Single-Cell Transcriptomic Evidence for Dense Intracortical Neuropeptide Networks

Stephen J Smith, Uygar Sümbül, Lucas Graybuck, Forrest Collman, Sharmishtaa Seshamani, Rohan Gala, Olga Gliko, Leila Elabbady, Jeremy A. Miller, Trygve Bakken, *Jean Rossier, Zizhen Yao, Ed Lein, Hongkui Zeng, Bosiljka Tasic, Michael Hawrylycz

Allen Institute for Brain Science, 615 Westlake Ave N, Seattle, WA 98109, USA
* Neuroscience Paris Seine, Sorbonne Université, 9 Quai Saint Bernard, 75005 Paris, France

Briefly

Analysis of single-cell RNA-Seq data from mouse neocortex exposes evidence for local neuropeptidergic modulation networks that involve every cortical neuron directly.

Data Highlights

- At least 98% of mouse neocortical neurons express one or more of 18 neuropeptide precursor proteins (NPP) genes.
- At least 98% of cortical neurons express one or more of 29 neuropeptide-selective G-protein-coupled receptor (NP-GPCR) genes.
- Neocortical expression of these 18 NPP and 29 NP-GPCR genes is highly neuron-type-specific and permits exceptionally powerful differentiation of transcriptomic neuron types.
- Neuron-type-specific expression of 37 cognate NPP / NP-GPCR gene pairs predicts modulatory connectivity within 37 or more neuron-type-specific intracortical networks.

Summary

Seeking insight into homeostasis, modulation and plasticity of cortical synaptic networks, we analyzed results from deep RNA-Seq analysis of 22,439 individual mouse neocortical neurons. This work exposes transcriptomic evidence that all cortical neurons participate directly in highly multiplexed networks of modulatory neuropeptide (NP) signaling. The evidence begins with a discovery that transcripts of one or more neuropeptide precursor (NPP) and one or more neuropeptide-selective G-protein-coupled receptor (NP-GPCR) genes are highly abundant in nearly all cortical neurons. Individual neurons express diverse subsets of NP signaling genes drawn from a palette encoding 18 NPPs and 29 NP-GPCRs. Remarkably, these 47 genes comprise 37 cognate NPP/NP-GPCR pairs, implying a strong likelihood of dense, cortically localized neuropeptide signaling. Here we use neuron-type-specific NP gene expression signatures to put forth specific, testable predictions regarding 37 peptidergic neuromodulatory networks that may play prominent roles in cortical homeostasis and plasticity.

Correspondence: Stephen J Smith, stephens@alleninstitute.org

10,293 words (main text, in-line citations and figure legends), 127 references

Keywords: transcriptomes, neuropeptides, mouse, neocortex, visual cortex, motor cortex, neuron types, neuron-type taxonomy, neurotaxonomy, mRNA-seq, GPCR, synaptic plasticity, neuromodulation, homeostasis, synaptic network homeostasis

Introduction

Neuromodulation - the adjustment of synapse and ion channel function via diffusible cell-cell signaling molecules - is a fundamental requirement for adaptive nervous system function (Abbott and Regehr, 2004; Bargmann, 2012; Bucher and Marder, 2013; Marder, 2012; Marder et al., 2015; Mccormick and Nusbaum, 2014; Nadim and Bucher, 2014; Nusbaum et al., 2017). Neuromodulator molecules take many different chemical forms, including diatomic gases such as nitric oxide, lipid metabolites such as the endocannabinoids, and amino acids and their metabolites such as glutamate, GABA, acetylcholine, serotonin and dopamine. By far the largest family of neuromodulator molecules, however, comprises the evolutionarily ancient proteinaceous signaling molecules known as neuropeptides (Baraban and Tallent, 2004; Burbach, 2011; Gonzalez-Suarez and Nitabach, 2018; Hökfelt et al., 2013; van den Pol, 2012; Wang et al., 2015). The most well-known and widely studied neuropeptides are the endogenous "opioid" peptides - enkephalins, endorphins and dynorphins - but there are nearly one hundred other NPP genes in the human genome and numerous homologs are present in all known animal genomes except for those of the sponges (*Porifera*) (Elphick et al., 2018; Jekely, 2013).

The broadest definition of "neuropeptide" would embrace any soluble peptide that serves as a messenger by diffusing from one neuron to another. A narrower but more common definition (Burbach, 2011) requires that (1) a neuropeptide precursor protein (NPP) transcript be translated as an NPP into the lumen of a source neuron's rough endoplasmic reticulum (rER), (2) packaged into dense-core vesicles (DCVs) and enzymatically cleaved into one or more neuropeptide (NP) products after passage through the rER–Golgi complex, (3) transported and stored within the source neuron in DCVs, (4) released upon demand by activity- and calcium-dependent exocytosis, and only then (5) diffuse interstitially to act upon a target neuron by binding to a specific receptor. This pathway enlarges the potential palette of distinct neuropeptides beyond that established simply by the large number of NPP genes, as a given NPP may be cleaved into alternative NP products during its intracellular and interstitial passage.

Most neuropeptide receptors are encoded by members of the very large superfamily of G-protein-coupled receptor (GPCR) genes (Hoyer and Bartfai, 2012; Krishnan and Schioth, 2015; Mains and Eipper, 2006; van den Pol, 2012). GPCRs are selective, high-affinity receptors distinguished by characteristic seven-transmembrane-segment atomic structures and signal transduction involving heterotrimeric G-proteins (hence the name). Phylogenomic evidence suggests that the earliest behaving animals relied exclusively upon early neuropeptide homologs and cognate neuropeptide-selective GPCRs (NP-GPCRs) for the slow intercellular communication sufficient to generate their slow and simple behaviors (Elphick et al., 2018; Grimmelikhuijzen and Hauser, 2012; Jekely, 2013; Krishnan and Schioth, 2015; Varoqueaux and Fasshauer, 2017). The later evolution of neurons, focal synaptic contacts, rapidly recycled small-molecule neurotransmitters, and numerous ionotropic receptors was likely driven by survival advantages of faster cell-cell signaling (Varoqueaux and Fasshauer, 2017). The fast synaptic transmission characteristic of contemporary higher animals is almost invariably based on recycling small molecule neurotransmitters and ionotropic receptors, but modulation of synaptic transmission and membrane excitability by NP-GPCRs remains very prominent in all extant behaving animals (Elphick et al., 2018; Grimmelikhuijzen and Hauser, 2012; Jekely, 2013; Krishnan and Schioth, 2015; Varoqueaux and Fasshauer, 2017).

Because modulatory neuropeptides are not subject to the rapid transmitter re-uptake and/or degradation processes necessary for fast synaptic transmission, secreted neuropeptides persist long enough (e.g., minutes) in brain interstitial spaces for diffusion to NP-GPCRs hundreds of micrometers distant from release sites (Ludwig and Leng, 2006; Nässel, 2009; Russo, 2017). Neuropeptide signaling in the CNS can thus be presumed "paracrine", with secretion from one neuron acting upon many others by diffusion over distance and signals likewise converging by diffusion from many neurons onto one. The degradation of active neuropeptides by extracellular peptidases in cortex is nonetheless generally expected to restrict signal diffusion to sub-millimeter scale local circuit volumes, such as cortical "columns" or "barrels" or other commonly envisioned small anatomic/functional subunit tiles of the cortical sheet.

The many receptors encoded by different NP-GPCR genes are each highly selective for specific peptides but show considerable conservation at the level of downstream cellular signal transduction effects. Although GPCR signaling has long been recognized as complex and many faceted (Hamm, 1998), most

neuronal NP-GPCR actions reflect phosphorylation of ion channel or synaptic proteins, mediated by protein kinases dependent on the second messengers cyclic AMP and calcium (Mains and Eipper, 2006; Nadim and Bucher, 2014; van den Pol, 2012). Primary effects of NP-GPCRs, in turn, fall into just three major categories distinguished by G-protein alpha subunit class. The Gαi class (i) inhibits cAMP production, the Gαs class (s) stimulates cAMP production, and the Gαq class (q) amplifies calcium signaling dynamics (Syrovatkina et al., 2016). For most NP-GPCR genes, the primary G-protein α-subunit class (i.e., i, s or q) is now known (Alexander et al., 2017) and offers a good first-order prediction of the encoded GPCR's signal transduction activity. The profound functional consequences of neuromodulation by GPCRs range from adjustment of neuronal firing properties and calcium signaling dynamics through regulation of synaptic weights and synaptic plasticity (Bargmann, 2012; Markram et al., 2013; Mccormick and Nusbaum, 2014).

It is well established that particular neuropeptides, including vasoactive intestinal peptide (VIP), somatostatin (SST), neuropeptide Y (NPY), substance P, and cholecystokinin (CCK), are detectible at high levels in particular subsets of GABAergic cortical neurons (Tremblay et al., 2016). These neuropeptides, consequently, have come into broad use as markers for GABAergic interneuron classes, while the corresponding NPP and NP-GPCR genetics have provided molecular access to these and other broad neuron type classes (Daigle et al., 2018; Maximiliano José et al., 2018). *In situ* hybridization and microarray data (e.g., the Allen Brain Atlases (Hawrylycz et al., 2012; Lein et al., 2007)) have also established that mRNA transcripts encoding these five NPPs and that many other NPPs and cognate NP-GPCR genes are expressed differentially in different brain regions. There has been a critical lack, however, of comprehensive expression data combining whole-genome depth with single-cell resolution. Absent such data, it has been difficult to generate specific and testable hypotheses regarding cortical neuropeptide function and to design repeatable experiments to test those hypotheses (Tremblay et al., 2016; van den Pol, 2012).

Here we describe new findings regarding NPP and NP-GPCR gene expression in single cortical neurons, based on analysis of deep mRNA-Seq data acquired from 22,439 isolated mouse cortical neurons as described fully in a recent publication (Tasic et al., 2018). We begin by leveraging only the genomic depth and single-cell resolution of this dataset. Then, we briefly introduce the transcriptomic neurotaxonomy (i.e., neuron-type taxonomy) also developed in the Tasic 2018 publication and explore the additional analytical power of a taxonomic framework. Finally, we distill these findings into specific and testable predictions concerning intracortical peptidergic modulation networks.

Results

The present study is based on analysis of a resource single-cell mRNA-Seq dataset acquired at the Allen Institute (Tasic et al., 2018) and available for download at *http://celltypes.brain-map.org/rnaseq/*. These RNA-Seq data were acquired from a total of 22,439 isolated neurons, with detection of transcripts from a median of 9,462 genes per cell (min = 1,445; max = 15,338) and an overall total of 21,931 protein-coding genes detected. Neurons were sampled from two distant and very different neocortical areas: 13,491 neurons from primary visual cortex (VISp), and 8,948 neurons from anterior lateral motor cortex (ALM). Tasic, et al., harvested tissue specimens from a variety of transgenic mice expressing fluorescent proteins to enable enrichment of samples for neurons and for relatively rare neuron types by FACS sorting after dissociation. This enrichment procedure resulted, by design, in a disproportionate representation of GABAergic neurons, canonically ~20% of neurons (Sahara et al., 2012), such that the sampled neuron population is roughly half GABAergic (47%) and half glutamatergic (53%). The resource publication (Tasic et al., 2018) should be consulted for full details of neuronal sample and library preparation, sequencing and data processing.

The resource single-cell RNA-Seq data tables (Tasic et al., 2018) report the abundance of transcripts from individual neurons in both "counts per million reads" (CPM) and "fragments per kilobase of exon per million reads mapped" (FPKM) units. Our analysis of this data compares gene expression levels quantitatively, with two distinct use cases: (1) comparisons across large sets of different genes, and (2) comparisons of the same gene across different individual cells, cell types and brain areas. We have relied upon FPKM data (Mortazavi et al., 2008; Pimentel, 2014), for use case 1 (i.e., the Table 1 and 2

comparisons across genes). For use case 2 (as in all figures below), we have preferred the CPM units, because these units were used to generate the Tasic 2018 neurotaxonomy. While choice of units here seems unlikely to make any significant difference, it would seem inconsistent to use FPKM units to compare across cell types discerned from CPM data.

The NP signaling genes upon which the present analysis focuses are expressed very differentially across the sampled populations of individual mouse cortical neurons. That is, each gene is expressed at a high level in some subset of cells but at zero or very low levels in the remainder of the population. To compactly characterize such expression, we developed a "Peak FPKM" metric. This metric is generated by ranking single-cell FPKM values for a given gene across the entire population of 22,439 neurons sampled, then designating the FPKM value at the ascending 99.9th percentile point as "Peak FPKM". This metric was designed to minimize effects of sporadic outliers while still closely approximating the actual peak expression value in even very small subsets of neurons expressing the gene in question.

18 Neuropeptide Precursor Protein (NPP) genes are extremely highly expressed in mouse neocortex. **Table 1** lists 18 NPP genes highly expressed in varied subsets of the 22,439 individual neurons sampled from cortical areas VISp and ALM. This gene list was circumscribed by two requirements: (1) that the included NPP gene be highly expressed (top quartile Peak FPKM, across all protein-coding genes) in both VISp and ALM cortical areas, and (2) that at least one NP-GPCR gene cognate to a candidate NPP gene also be highly expressed in neurons within the same cortical local areas. Requirement (2) was imposed here to focus on prospects for intracortical paracrine neuropeptide signaling as noted in **Introduction** above. **Table 1** also lists Peak FPKM values for each NPP gene, percentile and absolute ranks of that Peak FPKM value across all protein-coding genes, the fraction of cells sampled in which expression of the listed gene is detectible, predicted neuropeptide product(s) encoded, and the NP-GPCR gene(s) fulfilling requirement (2) for that NPP gene. Transcripts of no other known NPP genes met the criteria specified above.

	NPP Gene	peak FPKM	pFPKM Percentile	pFPKM Rank	% Cells	Predicted Neuropeptides	Cognate NP-GPCR Genes
1	Npy	108,865	100.00	1	42	Neuropeptide Y	Npy1r, Npy2r, Npy5
2	Sst	70,274	99.99	2	26	Somatostatins	Sstr1, Sstr2, Sstr3, Sstr4
3	Vip	48,747	99.99	3	33	Vasoactive Intestinal Peptide	Vipr1, Vipr2
4	Tac2	18,284	99.98	4	15	Neurokinin B	Tacr3
5	Cck	16,396	99.97	6	69	Cholecystokinins	Cckbr
6	Penk	11,160	99.96	8	26	Enkephalins	Oprd1, Oprm1
7	Crh	9,118	99.95	10	17	Corticotropin-Releasing Hormone	Crhr1, Crhr2
8	Cort	7,477	99.93	15	32	Cortistatin	Sstr1, Sstr2, Sstr3, Sstr4
9	Tac1	5,728	99.92	18	11	Substance P, Neurokinin A	Tacr1
10	Pdyn	2,813	99.69	68	8	Dynorphins	Oprd1, Oprk1, Oprm1
11	Pthlh	1,656	99.29	156	18	Parathyroid-Hormone-Like Hormone	Pth1r
12	Pnoc	698	97.68	509	23	Nociceptins	Oprl1
13	Trh	510	96.51	766	3	Thyrotropin-Releasing Hormone	Trhr, Trhr2
14	Grp	435	95.59	968	12	Gastrin-Releasing Peptide	Grpr
15	Rln1	258	91.99	1757	7	Relaxin 1	Rxfp1, Rxfp2, Rxfp3
16	Adcyap1	165	87.29	2788	26	Adenylate Cyclase-Activating Polypeptides	Adcyap1r1, Vipr1, Vipr2
17	Nts	121	82.14	3917	1	Neurotensin	Ntsr1, Ntsr2
18	Nmb	112	80.53	4270	14	Neuromedin B	Nmbr

Table 1. 18 neuropeptide precursor protein (NPP) genes are highly expressed in mouse cortex. These genes are tabulated here along with peak single-cell expression levels as pFPKM (Peak FPKM, see main text), percentile and absolute ranking of these pFPKM values across pFPKMs for all 21,931 protein-coding genes, and the percentage of cells sampled in which transcripts of the specified NPP gene were detected. The table also lists predicted neuropeptide products, and genes encoding G-protein-coupled receptors (NP-GPCRs) that are cognate to the listed NPP genes and expressed cortex (see **Table 2**). NPP genes are listed here in descending order of Peak FPKM.

The Peak FPKM ranking columns in **Table 1** show that expression levels of most of the 18 NPP genes are extremely high in the range of Peak FPKM values for all 21,931 protein-coding genes detected in all

neurons sampled. Of these genes, Npy, Sst, Vip and Tac2 rank as the top four overall in peak FPKM values, while three more, Cck, Penk and Crh also rank in the top ten. Eleven of these NPP genes rank in the top percentile and all 18 rank above the 80th percentile in peak FPKM. The extremely high peak abundance of these NPP transcripts suggests that NPP products are likely synthesized in the highly expressing cells at correspondingly high rates. To maintain a steady state, the cell must therefore eliminate those protein products at a very high rate, with processing and secretion of active neuropeptides being the most likely route of elimination. The high abundance of transcripts encoding these 18 NPPs can thus be construed as evidence for secretion of the respective active neuropeptide products.

Expression of NPP genes by neocortical neurons is highly differential. **Figure 1A** characterizes differential expression of the 18 NPP genes of **Table 1**. Each of 18 color-coded solid curves represents the distribution of single-neuron CPM values for one NPP gene. Curves were generated by plotting CPM for each individual neuron in descending rank order along a sampled cell population percentile axis. Each curve shows an abrupt transition from very high to very low (commonly zero) expression across the sampled neuron population, but these transitions occur at very different population percentile points, providing clear evidence for highly differential single-cell expression of each gene. Percentages of the sampled neuron population expressing a given NPP gene (at greater than 1 CPM) range from more than 65% for Cck down to 1% for Nts. Recall that the cell population sampled here has been enriched for GABAergic cell types as noted above and described at length in the resource publication (Tasic et al., 2018).

