clustering reveals four 'classes' of POMC neurons

In order to determine whether P$^+$ neurons segregate into distinct classes, we first adjusted for batch effects and then performed unbiased clustering of the 118 cells based on the consensus matrix of Euclidean, Pearson, and Spearman correlations — using the 'SC3' R package [23]. We obtained 4 main clusters (Figure 4A), which could also be visualized using t-SNE [20] (Figure 4B). Cluster 1 was the largest with 69 cells, followed by cluster 4 (27) and cluster 2 (15). Cluster 3 had the least number of cells (7). Cluster 2 showed resemblance to cluster 1 (Figure 4A, B).

We identified 2773 candidate genes that were differentially expressed among the 4 clusters, with a false discovery rate (FDR) cut-off of 5%, that drove the segregation of cells into the different groups (Figure 4C). Interestingly, there was over-representation of membrane (530) and secretory genes (191) within the driver gene list (Figure 4C, D). These membrane genes encode ion channels, G-protein coupled receptors (GPCRs) and other transmembrane receptors, and transporters (Figure 4D).

Figure 5A lists the top 25 genes, ranked by differential expression that drove the clustering of the four groups (with the exception of group 2, which was a small group and only has 3 significantly differentially expressed genes). Shown are the expression level for each of the genes and how unique each gene was to each group. Additionally, Table 1 contained genes that are not in the top 25, encoding selected
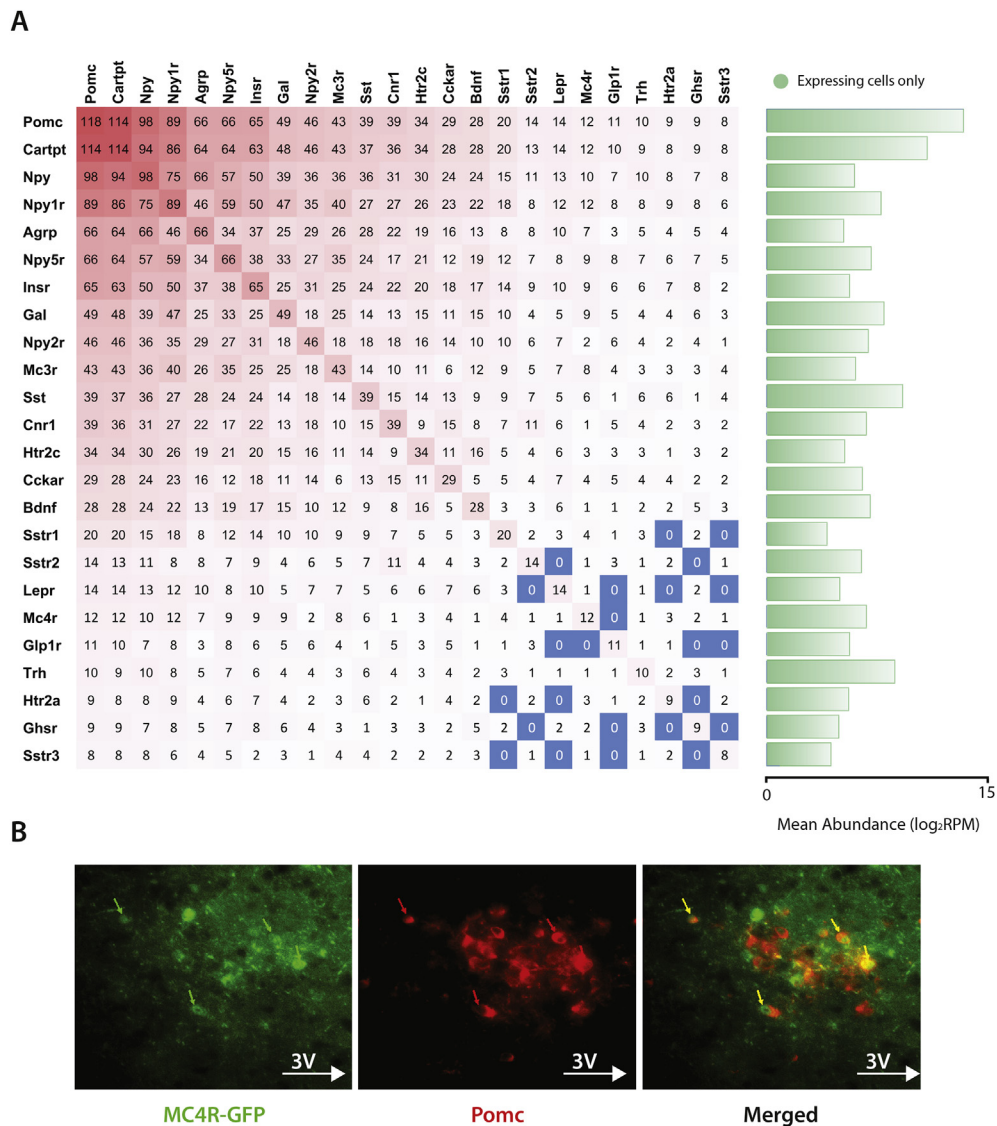
Original Article

