Neuroendocrinology of the lung revealed by single cell RNA sequencing

Christin S. Kuo[1,2], Spyros Darmanis[3*], Alex Diaz de Arce[2], Yin Liu[2], Nicole Almanzar[1], Timothy T.H. Wu[2], Stephen R. Quake[3,4], and Mark A. Krasnow[2]

[1]Department of Pediatrics, Stanford University School of Medicine, Stanford, California, 94305, USA

[2]Department of Biochemistry and Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California 94305-5307, USA

[3]Department of Bioengineering, Stanford University, Stanford, California, 94305, USA

[4]Chan-Zuckerburg Biohub, San Francisco, California, USA

*Current address: Department of Microchemistry, Proteomics & Lipidomics, Genentech, South San Francisco, California, USA
Correspondence to M.A.K. (krasnow@stanford.edu)

ABSTRACT

Pulmonary neuroendocrine cells (PNECs) are sensory epithelial cells that transmit airway status to the brain via sensory neurons and locally via CGRP and GABA. Several other neuropeptides and neurotransmitters have been detected in various species, but the number, targets, functions, and conservation of PNEC signals are largely unknown. We used scRNAseq to profile hundreds of the rare mouse and human PNECs. This revealed >40 PNEC neuropeptide and peptide hormone genes, most cells expressing unique combinations of 5-18 genes. Peptides are packaged in separate vesicles, release presumably regulated by the distinct, multimodal combinations of sensors expressed by each PNEC. Expression of cognate receptors predicts an array of local targets, and we show the new PNEC signal angiotensin directly activates innervating sensory neurons. Many signals lack lung targets so may have endocrine activity like those of PNEC-derived carcinoid tumors. PNECs are an extraordinarily rich and diverse signaling hub rivaling the enteroendocrine system.

INTRODUCTION

Pulmonary neuroendocrine cells (PNECs) are neuroepithelial cells scattered throughout the epithelium lining the airways, many as solitary cells and others in clusters of ~30 cells (called neuroepithelial bodies or NEBs) at airway branchpoints (Scheuermann, 1997). PNECs are thought to monitor airway oxygen and respiratory status, and rapidly signal this internal sensory information via afferent sensory neurons to the brain to regulate breathing and related functions (Adriaensen et al., 2009; Cutz et al., 2013; Nonomura et al., 2017; Xu et al., 2020). Such signaling may be mediated by the classical neurotransmitter serotonin (Lauweryns et al., 1973), but PNECs also produce GABA (Schnorbusch et al., 2013) and several neuropeptides including CGRP, which can signal locally to lung goblet (GABA) and immune cells (CGRP) and may contribute to asthma (Barrios et al., 2019; Sui et al., 2018). PNECs also function as reserve stem cells that repair the surrounding epithelium after injury (Ouadah et al., 2019; Song et al., 2012; Stevens et al., 1997), and they can be transformed by loss of tumor suppressor genes into the deadliest human lung cancer, small cell lung cancer (SCLC) (Ouadah et al., 2019; Park et al., 2011; Song et al., 2012; Sutherland et al., 2011), and likely other neuroendocrine tumors such as lung carcinoids accompanied by systemic symptoms from signals secreted by the tumor (Davila et al., 1993).

Despite the physiological and clinical significance of PNECs, molecular interrogation and understanding of PNEC function and diversity (Mou et al., 2021) has lagged because the cells are so rare, comprising just ~0.01% of human lung cells (Boers et al., 1996; Travaglini et al., 2020). Here we describe the isolation, expression profiling by single cell RNA sequencing (scRNA-seq), and analysis of hundreds of mouse and human PNECs. The results reveal an extraordinary molecular diversity of these cells, including the expressed sensors along with dozens of neuropeptides and peptide hormones whose predicted targets indicate they transmit

airway sensory information throughout the lung, to the brain, and potentially to the rest of the body.