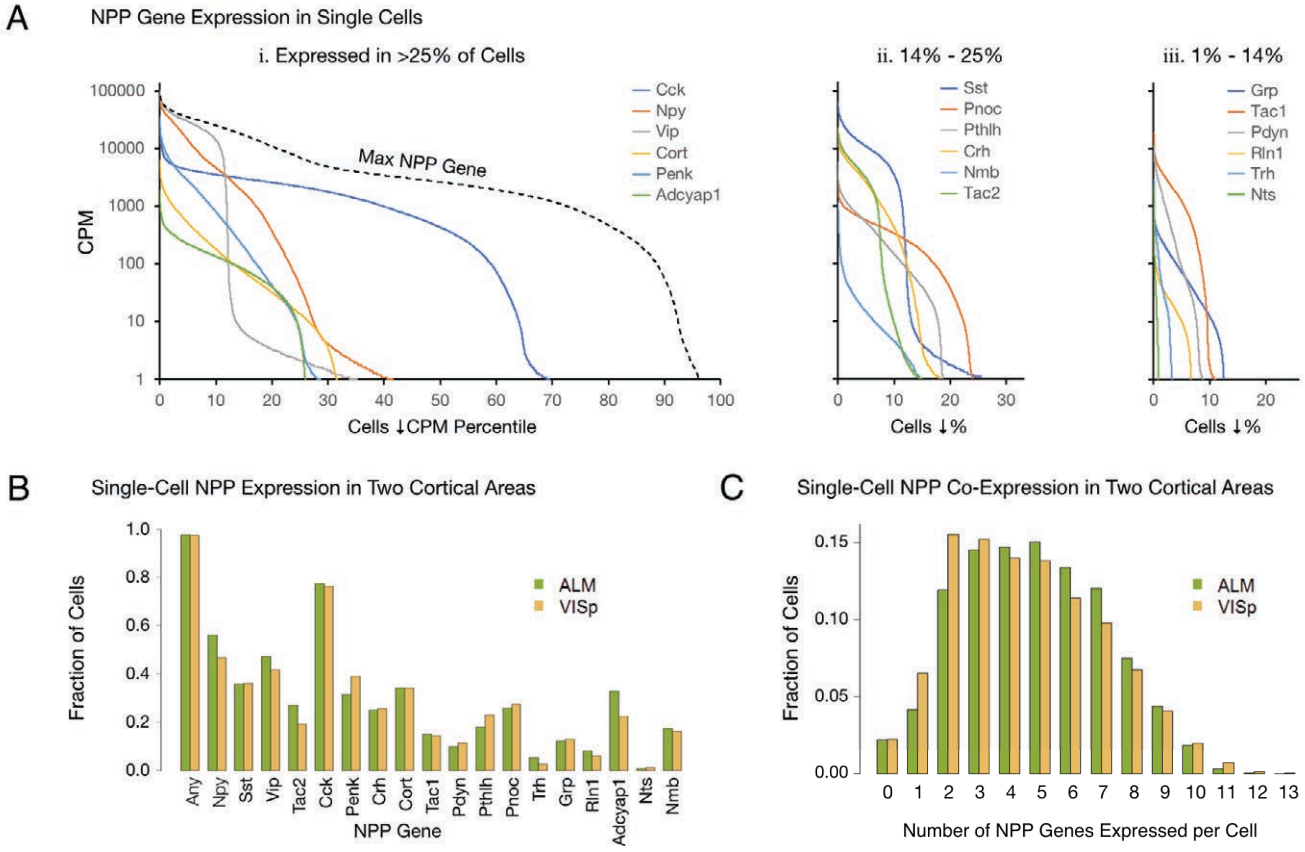


Figure 1. Expression and co-expression of NPP genes is highly differential with statistics conserved between two distant cortical areas. Single-cell expression analysis of 18 NPP genes in 22,439 individual neurons from cortical areas VISp and ALM. **(A)** Distributions of single-cell expression levels in neurons pooled from VISp and ALM. Color-coded solid curves plot single-cell CPM values for the specified individual NPP genes in descending order along a cell population percentile axis. The 18 curves are segregated for clarity into three panels, sorted by cell population percentiles at which CPM values fall below 1. Large differences in fractions of cells expressing different NPP genes are evident. The dashed curve labeled "Max NPP Gene" in panel **Ai** was generated by plotting CPM values of the highest CPM NPP gene for each individual cell in descending order. **(B)** Fractions of cells expressing each NPP genes represented separately for 13,491 VISp neurons and 8,948 ALM neurons,

showing conservation between areas of the patterning of NPP expression fractions documented in panel A. **(C)** Histograms illustrating frequencies of various multiples of NPP gene co-expression in individual neurons, represented separately for VISp and ALM neurons. The paired vertical bars show strong conservation of co-expression patterns between the two areas.

Almost all (and possibly all) neocortical neurons express at least one NPP gene. The dashed curve in the **Fig. 1Ai**, labeled "Max NPP Gene", was generated by plotting CPM values of the NPP gene with the highest CPM in each individual cell in descending order along a cell population percentile axis. This curve therefore shows that 97% percent of the sampled mouse cortical neurons express at least one NPP gene at >1 CPM and that 80% express at least one NPP gene at >1,000 CPM, a very high level. When one takes into account the pulsatile nature of transcription (Suter et al., 2011) and the stochastic nature of RNA-Seq transcript sampling (Fu and Pachter, 2016; Kim et al., 2015; Tasic et al., 2016), these numbers must be understood as lower limits to percentages of cortical neurons expressing at least one of the 18 NPP genes. The results summarized in **Fig. 1A** may therefore be consistent with the proposition that every cortical neuron is peptidergic.

Statistics of differential NPP gene expression are highly conserved between different neocortical areas. **Figures 1B** and **1C** illustrate strong conservation of differential NPP expression profiles between VISp and ALM, two distant and very different neocortical areas. The paired bars in **Fig. 1B** represent fractions of cells expressing a given gene in each of the two areas. It is obvious that the differential expression profiles in VISp and ALM are highly similar (ρ=0.972, p<1.72E-11), in spite of stark differences in function and cytoarchitecture between these two areas. Conservation of expression fractions across so many genes in such divergent cortical areas may suggest that these patterns have strong connections to conserved features of cortical function and argues against these patterns being secondary to more ephemeral variables such as neuronal activity patterns, which seem unlikely to be highly conserved between VISp and ALM areas.

Multiple NPP genes are co-expressed in almost all cortical neurons. **Figure 1C** represents frequencies with which transcripts of various multiples drawn from the set of 18 NPP genes were detected in individual neurons. These data establish a remarkable degree of NPP gene co-expression in almost all individual cortical neurons. The modal number of co-expressed NPP genes detected is 2 in VISp and 5 in ALM, but both distributions are actually quite flat between 2 and 5, with broad plateaus out to 7 co-expressed NPP genes per cell and a substantial tail out to 10. **Fig. 1C** also profiles strong similarities of NPP co-expression distributions between VISp and ALM.

29 Neuropeptide-selective G-protein-coupled receptor (NP-GPCR) genes are highly expressed in mouse neocortex. **Table 2** lists 29 NP-GPCR genes that are highly expressed in varied subsets of the 22,439 individual neurons sampled from cortical areas VISp and ALM. These 29 genes encode receptor proteins selective for neuropeptide products encoded by the 18 NPP genes of **Table 1** (cross-referenced in that table as "Cognate NP-GPCR Genes"). **Table 2** provides quantitative information on expression levels of these 29 NP-GPCR genes, names the receptor proteins they encode, indicates the A-F GPCR class and expected primary G-protein signal transduction type and cross-references the cognate cortically-expressed NPP genes. As noted above, the 18 NPP genes and 29 NP-GPCR genes listed in **Tables 1** and **2** were selected for focused analysis here due to their cognate pairing relationships and the consequent prospect that they may transmit local intracortical signals.

The "pFPKM Percentile" column in **Table 2** shows that most of these 29 NP-GPCR genes are expressed in cortex with Peak FPKM values well above median (50th percentile) for all protein coding genes. The high end of the range of cortical neuron pFPKM values for NP-GPCR genes does not match the extreme values noted for NPP genes, but this is as expected given that NP-GPCR gene products are thought to be durable cellular components, unlikely to be rapidly disposed by secretion as expected for NPP gene products. Peak FPKM values for NP-GPCR transcripts are nonetheless quite high in the range of transcripts of other likely durable cellular component genes, suggesting a strong likelihood that they are indeed translated into functionally important protein products.

	NP-GPCR Gene	Peak FPKM	pFPKM Percentile	% Cells	Neuropeptide Receptor	GPCR Class	G protein Class	Cognate NPP Genes
1	Sstr2	413	95.3	42	Somatostatin Receptor 2	A4	i	Sst, Cort
2	Npy2r	291	93.1	10	Neuropeptide Y Receptor Y2	A9	i	Npy
3	Npy1r	272	92.4	50	Neuropeptide Y Receptor Y1	A9	i	Npy
4	Grpr	231	91	10	GRP Receptor	A7	q	Grp
5	Cckbr	210	90	52	Cholecystokinin B Receptor	A6	q	Cck
6	Ntsr2	161	86.9	17	Neurotensin Receptor 2	A7	q	Nts
7	Npy5r	152	86.1	28	Neurpeptide Y Receptor Y5	A9	i	Npy
8	Nmbr	123	82.4	8	Neuromedin B Receptor	A7	q	Nmb
9	Rxfp1	121	82	22	Relaxin Family Receptor 1	A5	s	Rln1
10	Sstr4	106	79.5	28	Somatostatin Receptor 4	A4	i	Sst, Cort
11	Trhr	101	78.4	3	TRH Receptor	A7	q	Trh
12	Sstr1	90	76	38	Somatostatin Receptor 1	A4	i	Sst, Cort
13	Adcyap1r1	89	75.8	71	ADCYAP1 Receptor 1	B1	s	Adcyap1, Vip
14	Crhr1	86	74.9	28	CRH Receptor 1	B1	s	Crh
15	Rxfp3	85	74.7	5	Relaxin Family Receptor 3	A5	s	Rln1
16	Oprl1	82	73.8	48	Opioid Receptor-Like 1	A4	i	Pnoc
17	Crhr2	72	70.7	3	CRH Receptor 2	B1	s	Crh
18	Tacr3	65	68	3	Tachykinin Receptor 3	A9	q	Tac2
19	Oprk1	64	67.4	3	Kappa-Opioid Receptor	A4	i	Pdyn
20	Tacr1	56	64.2	3	Tachykinin Receptor 1	A9	q	Tac1
21	Pth1r	51	61.6	15	PTH 1 Receptor	B1	q	Pthlh
22	Vipr1	41	56.1	28	VIP Receptor 1	B1	s	Vip, Adcyap1
23	Oprm1	35	52.1	43	Mu-Opioid Receptor	A4	i	Penk, Pdyn
24	Trhr2	30	48.9	10	TRH Receptor 2	A7	q	Trh
25	Vipr2	30	48.4	0.5	VIP Receptor 2	B1	s	Vip, Adcyap1
26	Rxfp2	28	47.3	4	Relaxin Family Receptor 2	A5	s	Rln1
27	Oprd1	26	45.8	13	Delta-Opioid Receptor	A4	i	Penk, Pdyn
28	Ntsr1	24	44.3	10	Neurotensin Receptor 1	A7	q	Nts
29	Sstr3	17	39.5	21	Somatostatin Receptor 3	A4	i	Sst, Cort

Table 2. Genes encoding 29 NP-GPCRs expressed in mouse cortical areas VISp and ALM and cognate to the 18 neuropeptide precursor protein (NPP) genes listed in Table 1. Peak FPPM values, Peak FPKM percentile and absolute ranks and the percentage of cells sampled in which transcripts of the specified NP-GPCR gene were detected. In addition, the table names the encoded neuropeptide-selective GPCR protein, lists the GPCR A-F class (Alexander et al., 2017), the primary Gα signal transduction class associated with each predicted NP-GPCR as "i" for $G\alpha_{i/o}$ family, "s" for $G\alpha_s$ and "q" for $G\alpha_{q/11}$ family (Alexander et al., 2017), and indicates cognate NPP genes. The color fills of pastel red, green and blue in the "G protein Class" column correspond to i, s and q classes and will be used to highlight these classes consistently in all following figures. NP-GPCR genes are listed here in order of Peak FPKM values.

Expression of NP-GPCR genes by cortical neurons is highly differential. **Figure 2** represents expression patterns of the 29 NP-GPCR genes listed in **Table 2** in a manner that closely parallels the presentation for 18 NPP genes in **Fig.1**. **Figure 2A** establishes that each of the 29 NP-GPCR genes, like the 18 NPP genes, is expressed in highly differential fashion across the population of 22,439 mouse cortical neurons sampled. Each of 29 color-coded solid curves represents the distribution of single-neuron expression level values for one NP-GPCR gene. Curves were generated by plotting CPM for each individual neuron in descending order along a cell population percentile axis. As was noted for NPP genes in **Fig. 1**, each of the curves in **Fig. 2A** shows an abrupt transition from very high to very low (commonly zero) expression across the sampled neuron population. These transitions again occur at very different population percentile points, providing clear evidence for highly differential expression of NP-GPCR gene.

Percentages of the sampled neuron population expressing a given NP-GPCR gene (at greater than 1 CPM) range from more than 72% for Adcyap1r1 down to 0.7% for Vipr2.

Almost all (and possibly all) neocortical neurons express at least one NP-GPCR gene. The dashed curve in the left panel of **Fig. 2A**, labeled "Max NP-GPCR Gene", was generated by plotting CPM values of the NP-GPCR gene with the highest CPM in each individual cell in descending order along a cell population percentile axis. This curve shows that 98% percent of the sampled mouse cortical neurons express at least one NP-GPCR gene at >1 CPM and that 78% express at least one NP-GPCR gene at >100 CPM, lower than the comparable point for NPP genes (see **Fig. 1**) but still a very high value. Again, these numbers must be understood as lower limits to percentages of cortical neurons actually expressing at least one of the 29 NP-GPCR genes, after taking into account the pulsatile transcription and stochastic sampling issues cited above. The results summarized in **Fig. 2A** may thus be consistent with a conclusion that every cortical neuron expresses at least one NP-GPCR gene cognate to a cortically expressed NPP gene.

Statistics of differential NP-GPCR gene expression are highly conserved between different neocortical areas. **Figure 2B** provides evidence for strong conservation of differential NP-GPCR expression profiles between distant cortical areas VISp and ALM. The paired bars represent fractions of cells expressing a given gene in each of the two areas, again revealing strong similarities of differential expression profiles in the two very different neocortical areas (ρ=0.959, p<2.2E-16).

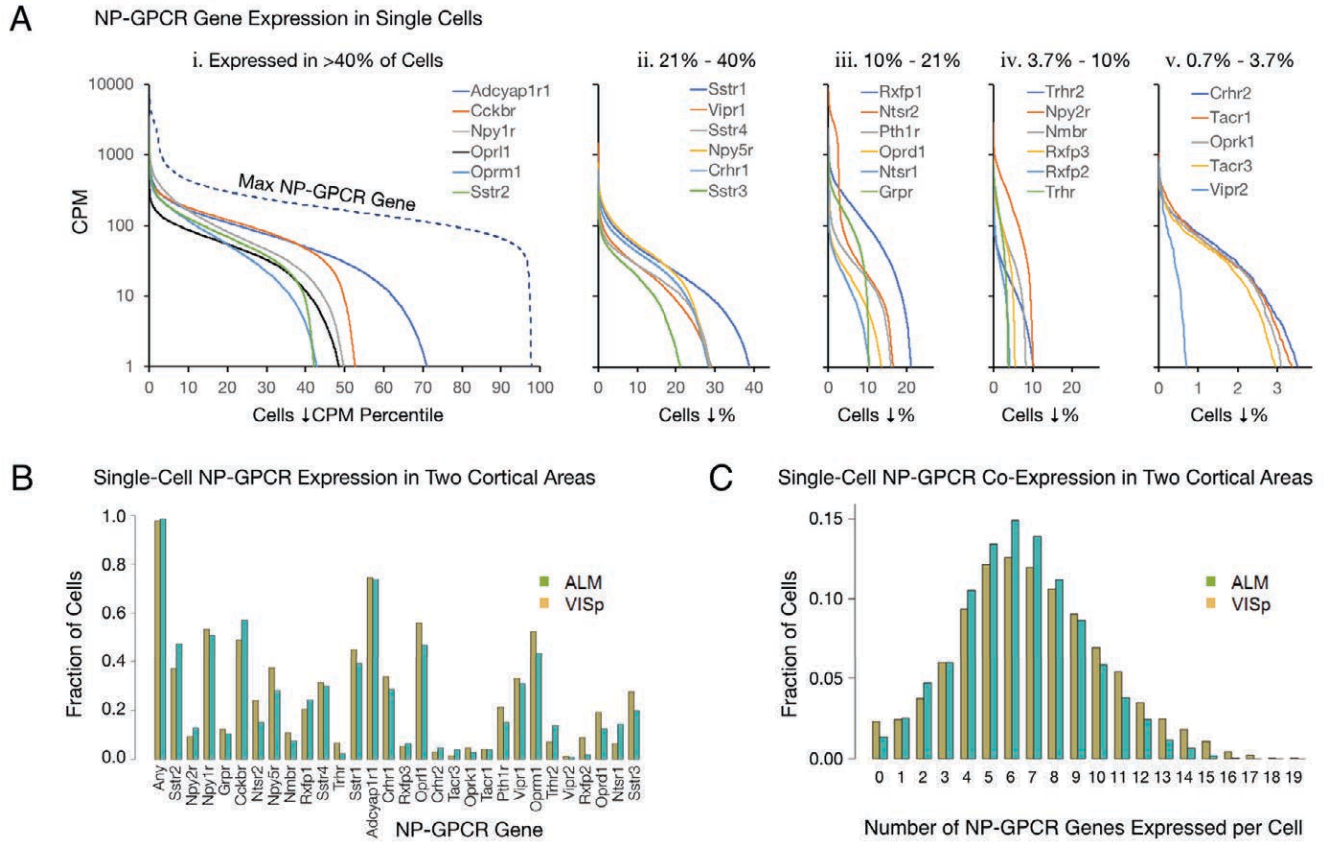


Figure 2. Expression and co-expression of NP-GPCR genes is highly differential with statistics conserved between two distant cortical areas. Single-cell expression analysis of 29 NP-GPCR genes in 22,439 individual neurons from cortical areas VISp and ALM. **(A)** Distributions of single-cell expression levels in neurons pooled from VISp and ALM. Color-coded solid curves plot single-cell CPM values for individual NP-GPCR genes in descending order along a cell population percentile axis. The 29 curves are segregated for clarity into five panels, sorted by cell population percentiles at which CPM values fall below 1. Large differences in fractions of cells expressing different NP-GPCR genes are evident. Dashed curve in panel **Ai** was generated by plotting CPM values of the highest CPM NP-GPCR gene for each individual cell in descending order. **(B)** Fractions of cells expressing each NP-GPCR genes represented separately for 13,491 VISp neurons and 8,948 ALM neurons, showing strong conservation between areas of the patterning of NP-GPCR expression fractions documented in panel A. **(C)**

Histograms illustrating frequencies of various multiples of NP-GPCR gene co-expression in individual neurons, represented separately for VISp and ALM neurons. The paired vertical bars illustrate strong conservation of co-expression patterns between the two cortical areas.

Multiple NP-GPCR genes are co-expressed in almost all cortical neurons. **Figure 2C** represents frequencies of NP-GPCR gene co-expression multiples detected in individual neurons. These data establish that multiple NP-GPCR genes are co-expressed in almost all cortical neurons and that numbers of genes co-expressed are even higher than those noted above for co-expression of NPP genes. Modal numbers of co-expressed NP-GPCR genes detected is 6 in both VISp and ALM with broad plateaus extending out to 12 co-expressed NP-GPCR genes per cell. The striking similarities of NP-GPCR co-expression distributions between the two otherwise divergent neocortical areas once again suggests that the patterning of NP-GPCR co-expression may have consequences for cortical function that are conserved because they are functionally important.

Transcriptomic neurotaxonomy enables the generation of testable predictions about neocortical neuropeptidergic signaling. Our analysis so far has relied solely upon the genomic depth and single-cell resolution characteristics of the 2018 Tasic transcriptomic data, without utilizing the transcriptomic neurotaxonomy derived as one major goal of that study (Tasic et al., 2018). This taxonomy was developed from a large body of single-cell mRNA-Seq data based on dimensionality reduction and iterative hierarchical clustering methods. Such a transcriptomic neurotaxonomy makes it possible to predict a protein "parts list" for any neuron that can be mapped to a given transcriptomic type. While additional work now under way (Cadwell et al., 2017; Daigle et al., 2018; Moffitt et al., 2016; Shah et al., 2016; Wang et al., 2018; Zeng and Sanes, 2017) will be needed to reconcile this transcriptomic neurotaxonomy to existing anatomical and physiological neurotaxonomies, this taxonomy already offers the prospect of genetic access to specific neuron classes and types for physiological and anatomical study and thereby the prospect of experimental test of transcriptomically generated hypotheses. The present analysis will make extensive use of a subset of the 2018 Tasic neurotaxonomy representing 115 types discriminated in VISp and ALM cortical areas, as summarized in **Supplementary Fig. 1**. This neurotaxonomy will be represented in the following figures by cladograms and/or color code strips that can be interpreted by reference to **Supplementary Fig. 1** or (Tasic et al., 2018).

Expression of the 18 NPP genes is highly neuron-type-specific. **Figure 3A** represents expression levels of the 18 NPP genes across all 115 VISp+ALM neuron types as a "heat map" matrix color coding $\log_{10}$ CPM values for each NPP and each neuron type. The CPM values so rendered are calculated as "trimmed means" (mean value after discarding the top 1% of distributions to reject outliers) of single-cell CPM values aggregated by each neuron-type cluster (commonly on the order of 100 cells, see **Supplementary Figs. 1A** and **1C** for actual cell counts). **Figure 3A** confirms and extends four reasonable expectations from the type-agnostic single-cell analyses of **Figs. 1 and 2** above: (1) neurons of every type express one or more of the 18 NPP genes, (2) each of the 18 NPP genes is expressed in multiple neuron types, (3) neurons of every type express multiple NPP genes, and (4) expression of NPP genes is highly differential across neuron types. Remarkably, **Fig. 3A** shows that type-to-type variations in expression level for every one of the 18 NPP genes span the full >10,000-fold dynamic range characteristic of the Tasic 2018 RNA-Seq data. Quite intriguingly, **Fig. 3A** also suggests that each of the 115 VISp+ALM cell types might be distinguished by a unique combinatorial pattern of NPP gene expression. This possibility will be explored quantitatively in connection with **Fig. 4** below.