A

Expressing cells only

	Pomc	Cartpt	Npy	Npy1r	Agrp	Npy5r	Insr	Gal	Npy2r	Mc3r	Sst	Cnr1	Htr2c	Cckar	Bdnf	Sstr1	Sstr2	Lepr	Mc4r	Glp1r	Trh	Htr2a	Ghsr	Sstr3
Pomc	118	114	98	89	66	66	65	49	46	43	39	39	34	29	28	20	14	14	12	11	10	9	9	8
Cartpt	114	114	94	86	64	64	63	48	46	43	37	36	34	28	28	20	13	14	12	10	9	8	9	8
Npy	98	94	98	75	66	57	50	39	36	36	36	31	30	24	24	15	11	13	10	7	10	8	7	8
Npy1r	89	86	75	89	46	59	50	47	35	40	27	27	26	23	22	18	8	12	12	8	8	9	8	6
Agrp	66	64	66	46	66	34	37	25	29	26	28	22	19	16	13	8	8	10	7	3	5	4	5	4
Npy5r	66	64	57	59	34	66	38	33	27	35	24	17	21	12	19	12	7	8	9	8	7	6	7	5
Insr	65	63	50	50	37	38	65	25	31	25	24	22	20	18	17	14	9	10	9	6	6	7	8	2
Gal	49	48	39	47	25	33	25	49	18	25	14	13	15	11	15	10	4	5	9	5	4	4	6	3
Npy2r	46	46	36	35	29	27	31	18	46	18	18	18	16	14	10	10	6	7	2	6	4	2	4	1
Mc3r	43	43	36	40	26	35	25	25	18	43	14	10	11	6	12	9	5	7	8	4	3	3	3	4
Sst	39	37	36	27	28	24	24	14	18	14	39	15	14	13	9	9	7	5	6	1	6	6	1	4
Cnr1	39	36	31	27	22	17	22	13	18	10	15	39	9	15	8	7	11	6	1	5	4	2	3	2
Htr2c	34	34	30	26	19	21	20	15	16	11	14	9	34	11	16	5	4	6	3	3	3	1	3	2
Cckar	29	28	24	23	16	12	18	11	14	6	13	15	11	29	5	5	4	7	4	5	4	4	2	2
Bdnf	28	28	24	22	13	19	17	15	10	12	9	8	16	5	28	3	3	6	1	1	2	2	5	3
Sstr1	20	20	15	18	8	12	14	10	10	9	7	5	5	3	3	20	2	3	4	1	3	0	2	0
Sstr2	14	13	11	8	8	7	9	4	6	5	7	11	4	4	3	2	14	0	1	3	1	2	0	1
Lepr	14	14	13	12	10	8	10	5	7	5	5	6	6	7	6	3	0	14	1	0	1	0	2	0
Mc4r	12	12	10	12	7	9	9	2	8	6	1	3	4	1	1	1	1	1	12	0	1	3	2	1
Glp1r	11	10	7	8	3	8	6	5	6	4	1	5	3	5	1	1	3	0	0	11	1	1	0	0
Trh	10	9	10	8	5	7	6	4	3	4	3	4	3	1	1	1	1	1			10	2	3	1
Htr2a	9	8	8	9	4	6	7	2	3	6	2	1	4	2	0	0	2	0	3	1	2	9	0	2
Ghsr	9	9	7	8	5	7	8	6	4	3	1	3	3	2	5	2	0	2	2	0	3	0	9	0
Sstr3	8	8	8	6	4	5	2	3	1	4	4	2	2	2	3	0	1	0	1	0	1	2	0	8

0   15

Mean Abundance (log₂RPM)

B

MC4R-GFP   Pomc   Merged

Figure 3: The expression matrix of selected neuroendocrine genes. (A) An expression matrix showing the number of cells expressing ≥ 5 linear RPM (2.32 log₂RPM) of a selection of neuroendocrine genes, the *Agrp/Npy* cells were excluded from the matrix. One can look up to co-expression of up to 2 genes by looking at the intersections between rows and columns, e.g. there are 43 *Mc3r* expressing cells (9th row), and 25 of them also expressed *Insr* (7th column). The average normalized gene abundance in expressing cells is shown in green bars. **(B)** We found 12 cells expressing *Mc4r* within our cohort and this was confirmed by immunostaining for POMC the arcuate nucleus of MC4R-GFP mice.

extracellular secreted hormones, GPCRs and ion channels, indicating which groups had the highest expression of any given gene. For example, while the neurons in cluster 4 expressed the highest levels of *Pomc* and *Cartpt*, both genes, as expected, were also expressed in Clusters 1—3. Cluster 4 neurons, however, were characterized by expression of genes encoding transthyretin (*Ttr*), the cholesystekinin A receptor (*Cckar*), neuromedin U receptor 2 (*Nmur2*), and nascent helix loop helix 2 (*Nhlh2*) (Figure 5A), as well as the GPCRs corticosterone releasing hormone receptor 1 (*Crhr1*), hypocretin receptor 1 (*Hcrtr1*), neuropeptide Y 2-receptor (*Npy2r*), and prolactin releasing hormone receptor (*Prlhr*) (Table 1) at far higher levels than in the other clusters. Cluster 1, in contrast, was characterized by high expression of the serotonin-2a receptor (*Htr2a*), the melanocortin-3 receptor (*Mc3r*) and a large number of ion channels (Table 1).

Thus, no single given 'driver' gene is responsible for these clusters, but rather the combined expression of a repertoire of genes drove the segregation. The entire repertoires of genes for each cluster are listed in Supplementary Table 1.

These clusters are, of course, not fixed. For example, Figure 5B shows the t-SNE plot with the cluster 4 neurons highlighted and dividing clearly into two groups; one larger group of 22 and a smaller one of 5 neurons. Comparing the transcriptomes of the two groups, we could identify differentially expressed genes that defined the two groups (Figure 5C, Supplementary Figure). The four clusters, as we report here, were achieved by using a specific statistical threshold. If we raised this threshold, and therefore statistical rigour, then all clusters, could be further segregated and, if taken to its extreme, could segregate into the original 163 individual neurons sequenced.