RESULTS

Enrichment and single cell RNA sequencing of mouse pulmonary neuroendocrine cells

Because PNECs are among the rarest of lung cell types, they are not found, or only poorly represented, in lung scRNA-seq studies (Han et al., 2018; Tabula Muris Consortium, 2018; Travaglini et al., 2020), www.lungmap.net). We therefore genetically labeled PNECs using a Cre recombinase-dependent fluorescent reporter (ZsGreen) by tamoxifen induction of *Ascl1*$^{CreER/+}$; *Rosa26*$^{ZsGreen/+}$ mice, and then depleted other, abundant lung cell populations and enriched for the labeled PNECs prior to scRNA-seq (Figs. 1A, S1A and B). Poly-adenylated mRNA from each sorted ZsGreen+ or control cell was reverse transcribed and PCR-amplified using Smart-seq2 protocol (Picelli et al., 2014), and the obtained cDNA was used to generate libraries sequenced to a depth of 10^5 - 10^6 reads/cell and quantified to determine expression levels of each gene in each cell. Cells with similar expression profiles were computationally clustered (Butler et al., 2018), and the cell type identity of each cluster was assigned based on expression of canonical lung cell type markers including *Ascl1*, *Calca*, and *Chga* for PNECs (Fig. S1C). After filtering low quality cells and cell doublets, we obtained high quality transcriptomes (2025 $\pm$450 genes (mean $\pm$SD) detected per cell) of 176 PNECs. We also obtained 358 other pulmonary cells (Fig. S1D) including *Ascl1* lineage-labeled glial cells.

We identified dozens of mouse PNEC markers (Fig. 1B, C, Table S1) beyond the canonical four (*Calca* (encoding calcitonin gene-related peptide CGRP), *Ascl1*, *Syp* (synaptophysin), and *Chga* (chromogranin A)) (Table S2) using Wilcoxon rank sum (Seurat v2.3.4) to find differentially-expressed genes selectively expressed by PNECs, then prioritizing those with high sensitivity (expressed in >85% of PNECs), specificity (<5% of the other lung cell types), and fold-difference in mean expression over other lung cell types. Many of the

identified markers were expressed at higher levels (*Resp18, Pcsk1, Scg5*) and were more sensitive (*Resp18, Pcsk1*) and specific (*Chgb*, *Sez6l2*) than the canonical PNEC markers (Fig. 1C). Among the top reported tracheal neuroendocrine cell (TNEC) markers, just over half (54% (19/35) (Montoro et al., 2018); 63% (31/49) (Plasschaert et al., 2018) were also top PNEC markers, defining a core set of airway neuroendocrine markers (Table S1). But there were also notable differences between TNECs and PNECs (e.g. TNEC-selective *Cib3, Cxcl13, Scna, Mthfd2, Slc38a5, Slc26a9, St8sai, Kcnj1*, and *Sox1*), consistent with their different development, distribution, innervation, and functions (Kuo and Krasnow, 2015; Montoro et al., 2018; Mou et al., 2021; Ouadah et al., 2019; Song et al., 2012). Selective expression of the most robust PNEC markers was validated by single molecule in situ hybridization (smFISH) (*Resp18*, *Scg5,* Fig. 1D) and immunohistochemistry (*Pcsk1,* see below).

PNECs express dozens of neuropeptide and peptide hormone genes

PNECS were discovered by their hallmark secretory vesicles (Scheuermann, 1997) and later shown to secrete serotonin (Lauweryns et al., 1972), GABA (Barrios et al., 2017), and several neuropeptides including CGRP (Johnson et al., 1988; Scheuermann et al., 1987) and chromogranin A (Lauweryns et al., 1987) (Table S2). To determine the full set of PNEC neurotransmitters and neuropeptides, we searched our scRNA-seq dataset for neurotransmitter biogenesis genes (Table S3) and neuropeptide and peptide hormone genes ("peptidergic" genes, Table S4) expressed in PNECs. This confirmed that PNECs are GABAergic because many (28%) expressed biosynthetic gene *Gad1*. Although none expressed serotonin biosynthetic genes *Tph1* or *Tph2*, many (24%) expressed the reuptake transporter *Slc6a4,* suggesting they can signal through serotonin imported from other sources. Some PNECs may also use other neurotransmitters because rare cells expressed key dopaminergic (*Th, Ddc)*, glutamatergic (*Slc17a6*), cholinergic (*Slc18a3/VAChT*), or histaminergic (*Hdc*) genes (Table S3).