Figure 3A provides for ready comparison of NPP gene expression patterns between glutamatergic and GABAergic neuron types. Clearly, GABAergic types are more prolific in the variety and strength of their NPP genes expression. While glutamatergic types express fewer NPP genes and do not match the extremely high NPP expression levels observed in almost every GABAergic type, each type nonetheless expresses at least one NPP gene, and generally more, at a very substantial level. This differential is consistent with a long history of neuroscientific use of neuropeptide products as protein markers of GABAergic neuron subsets (e.g., VIP, SST, NPY, Substance P), which has no parallel in the marking of glutamatergic neuron subsets.

Expression of the 29 NP-GPCR genes is highly neuron-type-specific. **Figure 3B** illustrates neuron type specificity of NP-GPCR expression in a manner identical to the treatment of NPP gene expression in **Fig. 3A** and invites analogous conclusions: (1) neurons of every neuron type express one or more of the 29 NP-GPCR genes at very high levels, (2) neurons of every type express multiple NP-GPCR genes, and (3) expression of NP-GPCR genes is highly differential across neuron types. **Figure 3B** also shows, however, that the stronger and more varied expression of NPP genes in GABAergic expression profiles that was evident in **Fig. 3A** is leveled or even reversed for NP-GPCR genes. That is, while GABAergic neurons clearly show the more prolific and varied expression of NPP genes, glutamatergic neurons may be somewhat more prolific expressors of NP-GPCR genes. Finally, it should be noted that there are cases where both an NPP gene and its cognate NP-GPCR receptor are expressed in the same neuron type, with the Cck / Cckbr and Adcyap1 / Adcyap1r1 pairs being prominent examples, with both being highly expressed in majorities of glutamatergic neuron types.

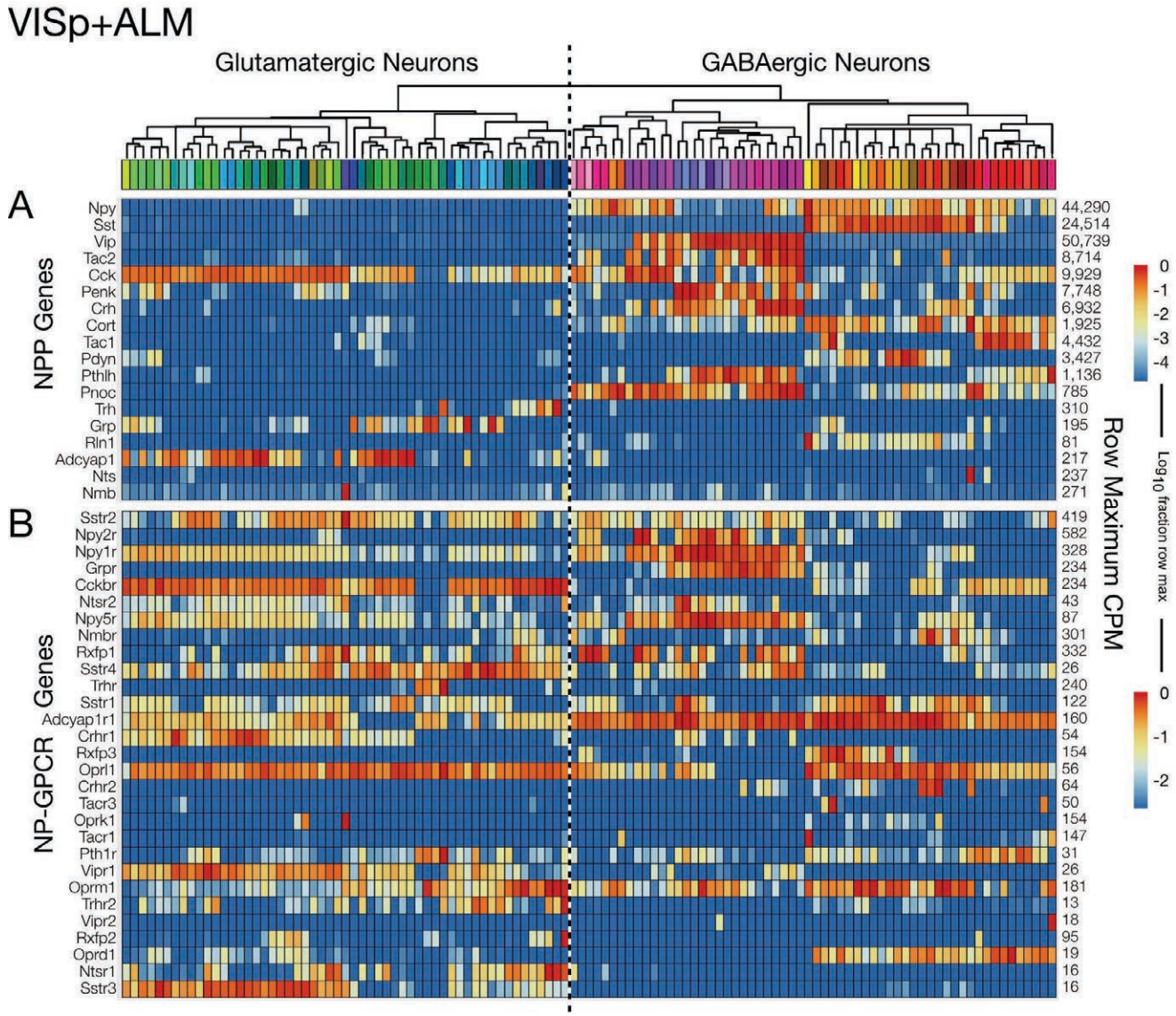


Figure 3. Profiles across 115 transcriptomic neuron types reveal very high neuron-type-specificity of NP gene expression. Heat maps representing CPM expression values of 18 NPP and 29 NP-GPCR genes (rows) aggregated for each of 115 neuron types (columns) distinguished by the resource neurotaxonomy (Tasic et al., 2018). Note that every gene spans a very wide range of expression values across all neuron types. **(A)** Expression profiles of 18 NPP genes, ordered as in Table 1. **(B)** Expression profiles of 29 NP-GPCR genes for all VISp and ALM neuron types, with NP-GPCR genes ordered as in **Table 2.** Color mapping (separate keys at right for panels **A** and **B**) encode $\log_{10}$ mean of CPM values for each gene and cell type after trimming away top 1% percent CPM values to suppress outlier effects. Expression values are normalized separately for each gene (row), with numbers at right representing row maximum trimmed mean CPM values.

Neurotaxonomic cladogram and color scheme at top of this map are those introduced in the 2018 Tasic, et al., publication and summarized in the present **Supplementary Fig. 2**.

A transcriptomic signature based upon just 47 NP-signaling genes (18 NPPs and 29 NP-GPCRs) permits exceptionally accurate classification of neocortical neurons. The strong marker patterning of the 47 NP gene expression profiles evident in **Fig. 3** suggests the possibility that each of the 115 neuron types profiled in that figure might be distinguished by a unique combination of these 18 NPP and 29 NP-GPCR genes. To explore this possibility quantitatively, we developed the analysis presented in **Figure 4**.

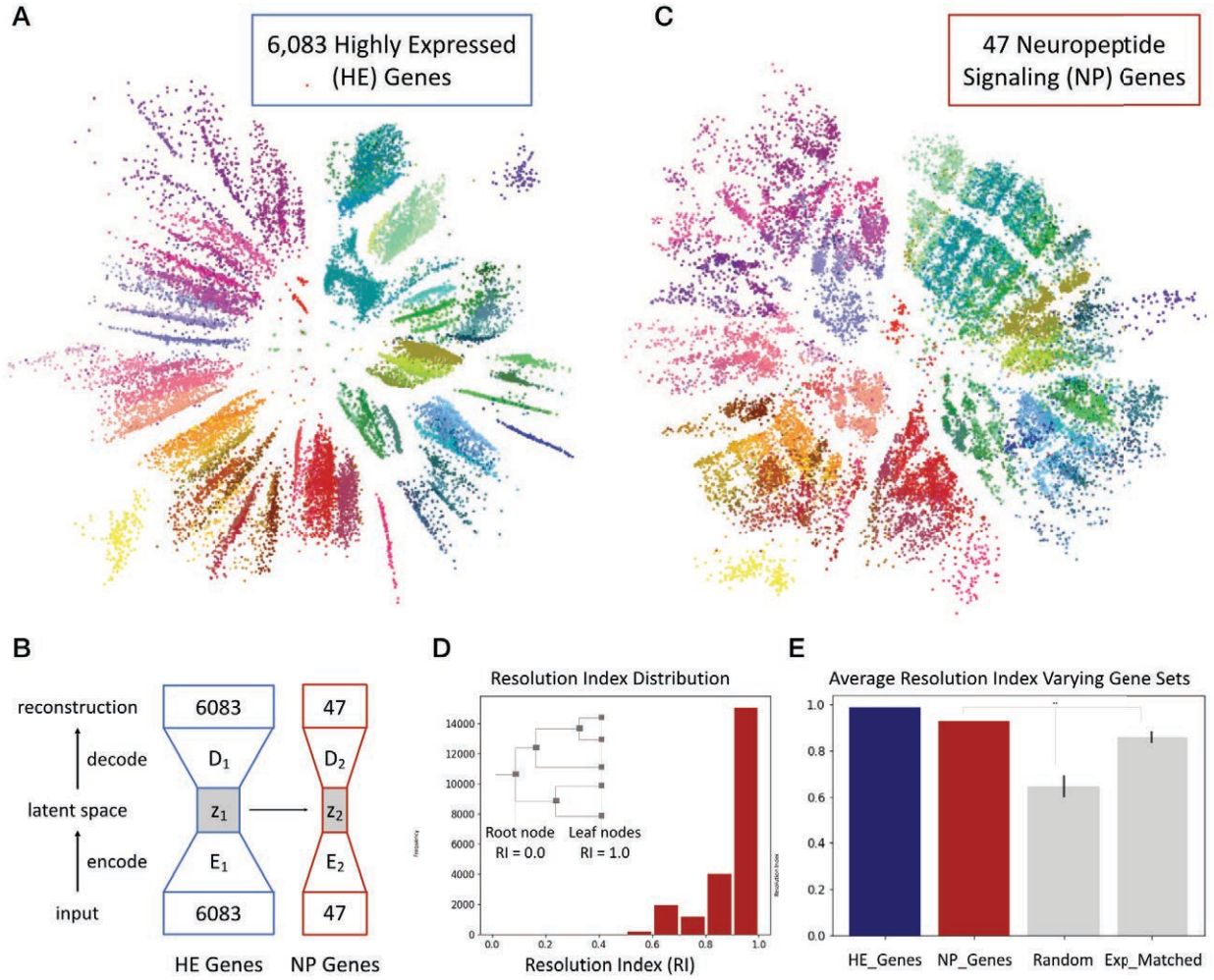


Figure 4. Neurons can be classified effectively based on just 47 NP genes (18 NPP and 29 NP-GPCR) alone. (A) A two-dimensional latent space representation of autoencoder network neuron type classification of 22,439 cells based on 6,083 highly expressed (HE) genes, where each dot represents a single cell colored according to the Tasic (Tasic et al., 2018) type-code coloring scheme, as introduced here in **Fig. 3** and **Supplementary Fig. 2**. The clustering of type-code colors to form distinct islands is indicative of good clustering performance. **(B)** Schematic representation of the linked autoencoder architecture used to optimize a second network to classify neurons as similarly as possible into a latent space based on much smaller gene sets. **(C)** Representation of latent space z_2 **(D)** A Gaussian mixture model of cell types is fit to a five-dimensional NP gene representation using 2018 classification. Progressively simplified taxonomies are obtained by iterative merging, and the most complex taxonomy for which each cell is correctly mapped is recorded as that cell's Resolution Index (RI). This histogram represents the distribution of RI scores corresponding to classification of the five-dimensional NP gene latent space representation. **(E)** Average RI scores for latent space representations encoded by various gene subsets. The average RI per cell obtained from the NP genes representation (0.928) was significantly higher than the average obtained from 100 distinct sets of 47 random genes (0.645), and the average of 100 sets of 47 genes that matched expression levels of the 47 NP genes (0.858), $p<0.01$. See **Supplementary Methods** for additional details of autoencoder development and performance metrics.

We began by asking whether there exists a low dimensional representation of gene expression that naturally separates neurons of different types into distinct parts of that low-dimensional space. The extent to which a neuron's location in such a space can be inferred from the expression of a limited subset of genes (such as our 47 NP genes) would then provide a measure of the sufficiency of that subset to classify a that neuron accurately. Hierarchical clustering methods to define neuron types based upon gene

expression are well established (Hastie et al., 2001; Oyelade et al., 2016) but have difficulty when comparing and making inferences between datasets. We therefore devised a machine learning approach based on linked multi-layer autoencoders (see **Supplementary Methods**) to address this question explicitly and quantitatively.

A single autoencoder network (Hinton and Salakhutdinov, 2006) was developed and trained to encode CPM values of the 6,083 most highly expressed genes represented in the Tasic 2018 dataset (the "HE" gene set). Results are illustrated in **Figure 4A**, where encoding coordinates in a two-dimensional latent space of 22,439 individual neurons are displayed as discrete dots, each colored according to the neuron's Tasic 2018 hierarchical classification (i.e., the neurotaxonomic color code introduced in **Fig. 3** and **Supplementary Fig. 2**). The tight grouping of type-code colors evident in **Fig. 4A** implicitly represents that position within this latest space corresponds well to the neuron types defined by hierarchical classification, in spite of the fact that the autoencoder was given no explicit prior information about how neurons were classified by Tasic, et al. We then trained a second, linked auto-encoder with the architecture schematized in **Fig. 4B** to classify cells using only the 47-gene subset, under a cost function constraint that latent spaces of the two auto-encoders be as similar as possible. This allowed us to test the extent to which any small gene subset by itself could match the encoding performance obtained using the much larger gene set. **Fig. 4C** displays a two-dimensional latent space resulting from encoding the same 22,439 neurons based only on 47 NP genes tabulated above, again projecting one dot for each cell using the Tasic 2018 type-code colors. Once again, the tight color grouping evident in **Fig. 4C** suggests qualitatively that these 47 genes indeed enable excellent type encoding of individual neurons.

For quantitative comparison of classification performance based on varied neocortical gene sets, we partitioned the autoencoder encodings into classes using a supervised Gaussian Mixture Model (see **Supplementary Methods**) and designed the resolution index schematized in **Fig. 4D** to evaluate consensus between classifications driven by autoencoder encoding with the resource hierarchical neurotaxonomy (Tasic et al., 2018). This index yields a value of 0 when a neuron is mapped incorrectly from the root node and 1 when a neuron is mapped correctly all the way to a terminal leaf node. By averaging this metric over all 22,439 neurons, we generated an overall figure of merit called a resolution index. This figure for the large HE 6083 gene set was 0.987, the same index for classification based on the NP 47-gene subset was 0.928. To place these resolution index numbers in context and test the significance of this correspondence, we compared resolution indices resulting from linked autoencoder classification based on 100 subsets of 47 genes drawn randomly from the Tasic 2018 expression dataset. The sets of 47 random genes yielded an average resolution index of 0.645 ± 0.047 (Fig. 4E), establishing clearly that NP genes yield classification greatly superior to random subsets of 47 genes. **Figure 4E** also shows results from encoding runs using 100 sets of 47 random genes chosen to approximate the same high expression statistics of the NP genes. Again, resolution indices from the random sets fell well below that yielded by the 47 NP genes (average = 0.858 ± 0.0242, with none reaching the NP gene index of 0.928 and the difference being significant at $p<0.01$). This demonstration of the exceptional power of NP genes to mark transcriptomic neuron types reinforces earlier indications of an especially close and fundamental connection between neuropeptide gene expression and neuron type identity.

Cell-type-specificity of differential NP gene expression is conserved between neocortical areas. **Figure 5** juxtaposes separate VISp and ALM expression profiles for NPP and NP-GPCR genes across 93 VISp neuron types (**Fig. 5A**) and 84 ALM neuron types (**Fig. 5B**). Similarities of expression profiles for the two areas are obvious in **Fig. 5**, but there are also visible differences. The latter are rooted primarily in the substantial divergence of glutamatergic neuron taxonomies discussed at length in Tasic, et al. (Tasic et al., 2018) and summarized here in **Supplementary Fig. 3**. Very strong similarities of both NPP and NP-GPCR expression profiles are most obvious for the GABAergic types, where the taxonomies are identical except for the absence of two GABAergic types in ALM (indicated by dark gray vertical placeholder bars in **Fig. 5B**). The general conservation of neuron-type-specific expression patterns between the two distant neocortical areas (NPP correlation: $\rho= 0.974$, $p<2.2e-16$, NP-GPCR: 0.877, $p<2.2e-16$) thus provides another indication of robust connection between NP gene expression and cortical neuron differentiation.

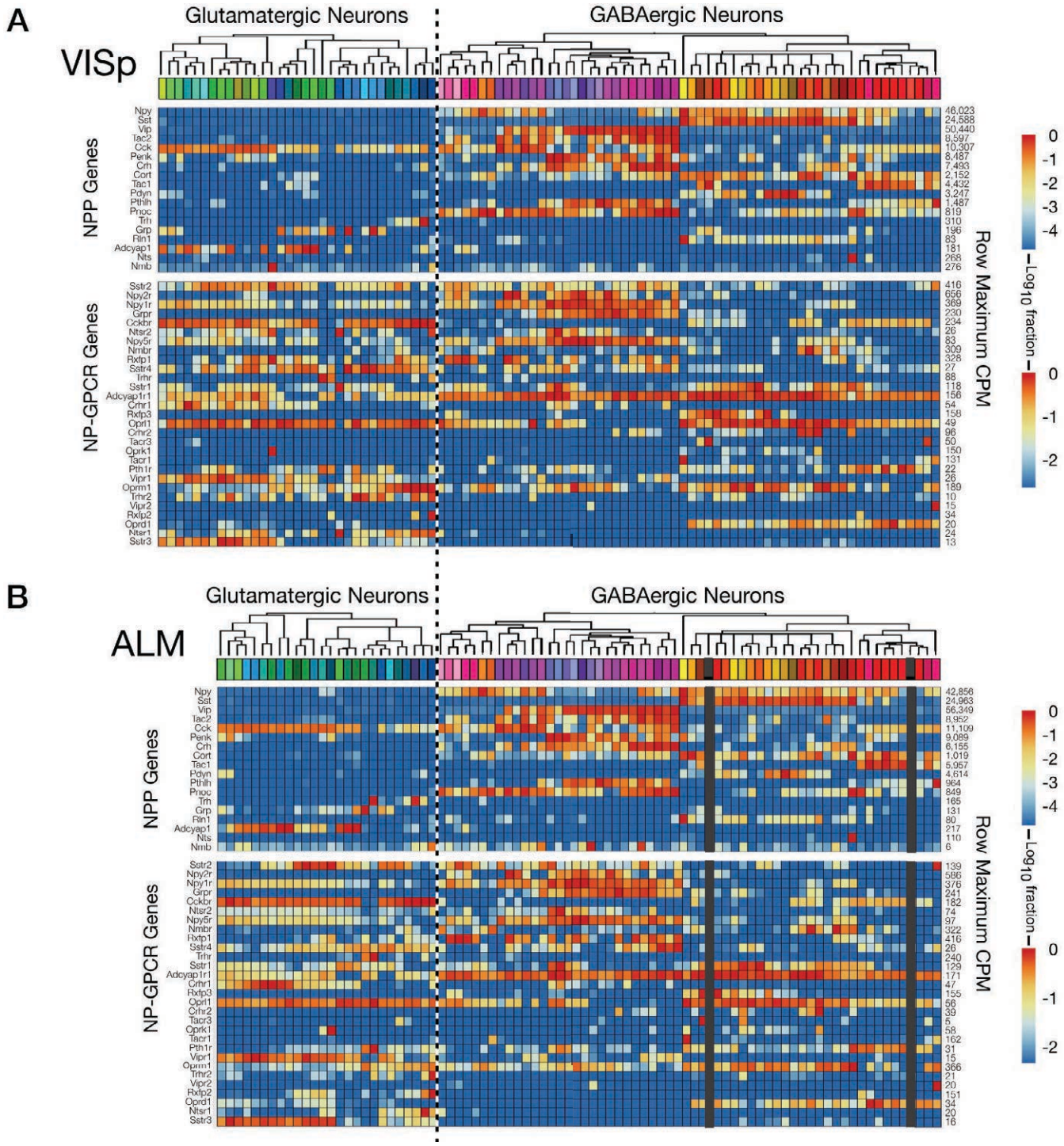


Figure 5. Expression profiles of NP genes restricted to cortical area VISp or to area ALM are very similar but differ somewhat due to divergence of VISp and ALM neurotaxonomies. Neuron-type-based NPP and NP-GPCR expression heat maps similar to those of **Fig. 3** but separating data from area VISp and ALM samples. **(A)** Expression profiles for 18 NPP and 29 NP-GPCR genes in area VISp, formulated, color coded and ordered as in Fig. 3. **(B)** Expression profiles for 18 NPP and 29 NP-GPCR genes in area ALM, again displayed as in **Fig. 3**. Heat maps are placed here to vertically align GABAergic neuron types that match between VISp and ALM areas. Vertical dark gray bars in **Fig. 5B** are spacers marking the two GABAergic cell types absent in ALM. Glutamatergic taxonomies differ more substantially (see resource publication (Tasic et al., 2018) and **Suppl. Fig. 1**).