2.5. Glutamatergic and GABAergic POMC neurons

Previous studies have shown that POMC neurons can also be classified by the release of stimulatory neurotransmitter glutamate or inhibitory
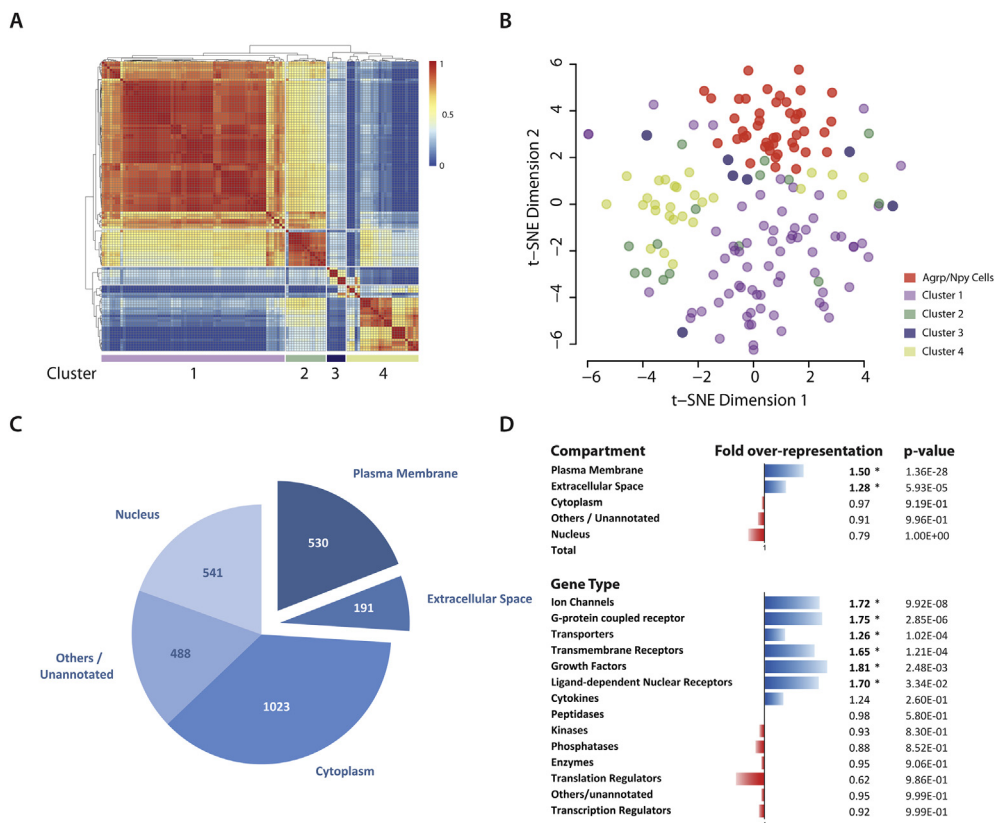
Figure 4: Clustering analysis and identification of driver genes for heterogeneity. **(A)** Clustering analysis was performed using the 'SC3' package in order to identify clusters of cells that show similarity in their transcriptomic profiles. Using k-means clustering, we found 4 clusters with an average silhouette width of 0.64. **(B)** A t-SNE plot showing the 4 clusters from the analysis. **(C)** Using the clustering results we identified 2773 candidate genes that drove heterogeneity. **(D)** There was over-representation of genes from the plasma membrane (1.50-fold) and extracellular space (1.28-fold) within the list of driver gene candidates. This enrichment was mainly driven by over-representation of plasma membrane protein encoding genes including ion channels, G-protein coupled receptors, transporters, and other transmembrane receptors. There was also a 1.8-fold enrichment of growth factors in the driver gene list.

gamma-aminobutyric acid (GABA) [9,10,24]. We examined the two subtypes by looking at expression of vesicular glutamate transporter type 2 (*Vglut2/Slc17a6*) as a proxy for glutamatergic POMC neurons and glutamate decarboxylase (*Gad67/Gad1*) for their GABAergic counterparts. In Figure 5D, we found that the majority of cells (87%) express high levels of *Gad1* (indicated in blue), and these included most of the A$^+$P$_+$ neurons as well as 70% of the P$^+$ neurons. On the other hand, only about half of the P$^+$ neurons have high levels of *Slc17a6* (indicated in red), and these tended to fall into clusters 1 and 4 (Figures 4B and 5D). 24% of P$^+$ neurons express both markers (indicated in purple) indicating the possibility of dual roles at different synaptic junctions.

2.6. Leptin receptor and insulin receptor expressing POMC neurons

Next we examined the two most well studied types of POMC neurons, those expressing the leptin (*Lepr*) and insulin receptors (*Insr*). Looking at all 163 neurons, our sequencing data showed that 40/163 (25%) neurons expressed the leptin receptor and 94/163 (57%) expressed the insulin receptor. 11 (7%) expressed only *Lepr* and *not Insr*, 65 (40%) expressed only *Insr* and *not Lepr*, while 29 (18%) expressed both receptors. 58 (36%) did not express either of the receptors (Figure 6A).

When we split the 163 cells into the high *Agrp* expressing A$^+$P$_+$ neurons and high *Pomc* expressing P$^+$ neurons and looked at the distribution of *Lepr* in these two groups separately, we saw that 26/45

(58%) of A$^+$P$_+$ neurons expressed *Lepr* as compared to only 14/118 (12%) of P$^+$ neurons, almost five-fold more. In contrast, the *Insr* is expressed in 29/45 (64%) of A$^+$P$_+$ neurons, which is comparable to its expression in 65/118 (55%) of P$^+$ neurons.