PNECs expressed dozens of genes that encode neuropeptides and/or peptide hormones. Thirty-one peptidergic genes were expressed in our dataset, and 30 of those plus an additional 12 were found in our second, independent dataset (see below), totaling 43 (47%) of the 91 peptidergic genes in mice (Fig. 1E, Table S4). The full set includes previously known *Calca* (encodes CGRP) and *Chga* but also some of the most biologically and medically significant neuropeptide and peptide hormone genes (e.g., *Pomc* (pro-opiomelanocortin), *Hcrt* (hypocretin/orexin), *Agt* (angiotensinogen), *Gnrh1* (gonadotropin releasing hormone), *Oxt* (oxytocin), *Lhb* (luteinizing hormone), *Crh* (corticotropin-releasing hormone), *Adm* (adrenomedullin), *Ghrl* (ghrelin), *Igf2* (insulin-like growth factor), *Inha*, *Inhba* and *Inhbb* (inhibins), *Edn1* and *Edn3* (endothelins), *Nppa* (atrial natriuretic peptide)). The seven granin genes (*Chga, Chgb, Scg2, Scg3, Scg5, Pcsk1n, Vgf*) and 16 others (*Calca*, *Calcb, Cartpt, Agt, Iapp, Adcyap1, Igf1, Igf2, Lhb, Inha, Inhba, Inhbb, Adm, Oxt, Pomc, Crh)* have also been detected in neuroendocrine (NE) cells outside the lung, but this is the first description of NE cell expression for 17 PNEC peptidergic genes (*Dbi, Nppa, Nppb, Nppc, Nmb, Npb, Npff, Npw, Pnoc, Prok2, Gal, Gnrh1, Hcrt, Ucn2, Agrp, Thpo, Uts2*).

Some of the peptidergic genes including *Calca, Igf2,* and granins (*Chga, Chgb, Scg2, Scg3, Scg5, Pcsk1n*) were expressed in most PNECs analyzed, but others were detected in minor subpopulations (Fig. 1E, Table S4). We analyzed peptidergic gene expression in our second scRNA-seq dataset of 92 PNECs that used a different strategy for PNEC labeling and capture (Ouadah et al., 2019) and obtained a similar distribution of peptidergic gene expression, with the exception of *Pcsk1n* (expressed in 15% of PNECs vs. 60% in original dataset) and *Ucn2* (35% vs 0%) (Fig. 1E, Table S4). We also validated expression of eight peptidergic genes and determined the distribution and abundance of the expressing cells in vivo by immunostaining and multiplex smFISH. The canonical mouse PNEC marker and neuropeptide CGRP (*Calca)* was detected in 95% of PNECs by scRNA-seq, 95% of PNECs by immunostaining (n=237 scored cells in 3

mice) (Fig. 1F), and 94% by smFISH (n=100 scored cells in 2 mice) (Fig. 1D). *Cartpt* was detected in 34% of PNECs by scRNA-seq and twice that by immunostaining (68%) (Fig. 1F) and smFISH (67%) (Fig. 1G). smFISH also confirmed PNEC expression of the six other peptidergic genes examined (*Scg5*, *Chga*, *Agt, Pomc, Nmb, Adcyap1*) but generally gave higher percentages of expressed PNECs especially for the less abundantly expressed genes, implying smFISH is more sensitive than scRNA-seq in detecting gene expression (Figs. 1G, 2B). Indeed, when PNECs with very low expression (1-4 mRNA puncta detected per cell) were not included in the smFISH quantification, there was excellent agreement between smFISH and scRNA-seq values (Fig. 1G). The broadly-expressed peptidergic genes *Calca, Scg5, Chga,* and *Cartpt* were detected in both solitary and clustered PNECs, and the four genes detected in smaller subpopulations (*Agt, Pomc, Nmb, Adcyap1*) were most commonly detected in NEBs at bronchial branchpoints, with local clustering observed within a NEB of the *Pomc*-expressing cells (Fig. S2A).