Expression of 37 cognate NPP/NP-GPCR pairs in cortex predicts the potential existence of 37 intracortical peptidergic networks. Expression of an NPP gene in one neuron and a cognate NP-GPCR gene in another nearby neuron implies the prospect of local paracrine signaling, with secretion of a specific neuropeptide by the first neuron activating the cognate specific neuropeptide receptor on a

second, nearby neuron. The present set of 47 cortical NP genes (18 NPP and 29 NP-GPCR) comprises the 37 distinct cognate NPP/NP-GPCR pairs enumerated in **Table 3** and predicts accordingly 37 distinct peptidergic neuromodulation networks. As noted in the **Introduction**, expected neuropeptide diffusion distances suggest that any neuron within a local cortical area (e.g., VISp or ALM) might signal by diffusion to any other neuron within that same local area, but almost surely not to more distant areas (e.g., from VISp to ALM). In the following, we therefore make predictions of 74 (37 x 2) peptidergic distinct signaling networks, keeping separate consideration of signaling within VISp and within ALM.

#	Cognate Pair Symbol	NPP Gene	NP-GPCR Gene	G protein Class	Fraction of Pairs
1	Npy→Npy1r	Npy	Npy1r	i	0.263
2	Npy→Npy2r	Npy	Npy2r	i	0.057
3	Npy→Npy5r	Npy	Npy5r	i	0.159
4	Sst→Sstr1	Sst	Sstr1	i	0.151
5	Sst→Sstr2	Sst	Sstr2	i	0.158
6	Sst→Sstr3	Sst	Sstr3	i	0.084
7	Sst→Sstr4	Sst	Sstr4	i	0.111
8	Penk→Oprd1	Penk	Oprd1	i	0.053
9	Penk→Oprm1	Penk	Oprm1	i	0.163
10	Cort→Sstr1	Cort	Sstr1	i	0.142
11	Cort→Sstr2	Cort	Sstr2	i	0.148
12	Cort→Sstr3	Cort	Sstr3	i	0.079
13	Cort→Sstr4	Cort	Sstr4	i	0.105
14	Pdyn→Oprd1	Pdyn	Oprd1	i	0.016
15	Pdyn→Oprk1	Pdyn	Oprk1	i	0.004
16	Pdyn→Oprm1	Pdyn	Oprm1	i	0.049
17	Pnoc→Oprl1	Pnoc	Oprl1	i	0.135
18	Rln1→Rxfp3	Rln1	Rxfp3	i	0.004
19	Vip→Vipr1	Vip	Vipr1	s	0.14
20	Vip→Vipr2	Vip	Vipr2	s	0.004
21	Crh→Crhr1	Crh	Crhr1	s	0.075
22	Crh→Crhr2	Crh	Crhr2	s	0.01
23	Rln1→Rxfp1	Rln1	Rxfp1	s	0.015
24	Rln1→Rxfp2	Rln1	Rxfp2	s	0.003
25	Adcyap1→Adcyap1r1	Adcyap1	Adcyap1r1	s	0.196
26	Adcyap1→Vipr1	Adcyap1	Vipr1	s	0.084
27	Adcyap1→Vipr2	Adcyap1	Vipr2	s	0.002
28	Tac2→Tacr3	Tac2	Tacr3	q	0.006
29	Cck→Cckbr	Cck	Cckbr	q	0.414
30	Tac1→Tacr1	Tac1	Tacr1	q	0.006
31	Pthlh→Pth1r	Pthlh	Pth1r	q	0.035
32	Trh→Trhr	Trh	Trhr	q	0.002
33	Trh→Trhr2	Trh	Trhr2	q	0.004
34	Grp→Grpr	Grp	Grpr	q	0.014
35	Nts→Ntsr1	Nts	Ntsr1	q	0.001
36	Nts→Ntsr2	Nts	Ntsr2	q	0.002
37	Nmb→Nmbr	Nmb	Nmbr	q	0.014

Table 3. The 18 NPP and 29 NP-GPCR genes of Tables 1 and 2 comprise 37 cognate NPP/NP-GPCR pairs and predict at least 37 potentially distinct peptidergic modulatory networks. The 37 pairs are enumerated here along with indications of the expected primary GPCR signal transduction class for each NP-GPCR (Alexander et al., 2017) and an indication of the gene expression prevalence for each cognate pair as a fraction of the total number of neuron pairs surveyed. Table fill colors echo G-protein class as in Table 2.

Type-specific NP gene expression profiles predict type-specific peptidergic coupling. **Figure 6** displays heat maps representing predictions of neuron-type-specific peptidergic coupling from a selection of the 37 cognate NP gene pairs and expression profiles of the paired NPP and NP-GPCR genes. The predictions of **Fig. 6** are based on cell-type-by-cell-type aggregation of binarized cell-pair-by-cell-pair products of the NPP and NP-GPCR gene CPM values. CPM values were first thresholded at the 50th percentile independently for each cell type. The coupling matrix is then defined as:

$$E(p,q) = \frac{1}{|C(p)||C(q)|} \sum_{j=1}^{|C(p)||C(q)|} I(e^j{}_p > t_p) * I(e^j{}_q > t_q) \qquad (1)$$

where e^j_p denotes the expression of individual cell j in cell type p, |C(p)| denotes the total number of expressing cells of type p, and I is the indicator function and is therefore the fraction of expressing pairs exceeding the 50th percentile threshold.

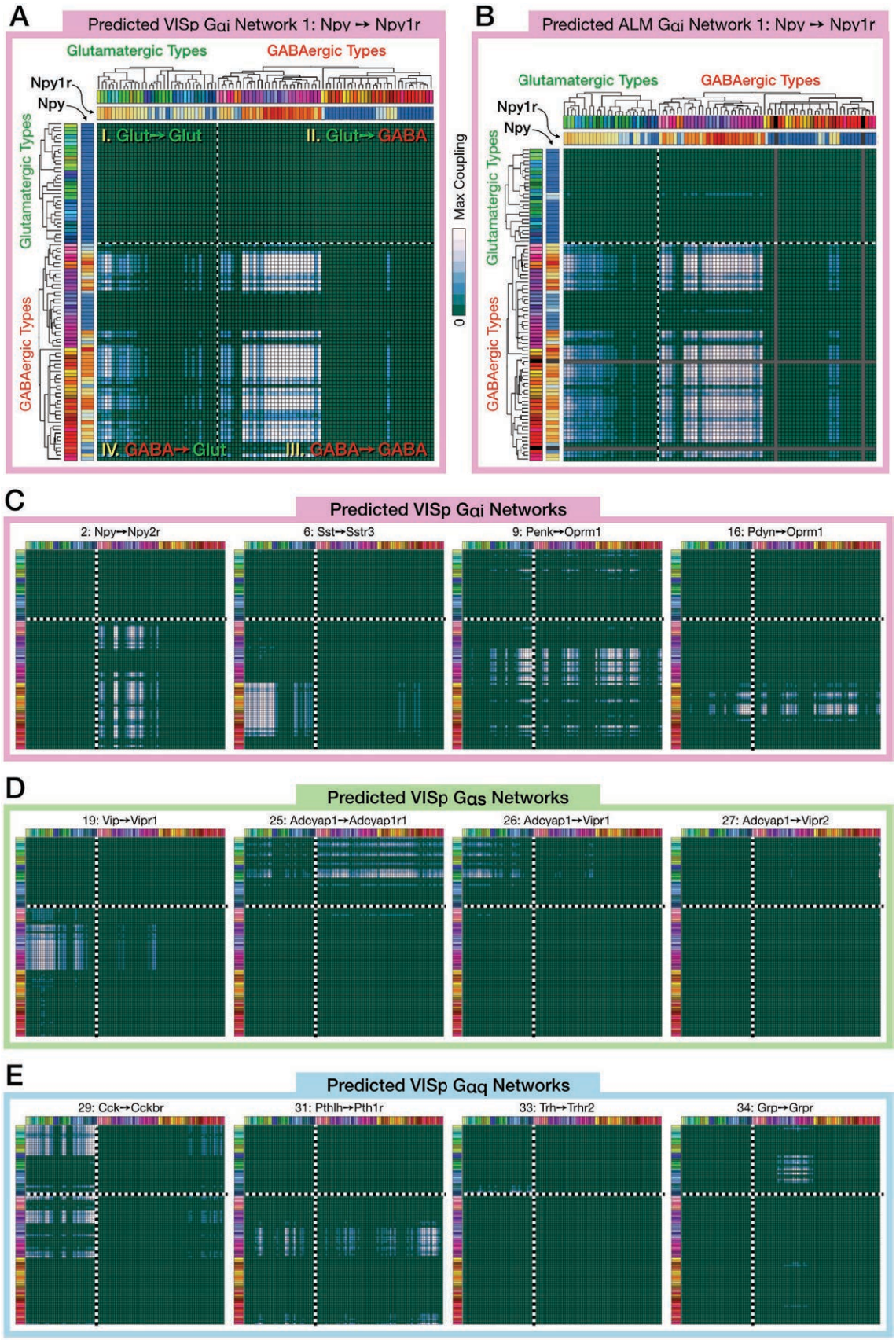


Figure 6. Expression profiles of conjugate NPP/NP-GPCR pairs predict neuropeptidergic signaling networks. Square matrices representing predictions of local peptidergic coupling between specific neuron types based on the resource data tables and neurotaxonomy (Tasic et al., 2018). Coupling values predicted from Eqn. 1, normalized to maximum values within each matrix, are plotted in all panels according to the color scale located between panels **A** and **B**. **(A)** A 93 x 93 matrix predicting coupling amongst the 93 VISp neuron types based on type-specific expression of the NPP gene Npy and the

cognate NP-GPCR gene Npy1r. Color strip nearest the left edge of the matrix is the expression profile for Npy (identical to the row so labeled in the expression heat map of **Fig. 5A**), while the color strip immediately to its left represents the 93 VISp neuron types according to the Tasic, et al., color scheme as summarized in **Suppl. Fig. 1**. The color strip nearest the top edge of the coupling matrix is the expression profile for the Npy1r NP-GPCR gene (identical to row so labeled in **Fig. 5B**), with the adjacent color strip again representing the 93 VISp neuron types. **(B)** An 84 x 84 square matrix representing Npy-Npy1r coupling as in panel **A**, except based on data from area ALM and thus representing only the 84 neuron types distinguished by Tasic, et al., in that area. Black bars within the type-code color strips and gray bars in heat map of panel **B** are spacers in place of the two VISp GABAergic neuron types not found in area ALM, there to preserve alignment of the 58 remaining neuron types common to the two areas. Glutamatergic type alignments differ more substantially (see Tasic, et al. (Tasic et al., 2018) and **Suppl. Fig. 3**). Dashed lines crossing on each plot (and all plots in this figure) demarcate the four quadrants of possible directed NPP/NP-GPCR pairings between glutamatergic and GABAergic neuron types: (*Glut*→*Glut*), (*Glut*→*GABA*), (*GABA*→*GABA*) and (*GABA*→*Glut*), as labeled in panel A. **(C-E)** Exemplary coupling matrix predictions for three additional sets of cognate NPP/NP-GPCR pairs, with the four plotted in **(C)** being representative of coupling to $G_{\alpha i}$-coupled NP-GPCRs, the four in **(D)** representing $G_{\alpha s}$-coupled NP-GPCRs and the four in **(E)** representing $G_{\alpha q}$-coupled NP-GPCRs. All are drawn from area VISp data and rendered as in panel **A**, except without copying the expression profile strips as in **Fig. 5**.

The exemplar matrix displayed in **Fig. 6A** predicts coupling in area VISp based on the expression profiles of Npy and Npy1r in VISp. **Figure 6B** represents a similar prediction for the same pair in area ALM. The dashed white crosses overlying both plots partition the matrices based on pairings of glutamatergic and GABAergic neuron types. Both matrices predict strong signaling from the canonical broad class of Npy-positive GABAergic neurons to a broad subset of GABAergic neurons that strongly express the Npy1r NP-GPCR: the strongest coupling thus falls in the GABA→ GABA quadrant. Weaker coupling is observed in the *GABA*→ *Glut* quadrant, where the Npy1r NP-GPCR gene is less strongly expressed in the Glutamatergic cell types. Strong similarities between the VISp and ALM coupling matrices are most notable. Apparent differences between VISp and ALM coupling predictions are mainly due to exclusive expression of different glutamatergic cell types in the two areas, and only in small part due to difference in same-type expression within the two cortical areas.

Figures 6C-E represent 12 more of the 37 cognate pair coupling matrices predicted for VISp using Eqn. 1. Along with Figs**. 6A** and **6B**, these exemplify the wide variety of neuron-type-specific coupling motifs resulting from transcriptomic prediction. Most coupling matrices (i.e., pairs 2, 6, 9, 16, 19, 25, 29, 31), predict significant coupling over wide swaths of type-pairs, approaching 20% of the entire matrix. A few matrices at the other extreme, such as 27 and 33, predict very sparse coupling. Other predictions are intermediate in sparsity. The full sets of 37 predicted coupling matrices enumerated in **Table 3** for both VISp and ALM are represented in **Suppl. Figure 2**, where strong similarities between the two cortical areas are again quite obvious.

Figure 6 and **Suppl. Fig. 2** also illustrate the tendency of coupling predictions from most cognate NP pairs to fall in contiguous "patches" of the full coupling matrix. This is a natural reflection of the strong tendency of both NPP and NP-GPCR expression to align with early nodes in the 2018 Tasic hierarchical clustering which was also evident in **Figs. 3** and **5**. The broadest example of coupling matrix patches reflecting hierarchical neurotaxonomy structure is provided by the observation of that most sizable coupling patches fall strictly within single quadrants of glutamatergic-GABAergic neuron type pairing.

Discussion

Light from single-cell transcriptomics is now beginning to illuminate dark corners of cellular neuroscience that have long resisted mechanistic and functional analysis (Fan et al., 2018; Földy et al., 2016; Gokce et al., 2016; Okaty et al., 2011; Paul et al., 2017; Shekhar et al., 2016; Tasic et al., 2018, 2016; Telley et al., 2016; Zeng and Sanes, 2017). Cortical neuropeptide signaling may be one such corner. While profound impacts of neuropeptide signaling are well-established in a wide range of non-mammalian and sub-cortical neural structures (Borbély et al., 2013; Burbach, 2011; Elphick et al., 2018; Grimmelikhuijzen and Hauser, 2012; Katz and Lillvis, 2014; Kuffler et al., 1979) and there certainly is an excellent literature on cortical neuropeptide signaling (Crawley, 1985; Férézou et al., 2007; Gallopin et al., 2006; Gomtsian et al., 2018; Hamilton et al., 2013; Liu et al., 2018; Mena et al., 2013; Rossier and Chapouthier, 1982; Williams and Zieglgänsberger, 1981), published physiological results are surprisingly rare given the breadth of neuroscientific interest in cortex. The new transcriptomic data analyzed here suggest a possible

explanation for this relative rarity. Though many NPP and cognate NP-GPCR genes are expressed abundantly in all or very nearly all neocortical neurons, such expression is highly differential, highly cell-type specific, and often redundant. These previously uncharted differential expression factors may have hindered repeatable experimentation. Our analysis supports this unwelcome proposition but may also point the way to more productive new perspectives on intracortical peptidergic neuromodulation.

Summary of findings. The present analysis establishes that mRNA transcripts from one or more of 18 NPP genes are detectible in over 97% of mouse neocortical neurons and that transcripts of one or more of 29 NP-GPCR genes are detectible in over 98%. Transcripts of at least one of the 18 NPP genes are present in the vast majority of cortical neurons at extremely high copy number, strongly suggesting brisk translation into neuropeptide precursor proteins. Brisk synthesis of precursor proteins further suggests brisk processing to active neuropeptide products and secretion of these products. Likewise, NP-GPCR transcripts rank high in abundance compared to transcripts of other cellular proteins, again strongly supporting product functionality. Our observations thus support the proposition that all, or very nearly all, neocortical neurons are both neuropeptidergic and modulated by neuropeptides. We are not aware of any previous empirical support for such a conclusion.

We have closely examined single-neuron expression patterns of sets of 47 NP genes (18 NPP and cognate 29 NP-GPCR) and find that these patterns are highly conserved between two distant and generally quite different areas of neocortex. Such conservation lends additional support to the proposition that NP gene products may have a very fundamental importance to cortical local circuit function and argues against these patterns reflecting more ephemeral variable such as recent activity patterns, which would seem unlikely to correlate so strongly between cortical areas with such different roles in brain function.

Following earlier indications that neurons may express multiple NPP genes, e.g., (Mezey et al., 1999), our analysis establishes that expression of multiple NPP genes in individual neurons may be the rule in cortex. Our analysis also establishes the generality of expression of multiple NP-GPCR genes in individual cortical neurons. The significance of these observations remains to be explored but should be viewed in light of recent discoveries of large numbers and great diversity of transcriptomic neuron types in neocortex and many other brain regions. Combinatorial expression of neuropeptide precursor and receptor genes obviously expands the prospects for molecular multiplexing that may allow selective communication amongst a multiplicity of distinct neuron types even though the signaling molecules propagate in diffuse paracrine fashion.

We also find that a modest set of 47 neuropeptide-signaling genes permits transcriptomic neuron type classification that is exceptionally precise in comparison to other similarly small gene sets. This tight alignment of neuron type classifications based solely on neuropeptide-signaling gene expression with classifications based on genome-wide expression patterns offers an intriguing suggestion of a very deep and fundamental connection between the expression of evolutionarily ancient neuropeptide-signaling genes and the differentiation of neuron type identities during metazoan speciation.

Prediction of cortical modulation networks. Our analysis delineates neuron-type-specific expression of 37 cognate pairs amongst the 18 NPP and 29 NP-GPCR genes expressed in mouse neocortex. Each of these pairs can be taken to predict a modulatory connection from cells expressing a particular NPP gene, via a secreted NP product, to cells expressing the particular NP-GPCR gene. Each pair thus establishes the prospect of a modulatory network with nodes defined by the neurotaxonomic identities of the transmitting NPP-expressing and the receiving NP-GPCR-expressing neurons. The analyses represented in **Figs. 1, 2, 3** and **5** and **Table 3** establish that at least one of the 37 pairs directly involves every neuron sampled, and that the vast majority of neurons are directly involved in more than one of the 37 predicted networks. Because of this saturated, multiplexed coverage of all neurons and neuron types, we refer to these predicted neuropeptidergic networks as “dense”.