To determine if we had introduced any selection bias during the flow sorting, we sought to define the number of *Lepr* and *Insr* expressing POMC neurons in the intact mouse hypothalamus. Attempts to use immunocytochemistry to detect either the leptin or the insulin receptor on whole brain slices proved to be technically unreliable (data not shown). We therefore measured levels of Stat3 phosphorylation (pStat3) after leptin administration and levels of Akt phosphorylation (pAkt) after insulin administration using immunohistochemistry, as a proxy for expression of the respective receptors. We performed these experiments on POMC-GFP mice, allowing us to co-localize pStat3 or pAkt immunofluorescence in neurons expressing GFP (Figure 6B, C). After leptin administration, pStat3 positive cells were largely concentrated to the base of the arcuate nucleus, surrounding the median eminence (Figure 6B). Cell counting across multiple brain slices from three mice indicated that 23.8% of POMC-GFP neurons were positive for pStat3 in response to leptin, which was consistent with our RNAseq data (Figure 6C). While the leptin responsiveness in POMC neurons was previously shown to be variable, dependent on the treatment protocols [25], the percentage we show here falls within the lower limit of what others have previously reported [26]. Insulin administration, however, resulted in a different pattern of activation, with pAkt
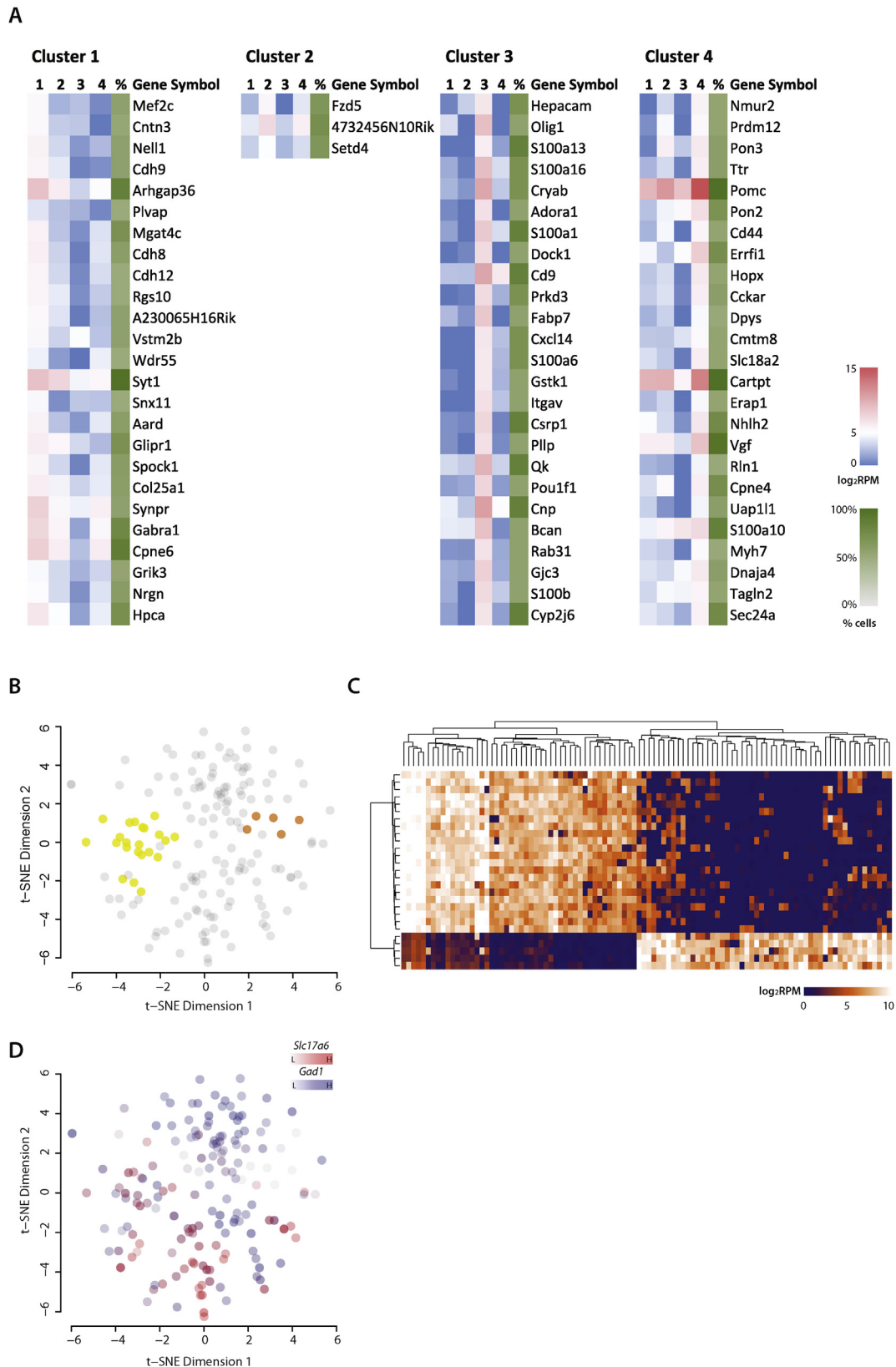
Figure 5: Top candidate driver genes for the resulting clusters. **(A)** The top 25 genes 'markers' were ranked by maximum fold difference and a differential expression FDR cut off of 0.05, and the lists for each of the clusters are shown. The number of cells expressing the gene is also included in the rightmost column (green). *Pomc* and *Cartpt* are expressed in all 4 clusters, but the expression of the two genes are highest in Cluster 4 cells (>1log$_2$FC from cluster 2 which is second highest). **(B)** The clusters can be further subdivided, e.g. cluster 4 shown here. **(C)** Top differentially expressed genes between the two subgroups of cluster 4 cells. A larger version of the heatmap is available in the supplementary material. **(D)** All of the A$^+$P$_+$ neurons express GABAergic marker *Gad1* (*Gad67*, blue) and the same applies to 70% of the P$^+$ neurons. We also found about 50% of the P$^+$ neurons that express glutamatergic marker *Slc17a6* (*Vglut2*, red), 24% of the P$^+$ neurons express both markers indicating dual roles at synaptic junctions.

Table 1 — A selective list of driver genes for the P$^+$ cell clusters.

Gene Symbol	Location	Gene Type	Enrichment in Expression (log$_2$FC >=1 FDR <0.05)				% Cells with Expression >= 5 Linear RPM			
			Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Gal	Extracellular Space	Hormone				✓	43.5%	13.3%	42.9%	51.9%
Pnoc	Extracellular Space	Hormone	✓				46.4%	26.7%	14.3%	33.3%
Prok2	Extracellular Space	Hormone				✓	2.9%	0.0%	0.0%	25.9%
Sema3f	Extracellular Space	Signalling protein				✓	5.8%	6.7%	14.3%	11.1%
Avpr1a	Plasma Membrane	G-protein coupled receptor	✓				15.9%	0.0%	0.0%	3.7%
Crhr1	Plasma Membrane	G-protein coupled receptor				✓	4.3%	6.7%	14.3%	11.1%
Hcrtr1	Plasma Membrane	G-protein coupled receptor				✓	4.3%	6.7%	0.0%	11.1%
Htr2a	Plasma Membrane	G-protein coupled receptor	✓				11.6%	0.0%	0.0%	3.7%
Mc3r	Plasma Membrane	G-protein coupled receptor	✓				46.4%	26.7%	0.0%	25.9%
Npr3	Plasma Membrane	G-protein coupled receptor				✓	2.9%	0.0%	0.0%	14.8%
Npy2r	Plasma Membrane	G-protein coupled receptor				✓	27.5%	60.0%	14.3%	63.0%
Ntsr2	Plasma Membrane	G-protein coupled receptor			✓		5.8%	0.0%	28.6%	7.4%
Prlhr	Plasma Membrane	G-protein coupled receptor				✓	4.3%	13.3%	0.0%	18.5%
Prokr1	Plasma Membrane	G-protein coupled receptor				✓	8.7%	6.7%	0.0%	44.4%
Sstr3	Plasma Membrane	G-protein coupled receptor	✓				11.6%	0.0%	0.0%	0.0%
Tas1r1	Plasma Membrane	G-protein coupled receptor	✓				8.7%	0.0%	0.0%	0.0%
Gabra1	Plasma Membrane	Ion channel	✓				78.3%	46.7%	14.3%	48.1%
Gabrg1	Plasma Membrane	Ion channel	✓				91.3%	100.0%	71.4%	70.4%
Gria4	Plasma Membrane	Ion channel			✓		88.4%	86.7%	85.7%	81.5%
Grik1	Plasma Membrane	Ion channel	✓				37.7%	20.0%	28.6%	7.4%
Grik3	Plasma Membrane	Ion channel	✓				55.1%	33.3%	14.3%	33.3%
Kcna1	Plasma Membrane	Ion channel			✓		7.2%	0.0%	42.9%	0.0%
Kcna2	Plasma Membrane	Ion channel	✓				17.4%	0.0%	14.3%	0.0%
Kcnd1	Plasma Membrane	Ion channel			✓		11.6%	6.7%	14.3%	11.1%
Kcnh8	Plasma Membrane	Ion channel			✓		24.6%	33.3%	57.1%	40.7%
Kcnip1	Plasma Membrane	Ion channel	✓				87.0%	93.3%	57.1%	74.1%
Kcns2	Plasma Membrane	Ion channel	✓				8.7%	0.0%	0.0%	3.7%
Scn5a	Plasma Membrane	Ion channel	✓				15.9%	0.0%	0.0%	3.7%
Trpc3	Plasma Membrane	Ion channel	✓				14.5%	0.0%	0.0%	3.7%
Trpc5	Plasma Membrane	Ion channel	✓				73.9%	46.7%	14.3%	37.0%