We conclude that PNECs express dozens of peptidergic genes, nearly half (47%) of all annotated peptidergic genes and over an order of magnitude more than previously known, with most expressed in rare PNEC subpopulations. Statistical modeling of our PNEC sampling in scRNA-seq indicated we likely achieved saturation of PNEC peptidergic genes (Fig. S2B), however the value of 43 expressed genes is a lower limit because even more sensitive methods of detecting gene expression in individual cells such as smFISH could identify additional PNEC peptidergic genes.

PNECs express myriad combinations of peptidergic genes

The scRNA-seq analysis showed that individual mouse PNECs expressed 7.2 ±1.9 (mean ±S.D.) peptidergic genes, with some cells expressing up to 13 (range 2-13). Remarkably, almost every PNEC expressed a different combination of peptidergic genes: 154 peptidergic patterns were identified among the 176 PNECs analyzed (Fig. 2A). Because the number of combinations

detected by scRNA-seq could be artificially inflated by "technical dropout" (failure to detect an expressed gene), we also scored peptidergic combinations by the more sensitive technique of multiplex smFISH (Fig. 2B). For the eight peptidergic genes probed in 100 PNECs, we identified 30 different cellular patterns of expression, similar to the 24 patterns detected by scRNA-seq for the same eight genes (Fig. 2C, D). Thus, each PNEC expresses multiple peptidergic genes and in an extraordinary number of combinations.

The number and diversity of neuropeptides and peptide hormones expressed by PNECs is further expanded by post-transcriptional processing. Some of the expressed genes are alternatively spliced to produce transcripts encoding different neuropeptides or hormones with distinct expression patterns and physiological functions. For example, *Calca,* which encodes the classical PNEC neuropeptide CGRP, can be alternatively spliced to generate transcripts encoding the thyroid hormone calcitonin that regulates calcium homeostasis (Amara et al., 1980, 1982). Mapping of PNEC scRNA-seq reads at the *Calca* genomic locus revealed that *Calca* transcripts are alternatively spliced in PNECs (Fig. 3A); most cells expressed both calcitonin and CGRP transcripts, although the ratio varied greatly across individual PNECs and some cells exclusively expressed CGRP (20% of cells) or calcitonin (10%) (Fig. 3B). Similar results obtained by co-staining for CGRP and calcitonin proteins (Fig. 3C).

Many neuropeptide and peptide hormone mRNAs are translated as larger pre-pro-peptides that are proteolytically processed and modified to generate up to eight different neuropeptides and/or peptide hormones, typically with different expression patterns and functions (Table S5), which would further increase the PNEC signaling repertoire. A classic example is POMC, cleaved by proproprotein convertases PCSK1 and PCSK2 to generate ACTH (adrenocorticotrophic hormone), MSH (melanocyte stimulating hormone), β-endorphin as well as others by additional processing events (Fig. 3D). We surveyed expression of the processing

enzyme genes in our scRNA-seq dataset (Fig. 3E) and found that nearly all PNECs expressed *Pcsk1* (Figs. 1C, 3F) and a subset of those expressed *Pcsk2*, predicting that POMC is processed in PNECs into multiple neuropeptides and peptide hormones including ACTH, α3-MSH, and β-endorphin. Confocal imaging of PNEC immunostains for POMC, CGRP, and calcitonin showed that each co-expressed neuropeptide or peptide hormone localized largely to its own secretory vesicles (Fig. 3G), even ones expressed from the same gene (CGRP and calcitonin, Fig. 3C), implying distinct vesicular targeting and packaging pathways. Thus, post-transcriptional and post-translational processing expands the number and diversity of peptidergic signals expressed by each PNEC, and their separate vesicular packaging raises the possibility of differentially-regulated release and impact on targets.