The logic of our prediction of multiplexed NP networks is summarized in the form of a simplified schematic by **Figure 7A-E. Figure 7** also suggests how multiple NP networks may align with neuron-type-based predictions of synaptic network architectures as the relevant empirical connectomic and neurotaxonomic

information, as schematized by (**Fig. 7F**) becomes available. **Figure 7G** integrates the fictitious network graphs of **Figs. 7E** and **7G**, articulating a schematic view of cortical circuitry as the superimposition of many and diverse modulatory and synaptic networks, with neuron types as common nodes uniting a heterogeneous multiplicity of slow and fast signaling networks.

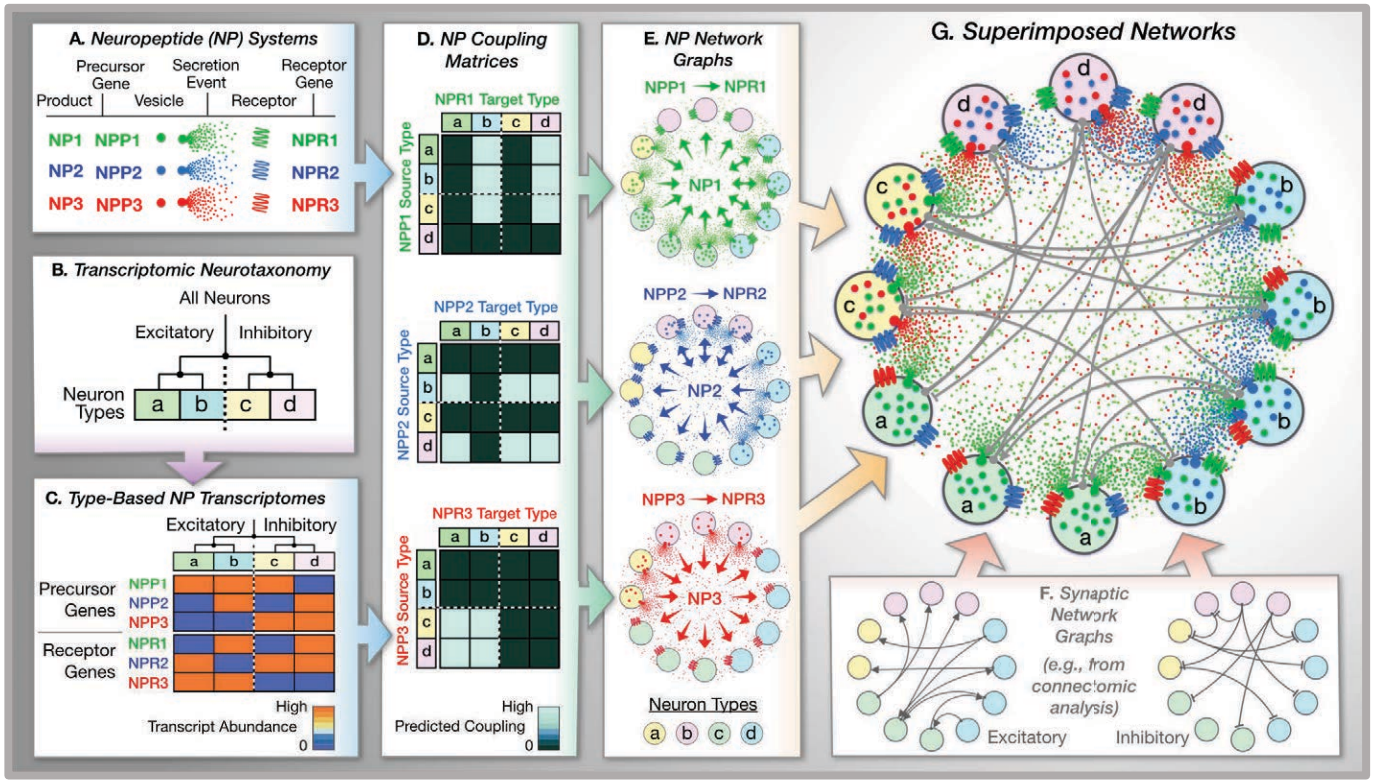


Figure 7. Schema for transcriptomic prediction of neuropeptide networks. (A) Three fictitious, conceptual neuropeptide systems (NP1, NP2 and NP3, color-coded red, green and blue) stand in here for the much larger numbers of actual peptidergic signaling elements listed in **Tables 1, 2** and **3** above. Biological literature cited in main text predicts paracrine intercellular signaling from neurons secreting products (NPx) of precursor genes (NPPx) to neurons expressing the receptor gene (NPRx). Paracrine diffusion of NP products is represented in this figure by scattered red, green and blue dots. **(B)** Four fictitious neuron types ("a" and "b" excitatory, "c" and "d" inhibitory, as coded by pastel shades) stand in here for the scores of transcriptomic cell types, some glutamatergic, some GABAergic, discerned by Tasic, et al., as tabulated in **Suppl. Fig. 1**. **(C)** A gene expression "heat" map, much smaller but otherwise analogous to those portrayed **Figs. 3** and **5** above, represents neuron-type-specific mRNA levels in the fictitious 3 NPPx and 3 NPRx genes of panel **A**. **(D)** Three neuron-type-based neuropeptidergic coupling matrices analogous to the 37 NP predicted coupling matrices represented in **Fig. 6** and **Suppl. Fig. 2,** derived here from the schematic heat map of panel **C**. **(E)** Schematic NP network coupling graphs cast from the NP coupling matrices of panel **D**. **(F)** Fictitious neuron-type-based synaptic network graphs as expected from the eventual convergence of empirical connectomic and connectome-based neurotaxonomic analysis. **(G)** A multiplex superimposing all three NP coupling graphs from panel E, with neuron types a-d serving as common nodes in this unified diffusion graph. The fictitious excitatory and inhibitory synaptic networks of panel F are overlaid here to represent the aspiration to eventually integrate empirical descriptions of synaptic and myriad modulatory networks.

Transcriptomic prediction of paracrine local signaling from GABAergic neuron sources is particularly compelling. Because few cortical GABAergic neurons have axons that project beyond the confines of a single cortical area, considerations of diffusion physics and the limited lifetime of peptides after secretion strongly imply that secreted neuropeptides must act locally, if at all. The extremely high levels of NPP expression in GABAergic neurons argue, in turn, that they must act somewhere. Most cortical glutamatergic neurons do emit long axons, so it is possible that neuropeptides secreted from such neurons may act in remote and perhaps extracortical projection target areas. Even so, most cortical glutamatergic neurons do have locally ramifying axons and may also secrete neuropeptides from their local dendritic arbors (Vila-Porcile et al., 2009). The high cortical expression of NP-GPCRs cognate to NPP genes expressed by glutamatergic neurons in the same local area suggests a scenario supportive of local modulatory signaling from glutamatergic neuron sources, even though this case may not be quite as strong as that for GABAergic neurons. That said, the much more profuse expression of NPP genes in

GABAergic neuron types along with the somewhat more profuse NP-GPCR expression in glutamatergic types still suggests a "prevailing wind" of peptidergic signaling, blowing predominantly from GABAergic to glutamatergic neurons, as presaged in an earlier microarray analysis of developing mouse cortex (Batista-Brito et al., 2008). Though our NP network predictions are entirely consistent with decades of pioneering work on peptidergic neuromodulation and cortical gene expression (Burbach, 2011; Hökfelt et al., 2013; van den Pol, 2012), it is perhaps only with the recent advent of data with single-cell resolution and genomic depth that it has become reasonable to propose the extreme neuron-type-specificity and density of network coverage suggested by our analysis . The cell-type-specific patterning of NP gene expression has allowed us to cast our predictions in testable form, and we believe emerging means to perturb and sense neuropeptide signaling, as discussed below, bring with them means to test these predictions critically.

Caveats to transcriptomic prediction. The present predictions of functional neuromodulatory coupling are based on analysis of cellular mRNA abundance, but prediction from such data depends upon (1) extrapolation from cellular mRNA census to inference about the synthesis, processing, localization and functional status of cellular NPP and NP-GPCR proteins, (2) assumptions about neuropeptide diffusion and lifetime in cortical interstitial spaces, (3) assumptions about signal transduction and effector consequences of neuropeptide binding in cortex to target cell NP-GPCR receptors. Though we have already discussed several factors mitigating such concerns, we stipulate here that these uncertainties remain substantial, and note the need for much further investigation.

Testing peptidergic network predictions. Physiological and anatomical experimentation will be essential to testing transcriptomic predictions of intracortical neuropeptide signaling. We have suggested that such work may have been frustrated in the past by irreproducibility due to the uncharted multiplicity, neuron-type-specificity, and redundancy of NPP and NP-GPCR expression. This conundrum may now be resolved with the emergence of transcriptomic neurotaxonomies and new tools for experimental access to specific cortical neuron types. Such access may be either *prospective*, using Cre driver lines (Daigle et al., 2018; He et al., 2016; Madisen et al., 2015) or viral vectors (Dimidschstein et al., 2016) of substantial neuron-type-specificity, or *retrospective* using highly multiplexed FISH (Lein et al., 2017; Zeng and Sanes, 2017) or immunostaining methods (He et al., 2016; Xu et al., 2010), patch-seq (Cadwell et al., 2017; Lein et al., 2017) or morphological neuron type classification methods (DeFelipe et al., 2013; Zeng and Sanes, 2017). By allowing the generation of highly specific predictions of peptidergic signaling between specific neuron types, these new molecular tools should enormously advance the prospects for decisive and repeatable tests of neuron-type-specific intracortical neuropeptide signaling hypotheses.

A vast pharmacopeia of well-characterized specific ligands and antagonists for most NP-GPCRs (Alexander et al., 2017) will be bedrock for the functional analysis of neuron-type-specific peptide signaling. For analysis of type-specific neuropeptide signaling in network context (i.e., *ex vivo* slices and *in vivo*), newer optophysiological methods of calcium imaging and optogenetic stimulation/inhibition will certainly join electrophysiology as foundations for measurement of neuropeptide impacts. In addition, many new tools more specific to neuropeptide signaling are emerging. Super-resolution 3D immunohistologies like array tomography (Smith, 2018) and 3D single-molecule methods (Jia et al., 2014; von Diezmann et al., 2017) will enable imaging of DCV localization and neuropeptide contents in type-specific network anatomical context. Genetically encoded sensors of extracellular GPCR ligands (Patriarchi et al., 2018; Sun et al., 2018), GPCR activation (Haider et al., 2019; Hill and Watson, 2018; Livingston et al., 2018; Ratnayake et al., 2017; Stoeber et al., 2018), G-protein mobilization (Ratnayake et al., 2017), cAMP concentration (Hackley et al., 2018; Ma et al., 2018), protein kinase activation (Chen et al., 2014) and protein phosphorylation (Haider et al., 2019) will enable fine dissection of NP dynamics and NP-GPCR signal transduction events (Spangler and Bruchas, 2017). In addition, new caged NP-GPCR ligands (Banghart et al., 2018) and antagonists (Banghart et al., 2013) will provide for precise spatial and temporal control for NP receptor activation. All of these tools have been demonstrated already in physiological applications, and all should be readily applicable to testing specific hypotheses derived from the type-specific peptidergic signaling predictions we have set forth.

Prospects for elucidating network homeostasis, modulation and plasticity. The original motivation for the present analysis was to deepen our understanding of the homeostasis, modulation and plasticity of cortical synaptic networks. Our work has raised the prospect that dense and highly multiplexed peptidergic neuromodulation could play very significant roles in these processes. Due to the clearly formidable complexity of cortical networks, however, a real grasp of the myriad network interactions implicated is certain to require theoretical and computational approaches, in addition to the biophysical approaches outlined in the preceding section. Perhaps most intriguing in the more theoretical directions are concepts that have emerged from work at the fertile intersection of the neuroscience of learning and memory and the computer science of machine learning and artificial neural networks (Dayan and Abbott, 2001; Huh and Sejnowski, 2017; Koch and Segev, 1998; Lillicrap et al., 2016; Marblestone et al., 2016; Shai and Larkum, 2017; Song et al., 2000).

Neuroscience and computer science efforts to model or engineer adaptive neural networks share the hard problem of optimal adjustment of large numbers of what both fields call "synaptic weights". At the heart of this challenge is "credit assignment", that is, the assignment of "credit" for progressive improvement during network development and learning processes to the correct subsets of synapses as needed to guide individualized synaptic weight adjustment. Neuroscientists struggle with the credit assignment problem as they search for the relevant biological learning rules. Computer scientists struggle with the excessive computational requirements of currently standard backpropagation-of-error-based credit assignment. One concept that has come into prominence as a candidate biologically plausible solution to the credit assignment problem is that of modulated "Hebbian" or "spike-timing-dependent" plasticity (STDP) (Bengio et al., 2016; Dan and Poo, 2006; Farries and Fairhall, 2007; Florian, 2007; Frémaux and Gerstner, 2016; Izhikevich, 2007; Marblestone et al., 2016; Pawlak et al., 2010; Poo et al., 2016; Roelfsema and Holtmaat, 2018; Xie and Seung, 2003). While most biological studies of modulated STDP so far have focused on the monoamine neuromodulator dopamine (Izhikevich, 2007; Kuśmierz et al., 2017; Schultz, 2015), known commonalities of signal transduction downstream from widely varying GPCRs suggest strongly that NP-GPCRs could play roles closely analogous to those postulated for dopamine-selective GPCRs (Hamilton et al., 2013; Roelfsema and Holtmaat, 2018).

A neurotaxonomic framework for integrating multiple, superimposed modulatory and synaptic networks, analogous to that schematized in very simple form by **Fig. 7,** may prove critical to advancing theoretical analyses of synaptic network homeostasis and plasticity. At present, efforts in this direction are limited by scant empirical information on synaptic connectomes and their neurotaxonomic annotation. It is very encouraging, however, that vigorous ongoing efforts, e.g., see (Daigle et al., 2018; Swanson and Lichtman, 2016; Tasic, 2018; Zeng and Sanes, 2017), suggest that such information is likely to materialize soon.

Prospects for neuropsychiatric drug development. Molecular components of neuropeptide signaling have beguiled as drug since targets the first wave of discovery that crested in the late twentieth century (Hökfelt et al., 2003; Hoyer and Bartfai, 2012). Many billions of dollars have been invested accordingly, but the returns seem to have been less than originally hoped. The present study raises the possibility that both NP-targeted drug discovery and the reproducibility of physiological experimentation have been hindered in similar ways by the same uncharted multiplicity, cell-type-specificity and redundancy of NPP and NP-GPCR expression. By charting these waters, single-neuron transcriptomic analysis may improve the odds substantially for both reproducible research and drug development.

Today's psychiatric pharmaceuticals almost all target signaling by the monoamine neuromodulators dopamine, serotonin, noradrenaline and/or histamine and their selective GPCR receptors (Data-Franco et al., 2017; Hamon and Blier, 2013; Millan et al., 2015; Urs et al., 2014). Because they are so numerous, neuropeptide signaling systems may be much more neuron-type specific than monoamines. Greater neuron-type-specificity may translate to NP-targeting drugs being less troubled by side-effects and compensation (Hoyer and Bartfai, 2012). Moreover, while GPCRs have long been known as among the most "druggable" of targets (Gurrath, 2001; Lundstrom, 2009), the "druggability" of GPCRs is currently advancing very rapidly due to advances in GPCR structural biology and molecular dynamic simulations (Hilger et al., 2018; Koehl et al., 2018; Weis and Kobilka, 2018). It seems likely that new knowledge of

peptide component neuron-type-specificity may substantially advance the development of NP-targeting pharmaceuticals.

Conclusions. Analysis of single-cell RNA-Seq data from mouse cortex reveals a new panoramic view of NPP and NP-GPCR gene expression. This view exposes an unexpected density and multiplicity of neuropeptide gene expression, as we have summarized and discussed. We have articulated many of findings into new and specific predictions regarding ways that cortical neurons may modulate one another's function. These predictions are just now subject to experimental test with the recent emergence of transcriptomic neurotaxonomies, means for genetic access to specific neuron types and powerful new tools for biophysical analysis of neuropeptide actions. Such tests are likely to greatly deepen our understanding of adaptive cortical function.

Acknowledgements

We wish to thank the Allen Institute for Brain Science founder, Paul G. Allen, for his vision, encouragement and support. This work was supported in part by award number R01NS092474 from the Office of the Director of National Institutes of Health and award number R01MH104227 from the National Institute of Mental Health. The content is solely the responsibility of the authors and does not necessarily represent official views of the National Institutes of Health.

Bibliography

Abbott LF, Regehr WG. 2004. Synaptic computation. *Nature* **431**:796–803. doi:10.1038/nature03010

Alexander SPH, Christopoulos A, Davenport AP, Kelly E, Marrion N V., Peters JA, Faccenda E, Harding SD, Pawson AJ, Sharman JL, Southan C, Davies JA. 2017. THE CONCISE GUIDE TO PHARMACOLOGY 2017/18: G protein-coupled receptors. *Br J Pharmacol* **174**:S17–S129. doi:10.1111/bph.13878

Banghart MR, He XJ, Sabatini BL. 2018. A Caged Enkephalin Optimized for Simultaneously Probing Mu and Delta Opioid Receptors. *ACS Chem Neurosci* **9**:684–690. doi:10.1021/acschemneuro.7b00485

Banghart MR, Williams JT, Shah RC, Lavis LD, Sabatini BL. 2013. Caged Naloxone Reveals Opioid Signaling Deactivation Kinetics. *Mol Pharmacol* **84**:687–695. doi:10.1124/mol.113.088096

Baraban SC, Tallent MK. 2004. Interneuron Diversity series: Interneuronal neuropeptides - Endogenous regulators of neuronal excitability. *Trends Neurosci* **27**:135–142. doi:10.1016/j.tins.2004.01.008

Bargmann CI. 2012. Beyond the connectome: How neuromodulators shape neural circuits. *BioEssays* **34**:458–465. doi:10.1002/bies.201100185

Batista-Brito R, MacHold R, Klein C, Fishell G. 2008. Gene expression in cortical interneuron precursors is prescient of their mature function. *Cereb Cortex* **18**:2306–2317. doi:10.1093/cercor/bhm258

Bengio Y, Lee D-H, Bornschein J, Mesnard T, Lin Z. 2016. Towards Biologically Plausible Deep Learning. doi:10.1007/s13398-014-0173-7.2

Borbély É, Scheich B, Helyes Z. 2013. Neuropeptides in learning and memory. *Neuropeptides* **47**:439–450. doi:10.1016/j.npep.2013.10.012

Bucher D, Marder E. 2013. XSnapShot: Neuromodulation. *Cell* **155**:482-482.e1. doi:10.1016/j.cell.2013.09.047

Burbach JPH. 2011. What are neuropeptides? *Methods Mol Biol* **789**:1–36. doi:10.1007/978-1-61779-310-3_1

Cadwell CR, Sandberg R, Jiang X, Tolias AS. 2017. Q&A: Using Patch-seq to profile single cells. *BMC Biol* **15**:1–7. doi:10.1186/s12915-017-0396-0

Chen Y, Saulnier JL, Yellen G, Sabatini BL. 2014. A PKA activity sensor for quantitative analysis of endogenous GPCR signaling via 2-photon FRET-FLIM imaging. *Front Pharmacol* **5 APR**:1–12. doi:10.3389/fphar.2014.00056

Crawley JN. 1985. Comparative Distribution of Cholecystokinin and Other Neuropeptides: Why is This Peptide Different from All Other Peptides? *Ann N Y Acad Sci* **448**:1–8. doi:10.1111/j.1749-6632.1985.tb29900.x

Daigle TL, Madisen L, Hage TA, Valley MT, Knoblich U, Larsen RS, Takeno MM, Huang L, Gu H, Larsen R, Mills M, Bosma-Moody A, Siverts LA, Walker M, Graybuck LT, Yao Z, Fong O, Nguyen TN, Garren E, Lenz GH, Chavarha M, Pendergraft J, Harrington J, Hirokawa KE, Harris JA, Nicovich PR, McGraw MJ, Ollerenshaw DR, Smith KA, Baker CA, Ting JT, Sunkin SM, Lecoq J, Lin MZ, Boyden ES, Murphy GJ, da Costa NM, Waters J, Li L, Tasic B, Zeng H. 2018. A Suite of Transgenic Driver and Reporter Mouse Lines with Enhanced Brain-Cell-Type Targeting and Functionality. *Cell* **174**:465-480.e22. doi:10.1016/j.cell.2018.06.035

Dan Y, Poo M-M. 2006. Spike Timing-Dependent Plasticity: From Synapse to Perception. *Physiol Rev* **86**:1033–1048. doi:10.1152/physrev.00030.2005

Data-Franco J, Singh A, Popovic D, Ashton M, Berk M, Vieta E, Figueira ML, Dean OM. 2017. Beyond the therapeutic shackles of the monoamines: New mechanisms in bipolar disorder biology. *Prog Neuro-Psychopharmacology Biol Psychiatry* **72**:73–86. doi:10.1016/j.pnpbp.2016.09.004

Dayan P, Abbott LF. 2001. Theoretical Neuroscience - Computational and Mathematical Modeling of

Neural Systems. MIT Press.