A checkmark indicates an enrichment of the gene in the cluster and the proportion of cells expressing the gene in each of the clusters is shown on the right. For example, *Gal* is enriched in cluster 4 and it is expressed in 51.9% of the cells in the cluster.
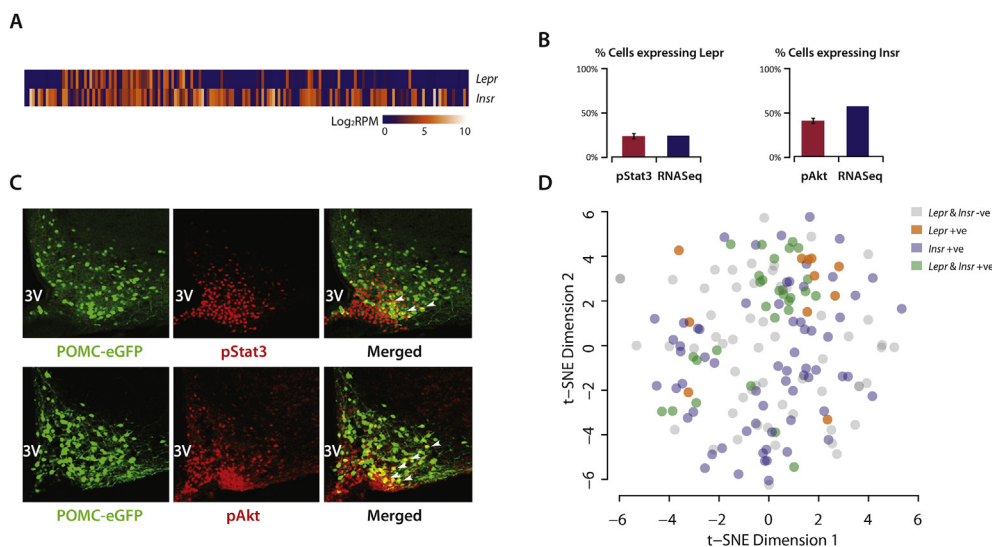

Figure 6: *Lepr* and *Insr* expression in hypothalamic POMC neurons. **(A)** We detected 30 *Lepr* and 83 *Insr* expressing cells out of the 163 cells in our sequencing cohort. **(B)** The percentages were compared to the number of cells that exhibited pStat3 and pAkt immunoreactivity post-administration of leptin and insulin (respectively). **(C)** Representative images showing POMC-eGFP, immunoreactivity of pStat3 (upper panel) and pAkt (lower panel). Cells that were double-stained were marked with white arrows. **(D)** The t-SNE plot shows a lack of clear clustering for *Lepr* and *Insr* expressing cells, indicating that expression of either receptor alone was not sufficient to drive segregation at the whole transcriptome level.

immunofluorescence evident more broadly through the arcuate nucleus (Figure 6B). We found that 41.1% of POMC-GFP neurons were positive for pAkt in response to insulin, which was not significantly different ($p = 0.357$) to the percentage of GFP neurons expressing *Insr* as determined by our sequencing data (Figure 2C). These findings increased confidence that the FACS sorting has captured a representative snapshot of the POMC neuronal population.

3. DISCUSSION

We describe here the use of single cell RNA sequencing to determine the transcriptomes of 163 individual murine POMC neurons. Unexpectedly, 25% of POMC positive neurons also express high levels of *Agrp*. The transcriptomes of these neurons are comparable to those of previously published pooled AgRP neurons [18], indicating that these are not an artifact but likely to be a real subpopulation of canonical AgRP neurons. Additionally, when we examine the raw data from Henry et al., we find that AgRP/NPY GFP neurons also express low levels of POMC [18]. This phenomenon likely reflects the shared developmental origins of POMC and AgRP cells [21]. POMC has been shown to be broadly expressed in hypothalamic progenitors, with up to 25% of the mature AgRP/NPY population sharing a common progenitor with POMC neurons [21]. The A^+P_+ population we describe here could be cells where POMC transcription is never fully silenced, even after the differentiation into AgRP/NPY neurons. Another possibility is that the high *Agrp/Npy* expressing neurons could be newly committed AgRP/NPY neurons from these progenitor cells, in which *Pomc* expression is yet to be fully suppressed. Regardless of the explanation, given that Pomc and AgRP genetic tools are in wide use, it is important for the field to know that a 'Pomc-specific' deletion of a given gene would also be removing the gene from a significant proportion of AgRP neurons. It is possible, for example, that the mild phenotype seen in the mice in which *Lepr* is 'specifically' deleted from POMC neurons [14] could be due in part to *Lepr* deletion from a proportion of AgRP neurons.