Diverse targets of PNEC signals

To predict the direct targets of the PNEC signals, we searched our molecular cell atlas of the lung, comprising the full expression profiles of nearly all lung cell types (Travaglini et al., 2020) and pulmonary sensory neurons (Liu et al., 2021), for cells that express the cognate receptors (Table S5). The only previously defined targets of PNEC signals are a subpopulation of immune cells (IL5 lineage-positive) proposed to be attracted to PNECs by secreted CGRP in a mouse model of asthma (Sui et al., 2018), and goblet cells, which are increased in macaque and mouse models of inflammation through neurotransmitter GABA from PNECs (Barrios et al., 2019; Sui et al., 2018). Recently, CGRP from tracheal TNECs has been found to support tracheal epithelial cells following hypoxic injury (Shivaraju et al., 2021).

Receptors for serotonin and GABA, the two major PNEC neurotransmitters, are expressed by the two pulmonary sensory neuron (PSN) subtypes that innervate NEBs (PSN4 (*Olfr78*+) and PSN7 (*Calb1*+)) (Fig. 4A), identifying the first signals from PNECs to these afferent fibers that communicate pulmonary sensory information to the brain. GABA receptor

genes were also expressed in goblet cells, supporting the conclusion from macaque and mouse models and identifying the specific GABA receptor subunit as GABRP, as well as in club cells, epithelial cells that neighbor PNECs (Fig. S4). Receptors for the neurotransmitters glutamate, dopamine and histamine predicted by our transcriptomic data to be produced by rare PNECs (see above) were also expressed in innervating PSNs (Fig. 4A) as well as in plasmacytoid dendritic cells (glutamate receptor *Grm8*) and basophils (histamine receptor *Hrh4*) (Fig. S4B).

Of the more than 90 neuropeptides and peptide hormones encoded by the 43 PNEC peptidergic genes, 36 have known receptors (Table S5). The lung expression patterns of these receptors are shown in Fig. 4B, identifying dozens of lung cell types that can directly receive PNEC peptidergic signals. Indeed, every lung cell type expressed a receptor for at least one PNEC peptidergic signal, and most expressed receptors for multiple signals. This suggests that PNECs can function as a signaling hub broadcasting airway sensory information to cells throughout the lung. The richest targets by far were the innervating pulmonary sensory neurons PSN4 and PSN7, which expressed receptors for 17 and 19 PNEC peptidergic signals, respectively. Other rich targets included goblet cells (4 signals) and club cells (3 signals). PNECs themselves expressed IGF receptors, implying autocrine signaling. The results also suggest a broad neuroendocrine-immune signaling axis, with CGRP potentially targeting at least five different types of immune cells (monocytes, dendritic cells, T cells, and alveolar and interstitial macrophages) plus airway smooth muscle and goblet cells (Fig. 4B), and other signals targeting NK cells, B cells, T cells, and dendritic cells (PCSK1N), basophils and neutrophils (IGFs), and dendritic cells (adrenomedullin).

We experimentally validated one of the inferred signaling interactions -- the predicted angiotensin signal from PNECs to pulmonary sensory neurons. Angiotensin is among the most medically important hormones because of its key role in vasoconstriction and blood pressure regulation (Fig. S5), and because one of its processing enzymes (angiotensin converting enzyme

2, ACE2) also serves as the entry receptor for SARS and Covid-19 coronaviruses. Indeed, the lung plays an essential role in this hormone pathway by providing angiotensin converting enzyme (ACE), the target of a ubiquitous class of anti-hypertensive drugs (ACE inhibitors), that proteolytically processes circulating angiotensin I peptide (Agt I) into the potent vasoconstrictor Agt II. Our discovery that PNECs express angiotensinogen (*Agt*), the preprohormone for Agt II, reveals a pulmonary source of the hormone, and our molecular cell atlas points to three potential lung targets: pericytes, and the PSN4 and PSN7 pulmonary sensory neurons that innervate NEBs, each of which selectively expressed Agt II receptor gene *Agtr1a* (Fig. 4B).