DeFelipe J, López-Cruz PL, Benavides-Piccione R, Bielza C, Larrañaga P, Anderson S, Burkhalter A, Cauli B, Fairén A, Feldmeyer D, Fishell G, Fitzpatrick D, Freund TF, González-Burgos G, Hestrin S, Hill SL, Hof PR, Huang ZJ, Jones EG, Kawaguchi Y, Kisvárday Z, Kubota Y, Lewis D a, Marín O, Markram H, McBain CJ, Meyer HS, Monyer H, Nelson SB, Rockland K, Rossier J, Rubenstein JLR, Rudy B, Scanziani M, Shepherd GM, Sherwood CC, Staiger JF, Tamás G, Thomson A, Wang Y, Yuste R, Ascoli G a. 2013. New insights into the classification and nomenclature of cortical GABAergic interneurons. *Nat Rev Neurosci* **14**:202–16. doi:10.1038/nrn3444

Dimidschstein J, Chen Q, Tremblay R, Rogers SL, Saldi GA, Guo L, Xu Q, Liu R, Lu C, Chu J, Grimley JS, Krostag AR, Kaykas A, Avery MC, Rashid MS, Baek M, Jacob AL, Smith GB, Wilson DE, Kosche G, Kruglikov I, Rusielewicz T, Kotak VC, Mowery TM, Anderson SA, Callaway EM, Dasen JS, Fitzpatrick D, Fossati V, Long MA, Noggle S, Reynolds JH, Sanes DH, Rudy B, Feng G, Fishell G. 2016. A viral strategy for targeting and manipulating interneurons across vertebrate species. *Nat Neurosci* **19**:1743–1749. doi:10.1038/nn.4430

Elphick MR, Mirabeau O, Larhammar D. 2018. Evolution of neuropeptide signalling systems. *J Exp Biol* **221**:jeb151092. doi:10.1242/jeb.151092

Fan X, Dong J, Zhong S, Wei Y, Wu Q, Yan L, Yong J, Sun L, Wang Xiaoye, Zhao Y, Wang W, Yan J, Wang Xiaoqun, Qiao J, Tang F. 2018. Spatial transcriptomic survey of human embryonic cerebral cortex by single-cell RNA-seq analysis. *Cell Res* **28**:730–745. doi:10.1038/s41422-018-0053-3

Farries MA, Fairhall AL. 2007. Reinforcement Learning With Modulated Spike Timing Dependent Synaptic Plasticity. *J Neurophysiol* **98**:3648–3665. doi:10.1152/jn.00364.2007

Férézou I, Hill EL, Cauli B, Gibelin N, Kaneko T, Rossier J, Lambolez B. 2007. Extensive overlap of mu-opioid and nicotinic sensitivity in cortical interneurons. *Cereb Cortex* **17**:1948–1957. doi:10.1093/cercor/bhl104

Florian R. 2007. Reinforcement learning through modulation of spike-timing dependent synaptic plasticity. *Neural Comput* **19**:1468–1502.

Földy C, Darmanis S, Aoto J, Malenka RC, Quake SR, Südhof TC. 2016. Single-cell RNAseq reveals cell adhesion molecule profiles in electrophysiologically defined neurons. *Proc Natl Acad Sci* **113**:E5222–E5231. doi:10.1073/pnas.1610155113

Frémaux N, Gerstner W. 2016. Neuromodulated Spike-Timing-Dependent Plasticity, and Theory of Three-Factor Learning Rules. *Front Neural Circuits* **9**. doi:10.3389/fncir.2015.00085

Fu A, Pachter L. 2016. Estimating intrinsic and extrinsic noise from single-cell gene expression measurements. *Stat Appl Genet Mol Biol 2016* **15**:447–471. doi:10.1515/sagmb-2016-0002

Gallopin T, Geoffroy H, Rossier J, Lambolez B. 2006. Cortical sources of CRF, NKB, and CCK and their effects on pyramidal cells in the neocortex. *Cereb Cortex* **16**:1440–1452. doi:10.1093/cercor/bhj081

Gokce O, Stanley GM, Treutlein B, Neff NF, Camp JG, Malenka RC, Rothwell PE, Fuccillo M V., Südhof TC, Quake SR. 2016. Cellular Taxonomy of the Mouse Striatum as Revealed by Single-Cell RNA-Seq. *Cell Rep* **16**:1126–1137. doi:10.1016/j.celrep.2016.06.059

Gomtsian L, Bannister K, Eyde N, Robles D, Dickenson AH, Porreca F, Navratilova E. 2018. Morphine effects within the rodent anterior cingulate cortex and rostral ventromedial medulla reveal separable modulation of affective and sensory qualities of acute or chronic pain. *Pain* **159**:1. doi:10.1097/j.pain.0000000000001355

Gonzalez-Suarez AD, Nitabach MN. 2018. Peptide-Mediated Neurotransmission Takes Center Stage. *Trends Neurosci* **41**:325–327. doi:10.1016/j.tins.2018.03.013

Grimmelikhuijzen CJP, Hauser F. 2012. Mini-review: The evolution of neuropeptide signaling. *Regul*

Pept **177**:S6–S9. doi:10.1016/j.regpep.2012.05.001

Gurrath M. 2001. Peptide-Binding G Protein-Coupled Receptors: New Opportunities for Drug Design. *Curr Med Chem* **8**:1605–1648.

Hackley CR, Mazzoni EO, Blau J. 2018. cAMPr: A single-wavelength fluorescent sensor for cyclic AMP. *Sci Signal* **11**. doi:10.1126/scisignal.aah3738

Haider RS, Godbole A, Hoffmann C. 2019. To sense or not to sense—new insights from GPCR-based and arrestin-based biosensors. *Curr Opin Cell Biol* **57**:16–24. doi:10.1016/j.ceb.2018.10.005

Hamilton TJ, Xapelli S, Michaelson SD, Larkum ME, Colmers WF. 2013. Modulation of Distal Calcium Electrogenesis by Neuropeptide Y1 Receptors Inhibits Neocortical Long-Term Depression. *J Neurosci* **33**:11184–11193. doi:10.1523/jneurosci.5595-12.2013

Hamm HE. 1998. The Many Faces of G Protein Signaling. *J Biol Chem* **273**:669–672.

Hamon M, Blier P. 2013. Monoamine neurocircuitry in depression and strategies for new treatments. *Prog Neuro-Psychopharmacology Biol Psychiatry* **45**:54–63. doi:10.1016/j.pnpbp.2013.04.009

Hastie T, Tibshirani R, Friedman J. 2001. The Elements of Statistical Learning: Data Mining, Inference, Prediction (pages 369-370). Springer US. doi:10.1198/jasa.2004.s339

Hawrylycz M, Lein E, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA, Van De Lagemaat LN, Smith KA, Ebbert A, Riley ZL, Abajian C, Beckmann CF, Bernard A, Bertagnolli D, Boe AF, Cartagena PM, Mallar Chakravarty M, Chapin M, Chong J, Dalley RA, Daly BD, Dang C, Datta S, Dee N, Dolbeare TA, Faber V, Feng D, Fowler DR, Goldy J, Gregor BW, Haradon Z, Haynor DR, Hohmann JG, Horvath S, Howard RE, Jeromin A, Jochim JM, Kinnunen M, Lau C, Lazarz ET, Lee C, Lemon TA, Li L, Li Y, Morris JA, Overly CC, Parker PD, Parry SE, Reding M, Royall JJ, Schulkin J, Sequeira PA, Slaughterbeck CR, Smith SC, Sodt AJ, Sunkin SM, Swanson BE, Vawter MP, Williams D, Wohnoutka P, Ronald Zielke H, Geschwind DH, Hof PR, Smith SM, Koch C, Grant SG, Jones AR. 2012. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**:391–399. doi:10.1038/nature11405

He M, Tucciarone J, Lee S, Nigro MJJ, Kim Y, Levine JMM, Kelly SMM, Krugikov I, Wu P, Chen Y, Gong L, Hou Y, Osten P, Rudy B, Huang ZJ. 2016. Strategies and Tools for Combinatorial Targeting of GABAergic Neurons in Mouse Cerebral Cortex. *Neuron* **92**:555. doi:10.1016/j.neuron.2016.10.009

Hilger D, Masureel M, Kobilka BK. 2018. Structure and dynamics of GPCR signaling complexes. *Nat Struct Mol Biol* **25**:4–12. doi:10.1038/s41594-017-0011-7

Hill SJ, Watson SP. 2018. Fluorescence Approaches Unravel Spatial and Temporal Aspects of GPCR Organisation, Location, and Intracellular Signalling. *Trends Pharmacol Sci* **39**:91–92. doi:10.1016/j.tips.2017.12.001

Hinton GE, Salakhutdinov RR. 2006. Reducing the dimensionality of data with neural networks. *Science (80-)* **313**:504–507.

Hökfelt T, Bartfai T, Bloom F. 2003. Neuropeptides: Opportunities for drug discovery. *Lancet Neurol* **2**:463–472. doi:10.1016/S1474-4422(03)00482-4

Hökfelt T, Ögren SO, Xu ZQD. 2013. Classical Neurotransmitters and Neuropeptides, Second Edi. ed, Handbook of Biologically Active Peptides. Elsevier Inc. doi:10.1016/B978-0-12-385095-9.00251-7

Hoyer D, Bartfai T. 2012. Neuropeptides and neuropeptide receptors: Drug targets, and peptide and non-peptide ligands: A tribute to prof. dieter seebach. *Chem Biodivers* **9**:2367–2387. doi:10.1002/cbdv.201200288

Huh D, Sejnowski TJ. 2017. Gradient Descent for Spiking Neural Networks.

Ioffe S, Szegedy C. 2015. Batch Normalization: Accelerating Deep Network Training by Reducing

Internal Covariate Shift. *arXiv Prepr*.

Izhikevich EM. 2007. Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb Cortex* **17**:2443–2452.

Jekely G. 2013. Global view of the evolution and diversity of metazoan neuropeptide signaling. *Proc Natl Acad Sci* **110**:8702–8707. doi:10.1073/pnas.1221833110

Jia S, Vaughan JC, Zhuang X. 2014. Isotropic three-dimensional super-resolution imaging with a self-bending point spread function. *Nat Photonics* **8**:302–306. doi:10.1038/nphoton.2014.13

Katz PS, Lillvis JL. 2014. Reconciling the deep homology of neuromodulation with the evolution of behavior. *Curr Opin Neurobiol* **29**:39–47. doi:10.1016/j.conb.2014.05.002

Kim JK, Kolodziejczyk AA, Illicic T, Teichmann SA, Marioni JC. 2015. Characterizing noise structure in single-cell RNA-seq distinguishes genuine from technical stochastic allelic expression. *Nat Commun* **6**. doi:10.1038/ncomms9687

Kingma DP, Ba J. 2014. Adam: A Method for Stochastic Optimization 1–15.

Koch C, Segev I. 1998. Methods in Neuronal Modeling: From Ions to Networks, Computational neuroscience. MIT Press. doi:10.1109/MCISE.1999.743629

Koehl A, Hu H, Maeda S, Zhang Y, Qu Q, Paggi JM, Latorraca NR, Hilger D, Dawson R, Matile H, Schertler GFX, Granier S, Weis WI, Dror RO, Manglik A, Skiniotis G, Kobilka BK. 2018. Structure of the μ-opioid receptor–Gi protein complex. *Nature* **558**:547–552. doi:10.1038/s41586-018-0219-7

Krishnan A, Schioth HB. 2015. The role of G protein-coupled receptors in the early evolution of neurotransmission and the nervous system. *J Exp Biol* **218**:562–571. doi:10.1242/jeb.110312

Kuffler SW, Jan LY, Jan YN. 1979. A peptide as a possible transmitter in sympathetic ganglia of the frog. *PNAS* **76**:1501–1505.

Kuśmierz Ł, Isomura T, Toyoizumi T. 2017. Learning with three factors: modulating Hebbian plasticity with errors. *Curr Opin Neurobiol* **46**:170–177. doi:10.1016/j.conb.2017.08.020

Lein E, Borm LE, Linnarsson S. 2017. The promise of spatial transcriptomics for neuroscience in the era of molecular cell typing. *Science (80-)*. doi:10.1126/science.aan6827

Lein E, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, Boe AF, Boguski MS, Brockway KS, Byrnes EJ, Chen Lin, Chen Li, Chen TM, Chin MC, Chong J, Crook BE, Czaplinska A, Dang CN, Datta S, Dee NR, Desaki AL, Desta T, Diep E, Dolbeare TA, Donelan MJ, Dong HW, Dougherty JG, Duncan BJ, Ebbert AJ, Eichele G, Estin LK, Faber C, Facer BA, Fields R, Fischer SR, Fliss TP, Frensley C, Gates SN, Glattfelder KJ, Halverson KR, Hart MR, Hohmann JG, Howell MP, Jeung DP, Johnson RA, Karr PT, Kawal R, Kidney JM, Knapik RH, Kuan CL, Lake JH, Laramee AR, Larsen KD, Lau C, Lemon TA, Liang AJ, Liu Y, Luong LT, Michaels J, Morgan JJ, Morgan RJ, Mortrud MT, Mosqueda NF, Ng LL, Ng R, Orta GJ, Overly CC, Pak TH, Parry SE, Pathak SD, Pearson OC, Puchalski RB, Riley ZL, Rockett HR, Rowland SA, Royall JJ, Ruiz MJ, Sarno NR, Schaffnit K, Shapovalova N V., Sivisay T, Slaughterbeck CR, Smith SC, Smith KA, Smith BI, Sodt AJ, Stewart NN, Stumpf KR, Sunkin SM, Sutram M, Tam A, Teemer CD, Thaller C, Thompson CL, Varnam LR, Visel A, Whitlock RM, Wohnoutka PE, Wolkey CK, Wong VY, Wood M, Yaylaoglu MB, Young RC, Youngstrom BL, Yuan XF, Zhang B, Zwingman TA, Jones AR. 2007. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**:168–176. doi:10.1038/nature05453

Lillicrap TP, Cownden D, Tweed DB, Akerman CJ. 2016. Random synaptic feedback weights support error backpropagation for deep learning. *Nat Commun* **7**:1–10. doi:10.1038/ncomms13276

Liu JJ, Sharma K, Zangrandi L, Chen C, Humphrey SJ, Chiu YT, Spetea M, Liu-Chen LY, Schwarzer C, Mann M. 2018. In vivo brain GPCR signaling elucidated by phosphoproteomics. *Science (80-)* **360**. doi:10.1126/science.aao4927

Livingston KE, Mahoney JP, Manglik A, Sunahara RK, Traynor JR. 2018. Measuring ligand efficacy at

the mu-opioid receptor using a conformational biosensor. *Elife* **7**:1–23. doi:10.7554/eLife.32499

Ludwig M, Leng G. 2006. Dendritic peptide release and peptide-dependent behaviours. *Nat Rev Neurosci*. doi:10.1038/nrn1845

Lundstrom K. 2009. An Overview on GPCRs and Drug Discovery: Structure-Based Drug Design and Structural Biology on GPCRs In: Leifert WR, editor. G Protein-Coupled Receptors in Drug Discovery. Totowa, NJ: Humana Press. pp. 51–66. doi:10.1007/978-1-60327-317-6_4

Ma L, Jongbloets BC, Melander JB, Xiong W-H, Qin M, Lameyer TJ, Harrison MF, Zemelman B V., Mao T, Zhong H. 2018. A Highly Sensitive A-Kinase Activity Reporter for Imaging Neuromodulatory Events in Awake Mice. *Neuron* **99**:665–679. doi:10.1016/j.neuron.2018.07.020

Madisen L, Garner ARR, Shimaoka D, Chuong ASS, Klapoetke NCC, Li L, van der Bourg A, Niino Y, Egolf L, Monetti C, Gu H, Mills M, Cheng A, Tasic B, Nguyen TNN, Sunkin SMM, Benucci A, Nagy A, Miyawaki A, Helmchen F, Empson RMM, Knöpfel T, Boyden ES, Reid RC, Carandini M, Zeng H, van der Bourg A. 2015. Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance. *Neuron* **85**:942–958. doi:10.1016/j.neuron.2015.02.022

Mains RE, Eipper BA. 2006. Chapter 18: Peptides In: Siegel GJ, Albers RW, Brady S, Price DL, editors. Basic Neurochemistry. Burlington, MA: Elsevier Academic Press. pp. 317–332.

Marblestone AH, Wayne G, Kording KP. 2016. Towards an integration of deep learning and neuroscience **10**:1–41. doi:10.3389/fncom.2016.00094

Marder E. 2012. Neuromodulation of Neuronal Circuits: Back to the Future. *Neuron* **76**:1–11. doi:10.1016/j.neuron.2012.09.010

Marder E, Goeritz ML, Otopalik AG. 2015. Robust circuit rhythms in small circuits arise from variable circuit components and mechanisms. *Curr Opin Neurobiol* **31**:156–163. doi:10.1016/j.conb.2014.10.012

Markram H, Gerstner W, Sjöström PJ. 2013. Spike Timing-Dependent Plasticity: A Comprehensive Overview, Frontiers in Synaptic Neuroscience. doi:10.3389/978-2-88919-043-0

Maximiliano José N, Hashikawa Y, Rudy B. 2018. Diversity and connectivity of layer 5 somatostatin-expressing interneurons in the mouse barrel cortex. *J Neurosci* **38**:2415–17. doi:10.1523/JNEUROSCI.2415-17.2017

Mccormick DA, Nusbaum MP. 2014. Editorial overview: Neuromodulation: Tuning the properties of neurons, networks and behavior. *Curr Opin Neurobiol* **29**:4–7. doi:10.1016/j.conb.2014.10.010

Mena JD, Selleck RA, Baldo BA. 2013. Mu-Opioid Stimulation in Rat Prefrontal Cortex Engages Hypothalamic Orexin/Hypocretin-Containing Neurons, and Reveals Dissociable Roles of Nucleus Accumbens and Hypothalamus in Cortically Driven Feeding. *J Neurosci* **33**:18540–18552. doi:10.1523/JNEUROSCI.3323-12.2013

Mezey VA, Glasgow E, Kusano K, Chin H, Iii WSY, Gainer H. 1999. Single Cell Reverse Transcription-Polymerase Chain Reaction Analysis of Rat Supraoptic Magnocellular Neurons : Neuropeptide Phenotypes and High Voltage-. *Endocrinoloy* **140**:5391–5401.

Millan MJ, Goodwin GM, Meyer-Lindenberg A, Ove Ogren S. 2015. Learning from the past and looking to the future: Emerging perspectives for improving the treatment of psychiatric disorders. *Eur Neuropsychopharmacol* **25**:599–656. doi:10.1016/j.euroneuro.2015.01.016

Moffitt JR, Hao J, Bambah-mukku D, Lu T, Dulac C, Zhuang X. 2016. High-performance multiplexed fluorescence in situ hybridization in culture and tissue with matrix imprinting and clearing. *Proc Natl Acad Sci* **113**:201617699. doi:10.1073/pnas.1617699113

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**:621–628.