The remaining 118 P^+ neurons display striking heterogeneity and fall into four clusters. A few important points should be kept in mind while interpreting these data. First, these groups have emerged in an unbiased fashion after achieving a certain statistical threshold, which means that they are not invariant. Thus, increasing the threshold would invariably result in more clusters. Second, these groupings are not solely characterised by small numbers of unique genes; rather they are driven by a 'signature' expression profile of a whole repertoire of genes. Thus, while the driver genes (see Figure 5) are, on average, the highest expressed in their particular group, many transcripts are not unique to a single group and are not necessarily expressed in all of the neurons in any given group. Third, these different clusters are likely to be dynamic, depending on the nutritional state of the animal at the time the brain was obtained, with the mice in this particular experiment being *ad libitum* fed.

The group with the highest expression of *Pomc* and *Cartpt* is cluster 4, and based on its expression profile of GPCRs, appears to be responsive to corticotrophin releasing hormone, hypocretin, prolactin-releasing hormone, Npy, CCK, and neuromedin U. The largest group is cluster 1, which, based on its repertoire of GPCR expression, responds to vasopressin, melanocortin peptides (via the MC3R), serotonin, and somatostatin. Cluster 1 is notable in that, compared to the other 3 groups, it expresses the highest levels of a wide variety of plasma membrane ion channels. We acknowledge that further experiments will be required to functionally characterise these neurons. However, we speculate that the enrichment of distinct classes of

surface receptors on the different neuronal clusters possibly speaks to the temporal nature of the downstream effector function, with signalling through ion channels typically bringing about rapid changes, and circulating humoural factors providing longer term signals.

Two drugs currently on the market that are used for weight-loss are the serotonin receptor agonist Lorcaserin [27] and the GLP1 analogue Liraglutide [28]. There is evidence that both mediate their effects on food intake via ARC POMC neurons expressing 5HT$_{2C}$R [12,13] and GLP1R [22], respectively. Our data provide two additional pieces of information (Figure 3). First, less than 10% (11/118) of ARC POMC neurons express the GLP1R. Second, of the POMC neurons expressing either 5HT$_{2C}$R or GLP1R, only 3 neurons express both receptors. Thus, it is likely that different subgroups of POMC neurons mediate the satiety effects of each compound, which poses the question of whether a potential combinatorial therapy, without increasing individual drug-dose, might possibly result in increased weight-loss.

As previously mentioned, the 45 A^+P_+ neurons are very distinct from the 118 P^+ neurons, clearly segregating into a separate group. This is particularly true when one examines the expression of the leptin and insulin receptors in these groups. The low percentage (12%) of *Lepr*-expressing POMC (P^+) neurons is consistent with previous evidence that the actions of leptin on these neurons are largely indirect [19,29]. Although we did not specifically or comprehensively evaluate an *Agrp* positive population of neurons, the much higher percentage of LepR positivity in the cells where we documented high levels of *Agrp* expression (57%) support the idea that AgRP neurons are more likely to be direct targets of leptin [19]. Our data also provide some evidence of autocrine/paracrine regulation by POMC neurons themselves, with 10% of P^+ neurons expressing *Mc4r* and nearly 40% expressing the *Mc3r*. One could speculate that there may be amplification of the melanocortin system, whereby P^+ neurons that respond directly to leptin go on to signal to *Mc3r/Mc4r* expressing neurons, but this remains to be determined.

In contrast, the expression of insulin receptors is comparable between the A^+P_+ and P^+ neurons, indicating that insulins' actions on both these populations of neurons are likely to be direct [26,30].

In conclusion, single cell RNAseq has revealed previously unappreciated heterogeneity in POMC neurons. We have provided a comprehensive atlas, which we hope will be of utility to basic and translational researchers attempting to better understand the fundamental nature of regulation of energy balance by the hypothalamus and to manipulate these systems for therapeutic benefit.

4. MATERIAL AND METHODS

4.1. Animals
7−9-week-old male *Pomc-eGFP* transgenic mice [19] (kind gift from Prof. Malcolm Low) were housed under a standard 12 h light−dark cycle and were fed *ad libitum*. All studies were approved by the local Ethics Committee and conducted according to the UK Home Office Animals (Scientific Procedures) Act 1986.

4.2. Isolation of single hypothalamic POMC-eGFP neurons using fluorescence-activated cell sorting (FACS)
Ad libitum fed *Pomc-eGFP* mice were sacrificed by cervical dislocation, and tissue from the medial basal hypothalamus (MBH) region was micro-dissected into Hibernate A medium without calcium and magnesium (BrainBits, Springfield, IL, USA). The tissue was then digested with Papain (20 U/ml, Worthington, Lakewood, NJ, USA) for 30 min at

37 °C with mild agitation, followed by trituration in Hibernate A medium containing DNase I (0.005%, Worthington). The cell suspension was then filtered through a 40 μm cell strainer into a new Falcon tube. Fluorescence-activating cell sorting was performed using an Influx Cell Sorter (BD Biosciences, San Jose, CA, USA) utilizing a protocol that was previously used for single-cell isolation [31,32]. The cell gating was set according to cell size (FSC), cell granularity (SSC), FSC pulse-width for singlets, fluorescence at 488 nm/532 nm for GFP and 647/670 nm for nuclear stain with DraQ5 (Biostatus, Shepshed, Leicester, UK). Cells were sorted direct into a 96 well plate containing lysis buffer with RNase inhibitor.

4.3. Whole-transcriptome amplification and single-cell RNA sequencing

RNA from the single cell lysate was reverse transcribed and PCR amplified (19 cycles) into double stranded cDNA using Clontech SMART-Seq v4 Ultra Low Input RNA Kit (Takara Clontech, Mountain View, CA, USA). 150 pg of amplified cDNA was then used to generate barcoded sequencing libraries using Illumina Nextera XT library preparation kit (San Diego, CA, USA) according to manufacturer's manual. Samples failing to show amplification were removed from the experiment.

The sequencing libraries were normalized to 10 nM concentration and combined into pools of 96. The pooled libraries were sequenced on 4 lanes of an Illumina HiSeq 4000 instrument at single-end 50bp (SE50), yielding an average of 15.7 million reads per cell. Library preparation was performed by the Genomics and Transcriptomic Core at the Institute of Metabolic Science. The sequencing was performed at the Genomics Core, Cancer Research UK Cambridge Institute. All the raw data in this manuscript has been submitted to GEO, accession number GSE92707.