We confirmed *Agt* expression in PNECs by smFISH, which showed expressing cells localized within NEBs (Fig. 4C). We also confirmed expression of its receptor *Agtr1a* in NEB-innervating sensory neurons by injecting *Agtr1a-2A-Cre* mice (Leib et al., 2017) with a Cre-responsive reporter virus (AAV-*flex-tdTomato*) in the vagal nodose ganglion where cell bodies of pulmonary sensory neurons reside, and found their tdTomato-labeled termini ramifying on NEBs at airway branchpoints (Fig. 4D). To determine if angiotensin can activate pulmonary sensory neurons, we labeled pulmonary sensory neurons by introducing a fluorescently-labeled wheat germ agglutinin (WGA-A647) into the lung, and allowing 4-5 days for label uptake at sensory neuron termini and retrograde transport to their cell bodies in the vagal ganglia (Fig. 4E). We then isolated and cultured vagal ganglion cells and visualized their neuronal activity by Fluo-4 calcium imaging during perfusion of Agt II into the imaging chamber. Agt II activated 1-15% of cultured pulmonary sensory neurons, as well as rare sensory neurons from other organs (Fig. 4F). We identified responding pulmonary sensory neurons as the PSN4 subtype: they expressed *Agtr1a-2A-Cre* lineage-label and stained positive for PSN4-specific marker neurotrophic receptor tyrosine kinase 1 (TRKA). PSN7 (*Calb1*+) neurons, which are larger but express *Agtr1a* at lower levels and do not express *TrkA* (Liu et al., 2021), were not activated under these conditions. These results suggest that in addition to its classical role as a circulating vasopressor,

Agt II can function as a local neuromodulator from PNECs directly to PSN4 sensory neurons, transmitting airway sensory information to the brain.

For one-third (36%) of the PNEC peptidergic signals (calcitonin/*Calca*, amylin/*Iapp*, ACTH/*Pomc*, neuropeptide W/*Npw*, galanin/*Gal*, urotensin 2/*Uts2*, gonadotropin-releasing hormone/*Gnrh1*, hypocretin/*Hcrt*, ghrelin/*Grhl, oxytocin/Oxt, prokinectin 2/Prok2, leutinizing hormone, subunit B/Lhb,thrombopoietin/Thpo*), we did not detect appreciable expression of their receptors in any lung cell type or PSN (Fig. 4B). All of these classically function as circulating hormones with targets throughout the body, such as calcitonin (bones, kidneys), ACTH (adrenal cortex, adipocytes), amylin (brain stem), and inhibin (pituitary), so PNECs too may secrete these hormones into circulation. However, we cannot exclude that some have local targets but their receptors were expressed below detection or in rare or fragile cells not captured in our lung cell atlas.

PNECs are diverse, multimodal sensors

Classical physiological studies of PNECs indicate that signal secretion is triggered by a variety of stimuli including hypoxia, hypercapnia, mechanical stimuli, and allergens (Lembrechts et al., 2012; Livermore et al., 2015; Sui et al., 2018; Youngson et al., 1993). However, the full diversity of PNEC sensory functions are unknown, and the molecules that mediate these functions have only recently begun to be identified (Nonomura et al., 2017). To more fully elucidate PNEC sensory functions and the molecules that mediate them, and to determine how sensors are paired with the myriad PNEC signals described above, we curated a list of over 1500 mouse genes encoding extant mammalian sensory receptors and their homologues (Table S6) including ones previously implicated in PNEC sensory function, then searched our scRNA-seq dataset for ones selectively expressed in PNECs.

The mechanically-activated channel PIEZO2 (*Fam38b*) gene was specifically and broadly (>90%) expressed by PNECs (Fig. 5A, Fig. S6A), as previously described (Nonomura et