Nadim F, Bucher D. 2014. Neuromodulation of neurons and synapses. *Curr Opin Neurobiol* **29**:48–56. doi:10.1016/j.conb.2014.05.003

Nässel DR. 2009. Neuropeptide signaling near and far: How localized and timed is the action of neuropeptides in brain circuits? *Invertebr Neurosci* **9**:57–75. doi:10.1007/s10158-009-0090-1

Nusbaum MP, Blitz DM, Marder E. 2017. Functional consequences of neuropeptide and small-molecule co-transmission. *Nat Rev Neurosci* **18**:389–403. doi:10.1038/nrn.2017.56

Okaty BW, Sugino K, Nelson SB. 2011. Cell Type-Specific Transcriptomics in the Brain. *J Neurosci* **31**:6939–6943. doi:10.1523/JNEUROSCI.0626-11.2011

Oyelade J, Isewon I, Oladipupo F, Aromolaran O, Uwoghiren E, Aameh F, Aachas M, Aadebiyi E. 2016. Clustering algorithms: Their application to gene expression data. *Bioinform Biol Insights* **10**:237–253. doi:10.4137/BBI.S38316

Patriarchi T, Cho JR, Merten K, Howe MW, Marley A, Xiong WH, Folk RW, Broussard GJ, Liang R, Jang MJ, Zhong H, Dombeck D, von Zastrow M, Nimmerjahn A, Gradinaru V, Williams JT, Tian L. 2018. Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. *Science (80-)* **360**. doi:10.1126/science.aat4422

Paul A, Crow M, Raudales R, He M, Gillis J, Huang ZJ. 2017. Transcriptional Architecture of Synaptic Communication Delineates GABAergic Neuron Identity. *Cell* **171**:522-539.e20. doi:10.1016/j.cell.2017.08.032

Pawlak V, Wickens JR, Kirkwood A, Kerr JND. 2010. Timing is not everything: Neuromodulation opens the STDP gate. *Front Synaptic Neurosci* **2**:1–14. doi:10.3389/fnsyn.2010.00146

Pimentel H. 2014. What the FPKM? A review of RNA-Seq expression units. https://haroldpimentel.wordpress.com/2014/05/08/what-the-fpkm-a-review-rna-seq-expression-units/

Poo M-M, Pignatelli M, Ryan TJ, Tonegawa S, Bonhoeffer T, Martin KC, Rudenko A, Tsai LH, Tsien RW, Fishell G, Mullins C, Gonçalves JT, Shtrahman M, Johnston ST, Gage FH, Dan Y, Long J, Buzsáki G, Stevens C. 2016. What is memory? The present state of the engram. *BMC Biol* **14**:1–18. doi:10.1186/s12915-016-0261-6

Ratnayake K, Kankanamge D, Senarath K, Siripurapu P, Weis N, Tennakoon M, Payton JL, Karunarathne A. 2017. Measurement of GPCR-G protein activity in living cells, 2nd ed, Methods in Cell Biology. Elsevier Inc. doi:10.1016/bs.mcb.2017.07.008

Roelfsema PR, Holtmaat A. 2018. Control of synaptic plasticity in deep cortical networks. *Nat Rev Neurosci* **19**:166–180. doi:10.1038/nrn.2018.6

Rossier J, Chapouthier G. 1982. Brain opiates. *Endeavour* **6**:168–176. doi:10.1016/0160-9327(82)90072-2

Russo AF. 2017. Overview of Neuropeptides: Awakening the Senses? *Headache* **57**:37–46. doi:10.1111/head.13084

Sahara S, Yanagawa Y, O'Leary DDM, Stevens CF. 2012. The Fraction of Cortical GABAergic Neurons Is Constant from Near the Start of Cortical Neurogenesis to Adulthood. *J Neurosci* **32**:4755–4761. doi:10.1523/jneurosci.6412-11.2012

Schultz W. 2015. Neuronal Reward and Decision Signals: From Theories to Data. *Physiol Rev* **95**:853–951. doi:10.1152/physrev.00023.2014

Shah S, Lubeck E, Schwarzkopf M, He T-F, Greenbaum A, Sohn CH, Lignell A, Choi HMT, Gradinaru V, Pierce NA, Cai L. 2016. Single-molecule RNA detection at depth by hybridization chain reaction and tissue hydrogel embedding and clearing. *Development* **143**:2862–2867. doi:10.1242/dev.138560

Shai AS, Larkum ME. 2017. Comment on: Towards deep learning with segregated dendrites. *Elife* **6**:1–2. doi:10.7554/eLife.22901

Shekhar K, Lapan SW, Whitney IE, Tran NM, Macosko EZ, Kowalczyk M, Adiconis X, Levin JZ, Nemesh J, Goldman M, McCarroll SA, Cepko CL, Regev A, Sanes JR. 2016. Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. *Cell* **166**:1308-1323.e30. doi:10.1016/j.cell.2016.07.054

Smith SJ. 2018. Q&A: Array tomography. *BMC Biol* **16**:1–18. doi:10.1186/s12915-018-0560-1

Song S, Miller KD, Abbott LF. 2000. Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nat Neurosci* **3**:919–926. doi:10.1038/78829

Spangler SM, Bruchas MR. 2017. Optogenetic approaches for dissecting neuromodulation and GPCR signaling in neural circuits. *Curr Opin Pharmacol* **32**:56–70. doi:10.1016/j.coph.2016.11.001

Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. 2014. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res* **1t5**:1929–1958.

Stoeber M, Jullié D, Lobingier BT, Laeremans T, Steyaert J, Schiller PW, Manglik A, von Zastrow M. 2018. A Genetically Encoded Biosensor Reveals Location Bias of Opioid Drug Action. *Neuron* 963–976. doi:10.1016/j.neuron.2018.04.021

Sun F, Zeng J, Jing M, Zhou J, Feng J, Owen SF, Luo Y, Li F, Wang H, Yamaguchi T, Yong Z, Gao Y, Peng W, Wang L, Zhang S, Du J, Lin D, Xu M, Kreitzer AC, Cui G, Li Y. 2018. A Genetically Encoded Fluorescent Sensor Enables Rapid and Specific Detection of Dopamine in Flies, Fish, and Mice. *Cell* **174**:481-496.e19. doi:10.1016/j.cell.2018.06.042

Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. 2011. Mammalian genes are transcribed with widely different bursting kinetics. *Science (80-)* **332**:472–474. doi:10.1126/science.1198817

Swanson LW, Lichtman JW. 2016. From Cajal to Connectome and Beyond. *Annu Rev Neurosci* **39**. doi:10.1146/annurev-neuro-071714-033954

Syrovatkina V, Alegre KO, Dey R, Huang XY. 2016. Regulation, Signaling, and Physiological Functions of G-Proteins. *J Mol Biol* **428**:3850–3868. doi:10.1016/j.jmb.2016.08.002

Tasic B. 2018. Single cell transcriptomics in neuroscience: cell classification and beyond. *Curr Opin Neurobiol* **50**:242–249. doi:10.1016/j.conb.2018.04.021

Tasic B, Menon V, Nguyen TNT, Kim TKTTK, Jarsky T, Yao Z, Levi BB, Graybuck L, Sorensen SA, Dolbeare T, Bertagnolli D, Goldy J, Shapovalova N, Parry S, Lee CC, Smith K, Bernard A, Madisen L, Sunkin SM, Hawrylycz M, Koch C, Zeng H. 2016. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat Neurosci* **19**:335–346. doi:10.1038/nn.4216

Tasic B, Yao Z, Smith KA, Graybuck L, Nguyen TN, Bertagnolli D, Goldy J, Garren E, Economo MN, Viswanathan S, Penn O, Bakken T, Menon V, Miller JA, Fong O, Hirokawa KE, Lathia K, Rimorin C, Tieu M, Larsen R, Casper T, Barkan E, Kroll M, Parry S, Shapovalova N V, Hirchstein D, Pendergraft J, Kim TK, Szafer A, Dee N, Groblewski P, Wickersham I, Cetin A, Harris JA, Levi BP, Sunkin SM, Madisen L, Daigle TL, Looger L, Bernard A, Phillips J, Lein E, Hawrylycz M, Svoboda K, Jones AR, Koch C, Zeng H. 2018. Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**:72–78. doi:10.1101/229542

Telley L, Govindan S, Prados J, Stevant I, Nef S, Dermitzakis E, Dayer A, Jabaudon D. 2016. Sequential transcriptional waves direct the differentiation of newborn neurons in the mouse neocortex. *Science (80-)* **351**:1443–1446. doi:10.1126/science.aad8361

Tremblay R, Lee S, Rudy B. 2016. GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. *Neuron* **91**:260–292. doi:10.1016/j.neuron.2016.06.033

Urs NM, Nicholls PJ, Caron MG. 2014. Integrated approaches to understanding antipsychotic drug

action at GPCRs. *Curr Opin Cell Biol* **27**:56–62. doi:10.1016/j.ceb.2013.11.002

van den Pol AN. 2012. Neuropeptide Transmission in Brain Circuits. *Neuron* **76**:98–115. doi:10.1016/j.neuron.2012.09.014

Varoqueaux F, Fasshauer D. 2017. Getting Nervous: An Evolutionary Overhaul for Communication. *Annu Rev Genet Vol 51* **51**:455–476. doi:10.1146/annurev-genet-120116-024648

Vila-Porcile É, Xu ZQD, Mailly P, Nagy F, Calas A, Hökfelt T, Landry M. 2009. Dendritic synthesis and release of the neuropeptide galanin: Morphological evidence from studies on rat locus coeruleus neurons. *J Comp Neurol* **516**:199–212. doi:10.1002/cne.22105

von Diezmann A, Shechtman Y, Moerner WE. 2017. Three-Dimensional Localization of Single Molecules for Super-Resolution Imaging and Single-Particle Tracking. *Chem Rev* acs.chemrev.6b00629. doi:10.1021/acs.chemrev.6b00629

Wang X, Allen WE, Wright MA, Sylwestrak EL, Samusik N, Vesuna S, Evans K, Liu C, Ramakrishnan C, Liu J, Nolan GP, Bava FA, Deisseroth K. 2018. Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science (80-)* **361**. doi:10.1126/science.aat5691

Wang Y, Wang M, Yin S, Jang R, Wang J, Xue Z, Xu T. 2015. NeuroPep: A comprehensive resource of neuropeptides. *Database* **2015**:1–9. doi:10.1093/database/bav038

Weis WI, Kobilka BK. 2018. The Molecular Basis of G Protein-Coupled Receptor Activation. *Annu Rev Biochem* **87**:897–919. doi:10.1146/annurev-biochem-060614-033910

Williams JT, Zieglgänsberger W. 1981. Neurons in the frontal cortex of the rat carry multiple opiate receptors. *Brain Res* **226**:304–308. doi:10.1016/0006-8993(81)91103-3

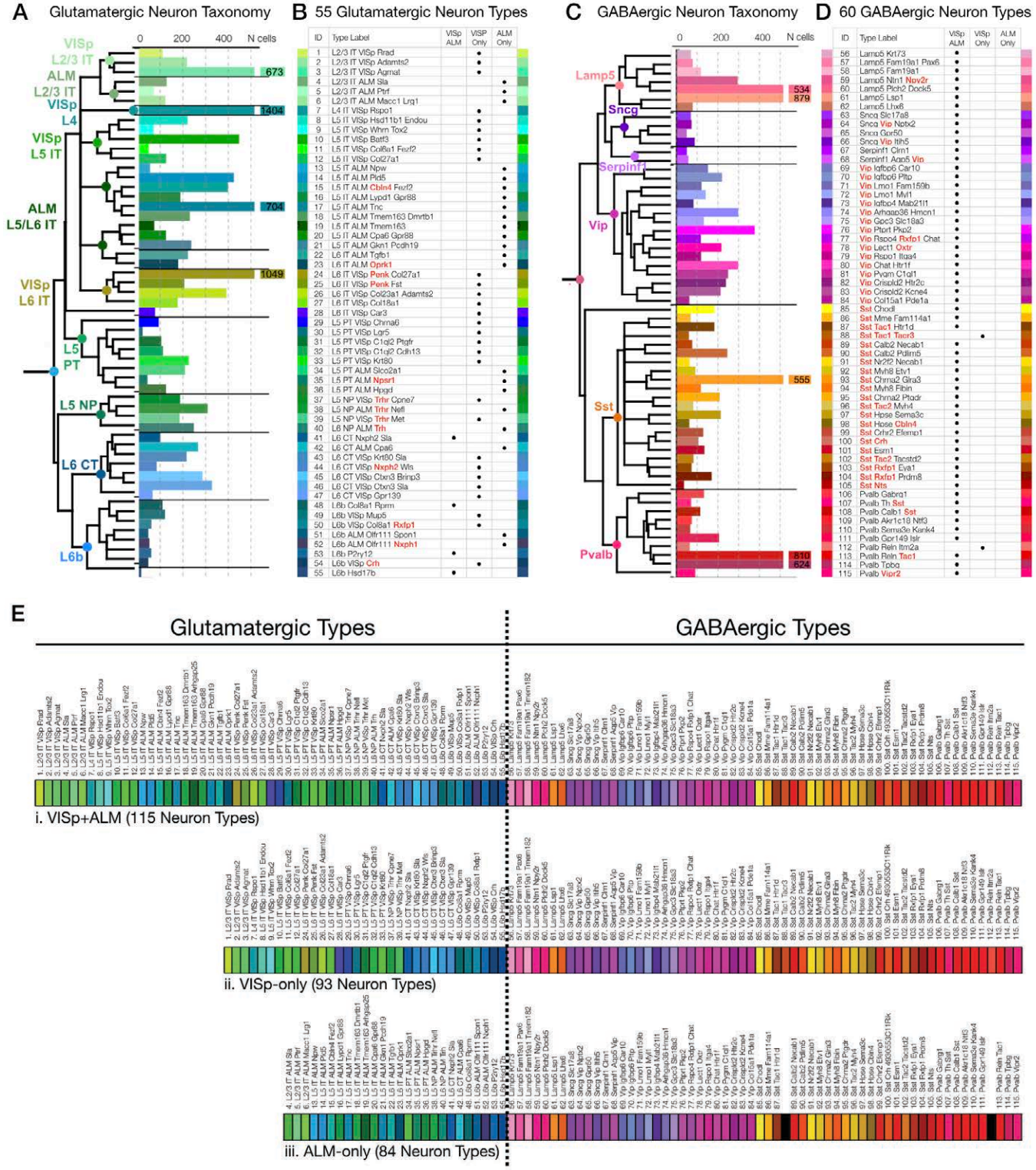
Xie X, Seung HS. 2003. Equivalence of backpropagation and contrastive Hebbian learning in a layered network. *Neural Comput* **15**:441–454. doi:10.1162/089976603762552988

Xu X, Roby KD, Callaway EM. 2010. Immunochemical characterization of inhibitory mouse cortical neurons: Three chemically distinct classes of inhibitory cells. *J Comp Neurol* **518**:389–404. doi:10.1002/cne.22229

Zeng H, Sanes JR. 2017. Neuronal cell-type classification: challenges, opportunities and the path forward. *Nat Rev Neurosci*. doi:10.1038/nrn.2017.85

Supplementary Materials

Supplementary Figure 1. Brief summary of the resource (Tasic et al., 2018) neurotaxonomy.



Transcriptomic neurotaxonomies enable neuron-type-specific profiling of NPP and NP-GPCR gene expression. Tasic and co-workers (Tasic et al., 2018) generated these taxonomies by iterative hierarchical clustering of data from single-cell RNA-Seq analysis of 22,439 mouse cortical neurons

dissociated from areas VISp and ALM. (A) Taxonomic cladogram of glutamatergic neurons. Horizontal bars represent numbers of analyzed cells falling into each neuron type category. (B) Neuron type labels and type-code colors for each glutamatergic neuron type. (C) Taxonomic cladogram for GABAergic neurons. (D) Neuron type labels and type-code colors for each GABAergic neuron type. (E) Concatenated glutamatergic and GABAergic type-code color strips and type labels for (i) the resource aggregate of VISp and ALM neuron types (VISp+ALM), (ii) VISp-only neuron types, and (iii) ALM-only neuron types. The neuron-type color codes and neurotaxonomy strips i-iii will be re-used consistently throughout the present publication.

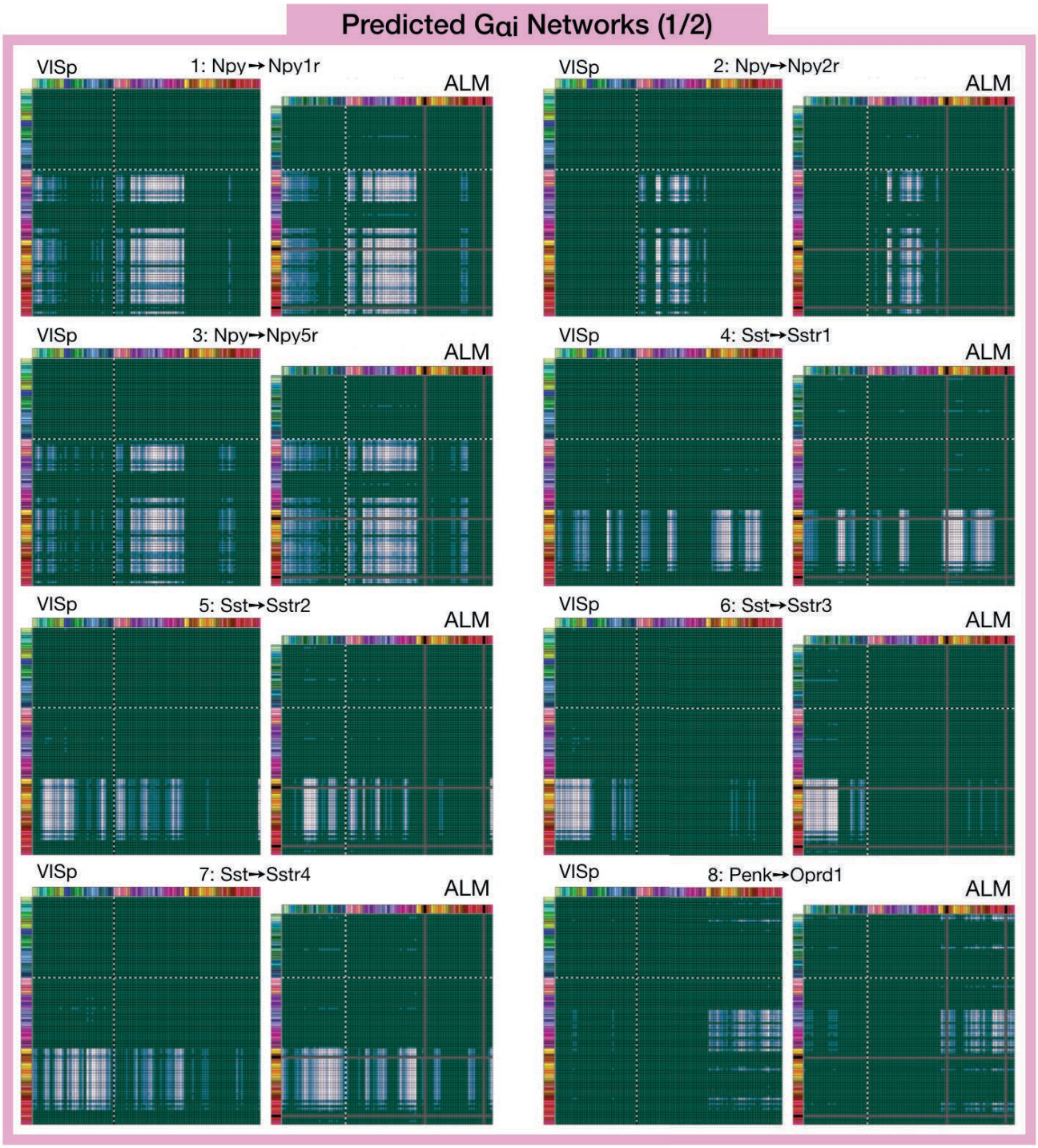
The cladogram in (A) represents the Tasic 2018 taxonomy of glutamatergic neurons in areas VISp and ALM and indicates the number of single-cell transcriptomes mapped to each “leaf” cluster. These cell numbers bear some relation to the actual proportions of cell types in the source tissues, but the relationship is imprecise. More reliable quantitative information about neuron-type proportions is likely to emerge from ongoing spatial transcriptomics studies and is certain to be useful but unlikely to be critical to the present analysis. Panels (A) and (B) also introduces the Tasic neuron-type-coding color scheme for glutamatergic neurons, which are used consistently in Figs. 3-6 of the main text. The table in (B) enumerates and labels each of the 55 glutamatergic neuron type leaf clusters graphed in (A). These neuron type labels were generated by Tasic, et al., to capture salient distinguishing feature for each neuron type. For glutamatergic types, these features always include cortical layer and a VISp vs ALM area designation wherever a cell type was found only in that one area. The labels also include names of key marker genes. Panel (B) shows that most of the 55 Tasic glutamatergic types are exclusive either to VISp or to ALM, with only four (*41. L6 CT Nxph2 Wls, 48. L6b Col8a1 Rxfp1, 53. L6b P2ry12 and 55. L6b Hsd17b*) common to both areas.