4.4. Leptin and insulin administrations

For pSTAT3 immuno-labelling, mice were fasted overnight and leptin (3 mg/kg, PeproTech, Rocky Hill, NJ, USA) or saline was injected intraperitoneally. 45 min after leptin administration, the animals were sacrificed by a lethal overdose of pentobarbital (60 mg/kg) and perfused transcardially with PBS followed by 4% paraformaldehyde. For pAKT immunofluorescence, overnight fasted animals were administered Actrapid (2 mU/kg, Novo Nordisk, Bagsværd, Denmark) intravenously. Animals were sacrificed 15 min after administration and perfused as described before.

4.5. Immunohistochemistry for POMC-eGFP, pStat3 and pAkt

Brains were collected immediately after perfusions and fixed for two additional hours in a 15% sucrose- 4% paraformaldehyde solution. The brains were cryo-preserved in 30% sucrose PBS overnight, embedded in OCT (Bright, Huntingdon, UK), and stored at −80 °C until further processing. Frozen coronal sections were sliced at 30 μm and processed for immunofluorescence. For pSTAT3 immunostaining, sections were pre-treated for 20 min in 1% NaOH and then placed overnight with a rabbit anti-pSTAT3 antibody (1:200, Tyr-705 Cell Signalling). For pAKT immunostaining, sections were incubated with anti-pAKT antibody (1:100, Ser-473 Cell Signalling) overnight. The primary antibodies were localized with Dylight 594 Goat anti-Rabbit IgGs (Vector; 1:500). Coronal images of the arcuate nucleus were taken using a Laser-scanning confocal microscope (LSM 510, Zeiss, Oberkochen, Germany), and eGFP positive cells and double labelled GFP-pSTAT3- or GFP-pAKT immune-reactive cells were counted using Image J.

4.6. Immunohistochemistry for Mc4r-MAPT, and Pomc

B6.Cg-Tg(Mc4r-MAPT/Sapphire)21Rck/J male mice were perfused transcardially via a 23-gauge needle placed in the left ventricle with 100 ml of 0.1 M heparinized PBS, pH 7.4, followed by 100 ml of 4% paraformaldehyde in PBS, and the fixed brains were cryoprotected in 30% sucrose for 48 h at 4C. Coronal hypothalamic sections of 30 μm thickness were prepared on a freezing microtome. Free floating sections were incubated for 15 min in 1% hydrogen peroxide, washed 2 times in PBS, blocked 2 h in 0.3% Triton X-100 and 5% NGS in PBS, before being incubated overnight at 4 °C in rabbit anti-proopiomelanocortin precursor (1:200, Phoenix Pharmaceuticals) in 0.3% Triton X-100 and 5% NGS. Sections were then washed and incubated for 2 h with Alexa Fluor 594 goat anti-rabbit IgG. Sections were washed and incubated overnight at 4 °C in chicken anti-GFP (1:500, Abcam). Sections were washed and incubated in Alex Fluor 488 goat anti-chicken IgG for 2 h. Sections were floated onto super-frost Plus microscope slides (Fisher), and coverslipped with Vecta-shield (Vector). Tissue sections were digitized using a Zeiss Axioplan fluorescent microscope, and areas of interest were outlined based on cellular morphology.

4.7. Bioinformatics

The quality of sequence reads was examined using FastQC, and Multiple Genome alignment (MGA) was used to rule out contamination from other DNA sources throughout the wet experimental procedures. Reads were aligned to the mouse genome (GRCm38, Ensembl v74) using Tophat version 2.0.11 [33], and the raw counts at the gene level were determined using Cufflinks version 2.2.1 [34]. The above analyses were performed at the High-performance computing cluster at the Research Computing Service, University of Cambridge.

The raw counts from cells were normalized and batch correction was performed using edgeR [35] and limma [36] packages. Rtsne package was used to generate t-SNE plots [20] and SC3 package was used to perform the clustering analysis [23]. GeneSpring GX software (Agilent) was used to generate the heatmaps. The normalized gene abundance is represented in $\log_2$RPM, which stands for the base-2 logarithm of Reads per Million. The plots were generated using ggplot2 package, Microsoft Excel and GeneSpring GX. Gene annotations and pathway analysis were perform using Ingenuity Pathway Analysis (Qiagen, Redwood City, CA, USA).

CONFLICT OF INTEREST

None declared.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.molmet.2017.02.007.

REFERENCES

[1] Yeo, G.S., Heisler, L.K., 2012. Unraveling the brain regulation of appetite: lessons from genetics. Nature Neuroscience 15:1343—1349.

Original Article

[2] Cone, R.D., 2005. Anatomy and regulation of the central melanocortin system. Nature Neuroscience 8:571–578.

[3] Bertagna, X., 1994. Proopiomelanocortin-derived peptides. Endocrinology Metabolism Clinics of North America 23:467–485.

[4] Castro, M.G., Morrison, E., 1997. Post-translational processing of proopiomelanocortin in the pituitary and in the brain. Critical Reviews in Neurobiology 11:35–57.

[5] Fan, W., Boston, B.A., Kesterson, R.A., Hruby, V.J., Cone, R.D., 1997. Role of melanocortinergic neurons in feeding and the agouti obesity syndrome. Nature 385:165–168.

[6] Ollmann, M.M., Wilson, B.D., Yang, Y.K., Kerns, J.A., Chen, Y., Gantz, I., et al., 1997. Antagonism of central melanocortin receptors in vitro and in vivo by agouti-related protein. Science 278:135–138.

[7] Woods, A.J., Stock, M.J., 1996. Leptin activation in hypothalamus. Nature 381:745.

[8] Williams, K.W., Margatho, L.O., Lee, C.E., Choi, M., Lee, S., Scott, M.M., et al., 2010. Segregation of acute leptin and insulin effects in distinct populations of arcuate proopiomelanocortin neurons. Journal of Neuroscience 30:2472–2479.

[9] Hentges, S.T., Otero-Corchon, V., Pennock, R.L., King, C.M., Low, M.J., 2009. Proopiomelanocortin expression in both GABA and glutamate neurons. Journal of Neuroscience 29:13684–13690.