al., 2017). In contrast, PIEZO1 (*Fam38a*) was expressed across all major lung compartments but excluded from PNECs (Fig. S6A). Mechanically-activated two-pore potassium channel TREK-2 (*Kcnk10*) gene (Fig. 5A, S6B) (Bang et al., 2000) and family member *Kcnk16* (Fig. S6B) were selectively expressed in PNECs, whereas TREK-1 (*Kcnk2*) was expressed in other airway epithelial cells and almost completely excluded from PNECs (Fig. S6B). Three TRP family cation channel genes also showed selective or enriched expression in PNECs: auditory hair cell stereocilium channel *Trpml3 (Mcoln3)*, *Trpc4*, and *Trpv2*, a noxious heat sensor also implicated in mechano- and osmoregulation (Fig. 5A, Fig. S6A). *Casr*, a G-protein coupled receptor implicated in ciliary mechanosensing and previously proposed to integrate NEB signals (Lembrechts et al., 2013), was also selectively expressed in PNECs (Fig. 5A, Fig. S6A), as was *Lhfp15*, an integral membrane protein of the auditory hair cell mechanotransduction complex. These results support the role of PNECs as specialized airway mechanosensors and suggest potentially new mechano- or thermosensory functions mediated by proteins encoded by *Kcnk10*, *Kcnk16*, *Trpml3* (*Mcoln3*), *Trpc4*, *Trpv2*, and *Lhfp15*.

PNECs are proposed to function in CO_2 sensing because they can be activated by hypercapnic challenge (Lauweryns et al., 1977), and by bicarbonate and acid in vitro (Ebina et al., 1997; Livermore et al., 2015), a response dependent on carbonic anhydrase, but the proteins that mediate this function are unknown. PNECs selectively expressed the classic acid-sensing potassium channel TREK (*Kcnk3*), and rare PNECs expressed acid-sensing sodium channels ASIC3 (*Accn3*) and ASIC4 (*Accn4*) (Fig. 5A, S6A). Expression of the widely-distributed cytoplasmic carbonic anhydrase *Car2* gene was not detected in PNECs, but some expressed membrane-bound *Car12* and the carbonic anhydrase-related gene *Car11* (Fig. 5A, S6A).

One of the first and still the most prominent proposed function of PNECs is as airway oxygen sensors because they can be activated by hypoxic challenge in vivo (Lauweryns et al., 1978) and in cultured lung slices or as isolated PNECs (Youngson et al., 1993). The oxygen

sensing mechanism is still uncertain but the dominant hypothesis proposes that low oxygen reduces H_2O_2 generation by a membrane-bound NADPH oxidase (heterodimer of gp91phox/*Cybb* and p22phox/*Cyba*, plus regulatory subunits p47phox/Ncf1, p67phox/Ncf2), which inhibits an oxygen-sensitive potassium channel (Kv3.3/Kcnc3 and Kv4.3/Kcnd3) that activates L-type voltage-gated calcium channels (Fu and Cutz, 2002), triggering neurosecretion that may act locally or be propagated to the brainstem breathing center. We did not detect PNEC expression of NADPH oxidase subunits p91phox/*Cybb*, p47-phox/*Ncf1*, or p67-phox/*Ncf2*, and p22phox/*Cyba* was broadly expressed in all cells (Fig. S6A). Kv3.3/*Kcnc3* and Kv4.3/*Kcnd3* were detected only at low levels or in rare cells, however we found robust and selective PNEC expression of many other potassium channel genes (Fig. S6B; e.g., voltage-gated: *Kcnc1, Kcnc2, Kcnb1, Kcnv1, Kcnf1, Kcnq2, Kcnq5, Kcnh2, Kcnh6, Kcnh7, Kcnh8*; cyclic-nucleotide gated Na/K channel: *Hcn1, Hcn2, Hcn3, Hcn4*; calcium-activated: *Kcnn3*; 2-pore: *Kcnk1, Kcnk3*) that could contribute to the hypoxia-sensitive potassium current required for PNEC secretion. We also did not detect selective expression in PNECs of any of the genes (mitochondrial respiratory complexes) implicated in the mitochondrial hypothesis of oxygen sensing (Mulligan et al., 1981; Quintana et al., 2012; Stettner et al., 2011) (Fig. S6A). Thus, the identity of the acute oxygen sensor in PNECs remains uncertain, although several of the newly identified PNEC potassium channels are appealing candidates. Chronic hypoxia also influences PNECs, and hypoxia inducible factor *Hif1a* is expressed though not selectively in PNECs, whereas *Hif3a* is a PNEC-selective family member (Fig. S6A).