Panels (C) and (D) represent the Tasic 2018 taxonomy of 60 GABAergic cortical neuron types just as (A) and (B) represented the 55 Glutamatergic types. These figures also introduce the Tasic type-coding color scheme for GABAergic neurons, which again will be used consistent in all following figures. As evident from comparison of (B) and (D), and discussed extensively by Tasic, et al., GABAergic neuron types are much more prevalently common to both VISp and ALM cortical areas, with all but two (*88. Sst Tac1 Tacr3 and 112. Pvalb Reln Itm2a*) of the 60 GABAergic types found in both areas. For GABAergic neurons, the Tasic type labels refer strictly to marker gene combinations. A notably large fraction of the marker genes appearing in the Tasic type labels are either NPP or NP-GPCR genes. Names of these genes are highlighted by bold red type in both (B) and (D). Most of these red-highlighted NP genes can be found in Table 1 but a few cannot because they did not meet our curation criterion of high cortical expression of both members of a conjugate NPP and NP-GPCR pair.

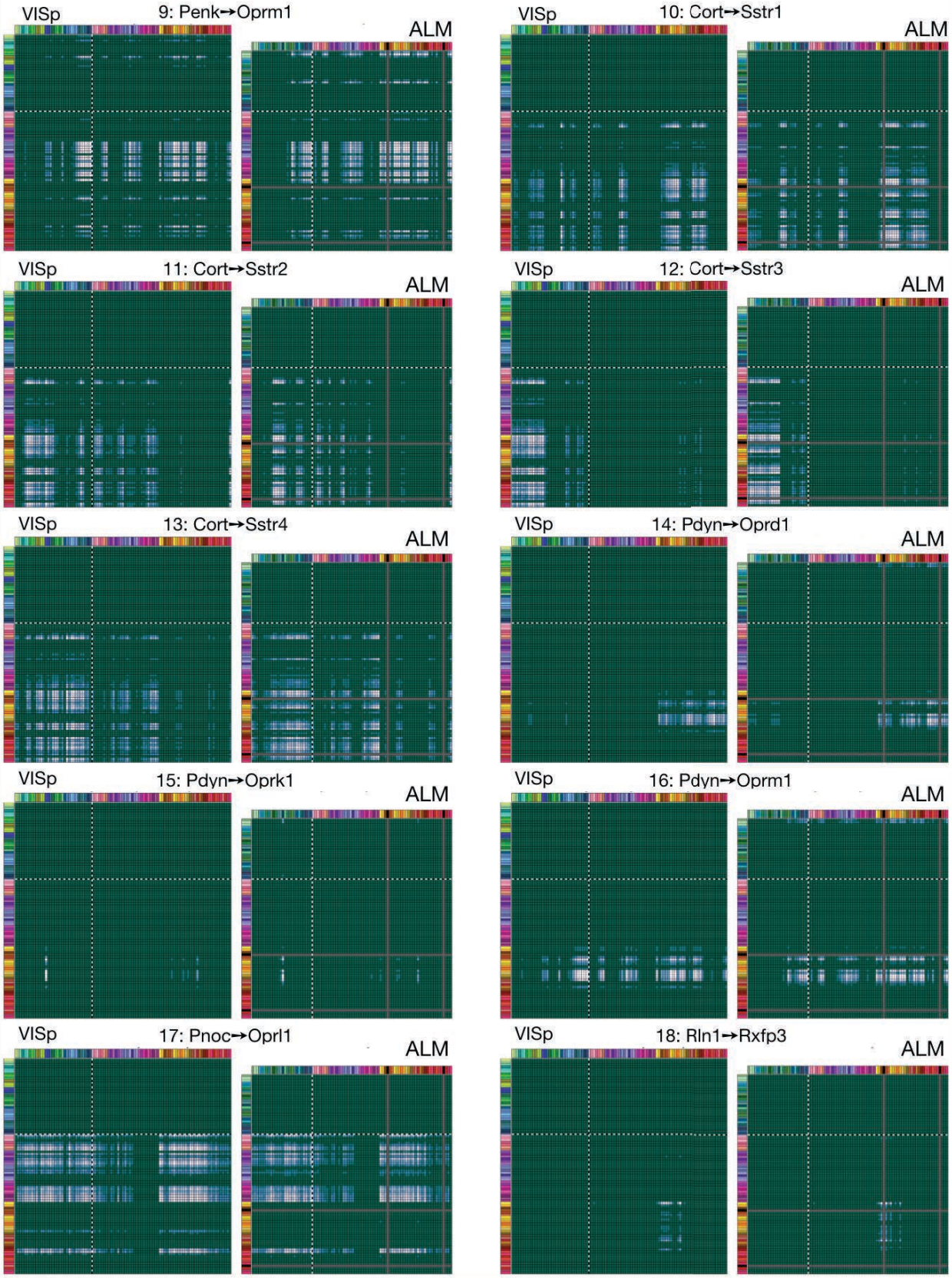
Panel (E) concatenates the Tasic glutamatergic and GABAergic neuron-type rosters and type-code color strips reproduced in Figs. 3A-D, tailored as needed to represent (i) all 115 neurons types in VISp and ALM, (ii) the 93 VISp-only types and (iii) the 84 ALM-only types. These edited rosters and code strips maintain a neuron type sequence dictated by the cladograms of (A) and (C). Though the sequence of glutamate and GABA types is somewhat arbitrary, we concatenated the rosters using glutamatergic-first ordering used by Tasic, et al. (Tasic et al., 2018). Note that (E) sub-panels i, ii and iii are aligned laterally to keep the conserved GABAergic cell types in horizontal register. Type-color pixels are replaced with black spacers for the two GABAergic cell types that are not present in ALM, again to maintain horizontal register of the conserved GABAergic types. Because few glutamatergic cell types are conserved between VISp and ALM, there is no such precise register for the glutamatergic types. The annotated color strips in this panel are provided to aid in interpretation of subsequent figures where space and font-size constraints render repeated typographic annotation impractical. It is hoped that even individuals with color vision anomalies will find these type-code color strips useful due to the redundancy of type order, luminance and hue information.

Supplementary Figure 2. Predicted Neuron-Type-Based Peptidergic Coupling Matrices.

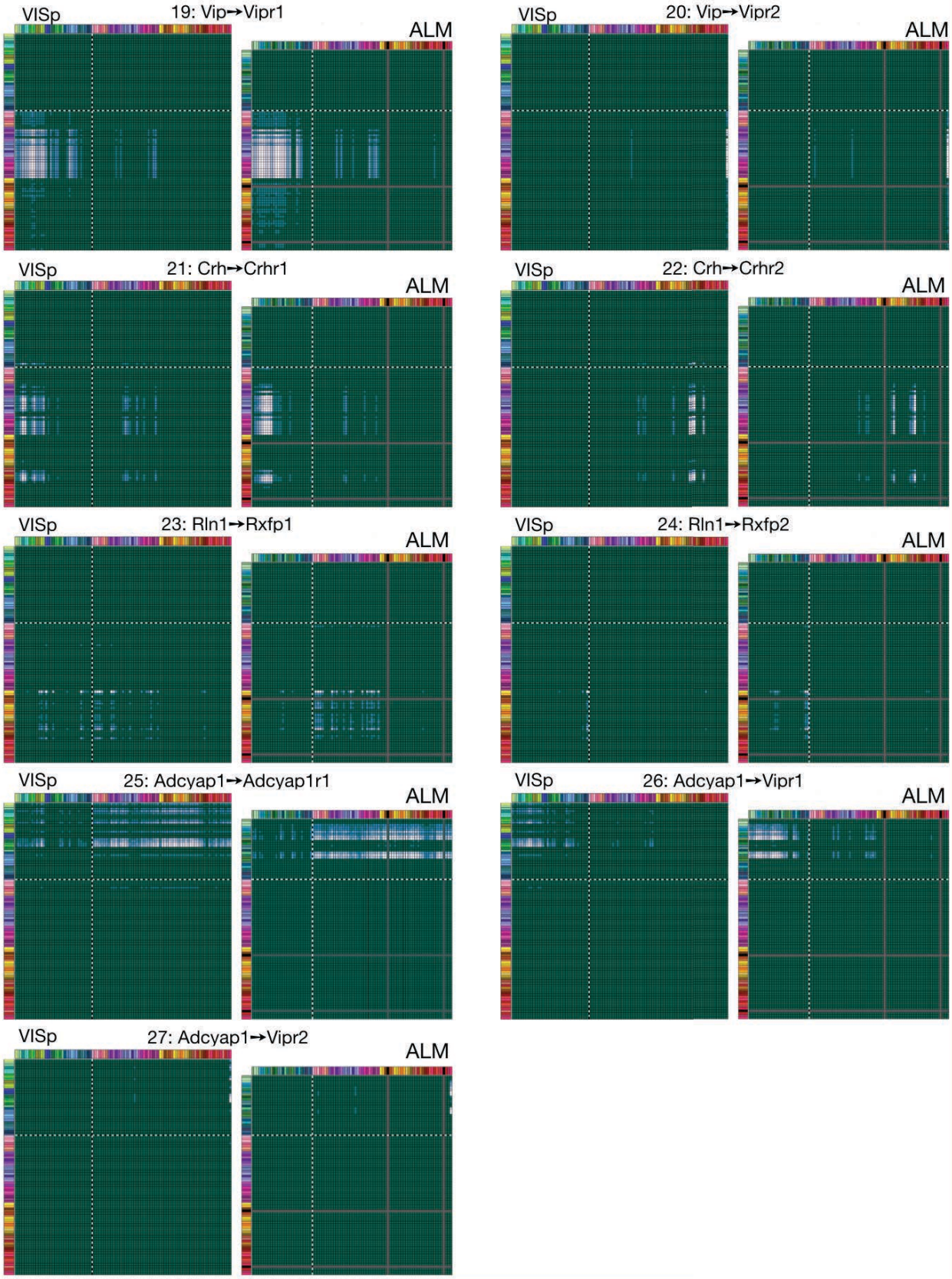
Predictions of neuron-type-based coupling matrices for all 37 of the cognate NPP/NP-GPCR pairs listed in Table 3, pairing predictions for VISp and ALM cortical areas, color grouped according to Gα class. All predictions were generated as described in connection with Fig. 6 and are represented as in Figs. 6C and 6D.



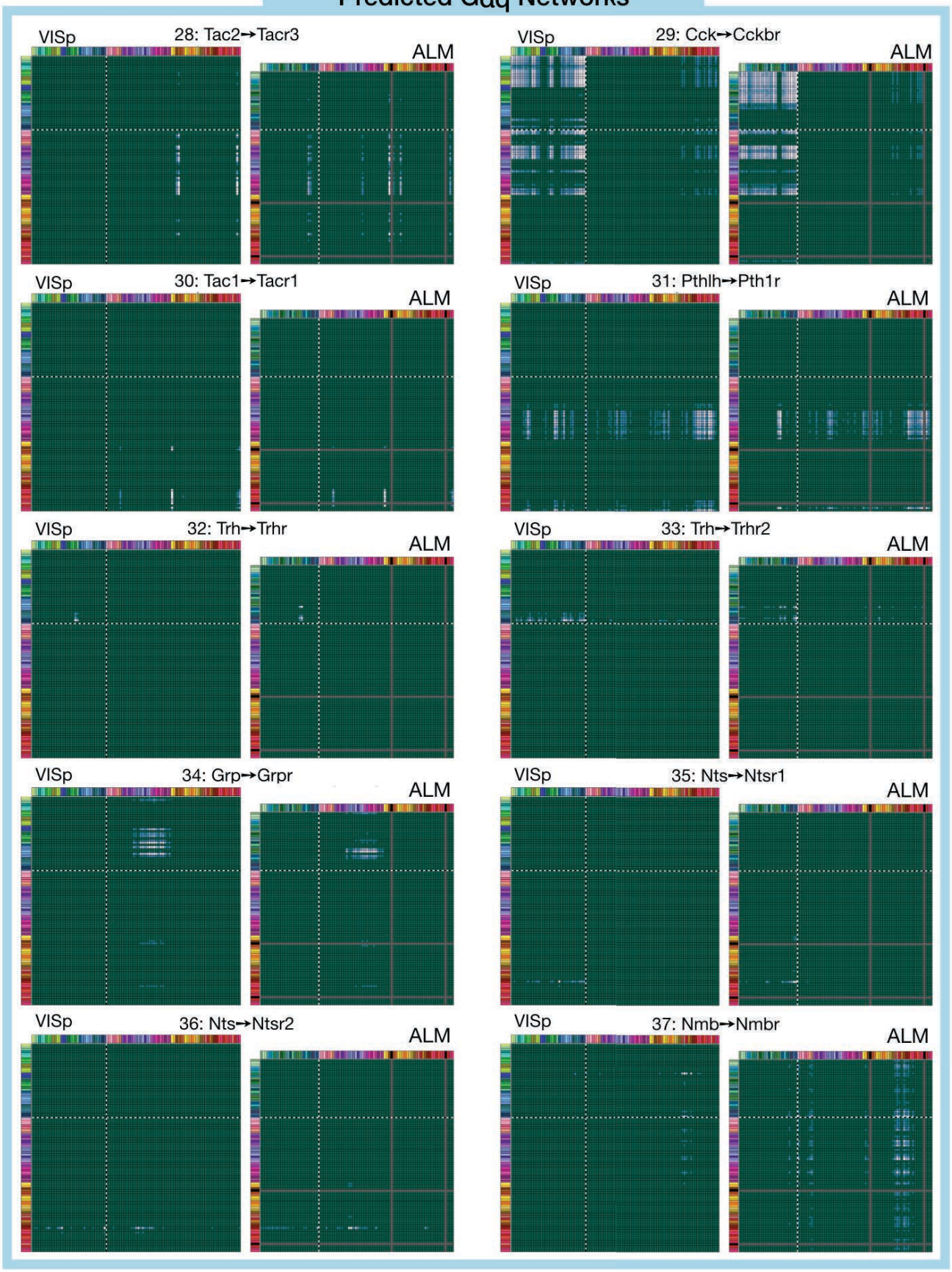
Predicted Gαi Networks (2/2)



Predicted Gαs Networks



Predicted Gαq Networks



Supplementary Methods

Autoencoder-based classifier development and evaluation methods. We used two types of gene datasets: 1) The "HE" gene set, which contains the expression of 6083 highly expressed neuronal genes in 22,439 neurons and 2) 47-gene sets, which contain the expression of 47 specific genes in 22,439 neurons (chosen either as the set of peptidergic precursor genes or random sets of 47 genes as explained in main text). Both "HE" and 47-gene datasets are divided into training and validation sets using a 92%-8% split.

Autoencoders are deep neural network models that consist of encoder/decoder subnetworks. In its basic form (Hinton and Salakhutdinov, 2006), the encoder subnetwork compresses the high dimensional input into a low dimensional representation, and the decoder subnetwork estimates the original input from that low dimensional representation. We constructed a network with two autoencoders, with 8 hidden layers each. The architecture of the first autoencoder("HE Genes autoencoder") is Input(6083) → Dropout(0.8) → Dense(100) → Dense(100) → Dense(100) → Dense(100) → Dense(*d*) → Batch Normalization (*latent representation* z_1) → Dense(100) → Dense(100) → Dense(100) → Dense(100) → Dense(6083), and the architecture of the second autoencoder ("NP Genes autoencoder") is Input(47) → Dropout(0) → Dense(50) → Dense(50) → Dense(50) → Dense(50) → Dense(*d*) → Batch Normalization (*latent representation* z_2) → Dense(50) → Dense(50) → Dense(50) → Dense(50) → Dense(47). Here, the numbers in parentheses denote the number of units in that layer, the numbers of input/output units in each network match the number of input genes, and the Dropout layer (Srivastava et al., 2014) is used to prevent overfitting in the first network. The 2-d representations shown in Fig. 4-a ($d=2$) and the 5-d representations used in Fig4-d,e ($d=5$) are the outputs of the Batch Normalization layer (Ioffe and Szegedy, 2015) for both networks. The Dense layers use the rectified linear (ReLU) function as the nonlinear transformation except for Dense(*d*) layers which do not use a nonlinear transformation. Both networks were iteratively trained using the backpropagation algorithm with the Adam optimizer (Kingma and Ba, 2014) and a batch size of 956. The "HE Genes" network was trained for 50,000 epochs using the mean squared error between the input and the output layers as the loss function. The "NP genes" network was trained for 10,000 epochs using $L=R+\lambda C$ as the loss function, where R denotes the mean squared reconstruction loss as in the HE Genes network, C denotes the coupling loss between the latent representations of the two networks, and $\lambda=100$ is the weighting scalar between the two terms. After training the HE genes network and obtaining the latent representation z_1 for each cell, C calculates the mean squared error between the latent representation of the NP genes network z_2 and z_1, normalized by the minimum eigenvalue of the 2-d representations of each batch during each training iteration. The two additive terms, R and C, together minimize the reconstruction error while attempting to match the representations learned based on the HE gene s*et.* The same procedure was used for all small gene subsets including NP and random gene sets. The Python implementations of the autoencoders using the Tensorflow [128] and Keras [129] libraries will be made publicly available upon acceptance.

We determined the optimal latent dimensionality ($d=5$) by varying the latent space dimensionality of the HE Genes network between 2 and 20 dimensions. The optimal dimensionality was chosen by maximizing the classification accuracy of a Gaussian Mixture Model (GMM) on a test set, whose cluster memberships in the training set are those of the resource taxonomy (Tasic et al., 2018). We used the adjusted Rand index to quantify the similarity between two different partitionings of the same test set (e.g., the Tasic 2018 taxonomy as the ground truth and the

predictions of the GMM), where a score of *1* corresponds to a perfect matching and a score of 0 corresponds to the chance level. At the optimal latent dimensionality of $d=5$, the GMM classifier achieved an adjusted Rand index of 0.8672 on the test set.

We quantified the performance of the GMM classifiers due to the different gene sets using the hierarchical dendrogram of the Tasic 2018 taxonomy by calculating the Resolution Index (RI) [33] for each cell. RI measures the depth of the first common ancestor of the predicted node and the original node in the taxonomy, from the lowest resolution at the root (RI = 0) to the finest resolution at the leaves (RI = 1). To account for the exclusion of all non-neuronal cell-types, scores were normalized over the resolution index corresponding to the first neuronal node on the Tasic 2018 taxonomy [33]. The performance of each gene set was quantified by taking the average RI score (across the cells) due to the respective GMM classifiers acting on the respective latent space representations. The RI scores reported in the main text are averages over all cells (training and test). The corresponding scores due to the test set only are 0.920 for the NP 47-gene subset, 0.768±0.039 for the expression-matched random sets of 47 genes, and 0.464±0.069 for the random sets of 47 genes, demonstrating an even wider performance gap.