[10] Jarvie, B.C., Hentges, S.T., 2012. Expression of GABAergic and glutamatergic phenotypic markers in hypothalamic proopiomelanocortin neurons. The Journal of Comparative Neurology 520:3863–3876.

[11] Meister, B., Gomuc, B., Suarez, E., Ishii, Y., Durr, K., Gillberg, L., 2006. Hypothalamic proopiomelanocortin (POMC) neurons have a cholinergic phenotype. European Journal of Neuroscience 24:2731–2740.

[12] Burke, L.K., Doslikova, B., D'Agostino, G., Greenwald-Yarnell, M., Georgescu, T., Chianese, R., et al., 2016. Sex difference in physical activity, energy expenditure and obesity driven by a subpopulation of hypothalamic POMC neurons. Molecular Metabolism 5:245–252.

[13] Heisler, L.K., Cowley, M.A., Tecott, L.H., Fan, W., Low, M.J., Smart, J.L., et al., 2002. Activation of central melanocortin pathways by fenfluramine. Science 297:609–611.

[14] Balthasar, N., Coppari, R., McMinn, J., Liu, S.M., Lee, C.E., Tang, V., et al., 2004. Leptin receptor signaling in POMC neurons is required for normal body weight homeostasis. Neuron 42:983–991.

[15] Challis, B.G., Coll, A.P., Yeo, G.S., Pinnock, S.B., Dickson, S.L., Thresher, R.R., et al., 2004. Mice lacking pro-opiomelanocortin are sensitive to high-fat feeding but respond normally to the acute anorectic effects of peptide-YY(3-36). Proceedings of the National Academy of Sciences of the United States of America 101:4695–4700.

[16] Yaswen, L., Diehl, N., Brennan, M.B., Hochgeschwender, U., 1999. Obesity in the mouse model of pro-opiomelanocortin deficiency responds to peripheral melanocortin. Nature Medicine 5:1066–1070.

[17] Chen, H., Charlat, O., Tartaglia, L.A., Woolf, E.A., Weng, X., Ellis, S.J., et al., 1996. Evidence that the diabetes gene encodes the leptin receptor: identification of a mutation in the leptin receptor gene in db/db mice. Cell 84:491–495.

[18] Henry, F.E., Sugino, K., Tozer, A., Branco, T., Sternson, S.M., 2015. Cell type-specific transcriptomics of hypothalamic energy-sensing neuron responses to weight-loss. Elife 4.

[19] Cowley, M.A., Smart, J.L., Rubinstein, M., Cerdan, M.G., Diano, S., Horvath, T.L., et al., 2001. Leptin activates anorexigenic POMC neurons through a neural network in the arcuate nucleus. Nature 411:480–484.

[20] Maaten, Lvd, Hinton, G., 2008. Visualizing data using t-SNE. Journal of Machine Learning Research 9:2579–2605.

[21] Padilla, S.L., Carmody, J.S., Zeltser, L.M., 2010. Pomc-expressing progenitors give rise to antagonistic neuronal populations in hypothalamic feeding circuits. Nature Medicine 16:403–405.

[22] Secher, A., Jelsing, J., Baquero, A.F., Hecksher-Sorensen, J., Cowley, M.A., Dalboge, L.S., et al., 2014. The arcuate nucleus mediates GLP-1 receptor agonist liraglutide-dependent weight loss. Journal of Clinical Investigation 124:4473–4488.

[23] Kiselev, V.Y., Kirschner, K., Schaub, M.T., Andrews, T., Chandra, T., Natarajan, K.N., et al., 2016. SC3-consensus clustering of single-cell RNA-Seq data. bioRxiv. http://dx.doi.org/10.1101/036558.

[24] Hentges, S.T., Nishiyama, M., Overstreet, L.S., Stenzel-Poore, M., Williams, J.T., Low, M.J., 2004. GABA release from proopiomelanocortin neurons. Journal of Neuroscience 24:1578–1583.

[25] Lam, D.D., Attard, C.A., Mercer, A.J., Myers Jr., M.G., Rubinstein, M., Low, M.J., 2015. Conditional expression of Pomc in the Lepr-positive subpopulation of POMC neurons is sufficient for normal energy homeostasis and metabolism. Endocrinology 156:1292–1302.

[26] Hill, J.W., Elias, C.F., Fukuda, M., Williams, K.W., Berglund, E.D., Holland, W.L., et al., 2010. Direct insulin and leptin action on proopiomelanocortin neurons is required for normal glucose homeostasis and fertility. Cell Metabolism 11:286–297.

[27] Smith, S.R., Weissman, N.J., Anderson, C.M., Sanchez, M., Chuang, E., Stubbe, S., et al., 2010. Multicenter, placebo-controlled trial of lorcaserin for weight management. The New England Journal of Medicine 363:245–256.

[28] Astrup, A., Rossner, S., Van Gaal, L., Rissanen, A., Niskanen, L., Al Hakim, M., et al., 2009. Effects of liraglutide in the treatment of obesity: a randomised, double-blind, placebo-controlled study. Lancet 374:1606–1616.

[29] Vong, L., Ye, C., Yang, Z., Choi, B., Chua Jr., S., Lowell, B.B., 2011. Leptin action on GABAergic neurons prevents obesity and reduces inhibitory tone to POMC neurons. Neuron 71:142–154.

[30] Konner, A.C., Janoschek, R., Plum, L., Jordan, S.D., Rother, E., Ma, X., et al., 2007. Insulin action in AgRP-expressing neurons is required for suppression of hepatic glucose production. Cell Metabolism 5:438–449.

[31] Macaulay, I.C., Svensson, V., Labalette, C., Ferreira, L., Hamey, F., Voet, T., et al., 2016. Single-cell RNA-sequencing reveals a continuous spectrum of differentiation in hematopoietic cells. Cell Reports 14:966–977.

[32] Schulte, R., Wilson, N.K., Prick, J.C., Cossetti, C., Maj, M.K., Gottgens, B., et al., 2015. Index sorting resolves heterogeneous murine hematopoietic stem cell populations. Experimental Hematology 43:803–811.

[33] Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S.L., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biology 14:R36.

[34] Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., et al., 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature Biotechnology 28:511–515.

[35] Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140.

[36] Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., et al., 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research 43:e47.