Primate PNECs have been proposed as volatile chemical sensors based on expression of olfactory receptors OR2W1 and OR2F1 in some solitary human PNECs, and the response of PNECs in tracheobronchial cultures to nonanal and other chemicals (Gu et al., 2014). We identified 19 olfactory and two pheromone receptor superfamily genes expressed in rare PNECs (Fig. 5A, S6A). We curiously also detected expression of photoreceptor opsin *Opn1sw* and non-

visual opsin *Opn3* (Fig. S6C) as well as two TNF receptor family genes (*Tnfrsf12a, Tnfrsf21*; Fig. S7A), suggesting possible light and immune sensing functions for PNECs.

Individual PNECs express multiple sensors and are predicted to sense multiple modalities. For example, one PNEC (combination 55) expressed mechanoreceptor/transducer genes *Piezo2*, *Casr*, and *Lhfpl5*, acid-sensitive channel *Kcnk3,* and chemoreceptors *Olfr90*, *Olfr92*, and *Vmn2r29* (Fig. 5A,B). Individual PNECs expressed different combinations of sensory genes, indicating diversity in their sensory roles (Fig. 5A,B). Comparison of the sensors and signals expressed in each PNEC did not identify any strong correlations, suggesting that specific sensory inputs are coupled to different output signals in different PNECs.

Human PNECs also show diverse sensory, signaling and target profiles

To explore the generality and biomedical significance of the properties of mouse PNECs uncovered by scRNA-seq, we performed a similar analysis of human PNECs. Although human PNECs are also extremely rare, in our scRNA-seq study of ~75,000 human lung cells (Travaglini et al., 2020) we obtained expression profiles of 55 PNECs. We analyzed these PNEC profiles as we did above for mouse PNECs and found that, even with this more limited sample, all the features uncovered for mouse PNECs are also apparent for human PNECs, although in more extreme form for some features and with species-specific specializations.

Human PNEC markers are largely conserved with mouse and include 26 that are more sensitive (e.g., *GRP*, *SCGN, SCG5, BEX1*), specific (*SLC35D3, CPLX2*), and/or highly expressed (*SCG2*) than the four common clinical markers (*CHGA, SYP, INSM1, ASCL1*) (Fig. S8). Some of the best markers are species-specific (human-specific: *GRP, SCGN*; mouse-specific: *Resp18*), and transcript isoform mapping of the human ortholog (*CALCA*) of the classic mouse PNEC marker CGRP (*Calca*) revealed alternative splicing such that only 60% of human PNECs expressed CGRP isoforms whereas all expressed calcitonin isoforms (Fig. S9),

explaining why calcitonin but not CGRP is a good human PNEC marker (Weichselbaum et al., 2005) whereas both are valuable for mouse (Figs. 1B, 3C, S3).

Like mouse, human PNECs have a large and diverse signaling output. Human PNECs express biosynthetic genes for neurotransmitters serotonin (*TPH2*) and GABA (*GAD1*) (Table S8), the major neurotransmitters of mouse PNECs. Some human PNECs are also likely glutamatergic because 14% expressed glutamate vesicular transporter *SLC17A6*, and some may be catecholaminergic or glycinergic because rare PNECs expressed key catecholamine synthetic enzymes (*DBH, PNMT*) or glycine re-uptake transporter *SLC6A5* (Table S8). Expression of dopaminergic genes was detected in rare mouse PNECs but none of the analyzed human PNECs.

Human PNECs expressed 40 different peptidergic genes (Fig. 6A), 45% of the 93 annotated human peptidergic genes (Table S4), with individual PNECs expressing 12.2 ±2.7 (mean ±SD, range 6-18), twice as many as mouse PNECs (Fig. 6A). Like mouse, almost every PNEC (54 of 55 cells, 98%) expressed a different combination (Fig. 6B). Two-thirds (26/40, 65%) of the expressed peptidergic genes are also expressed in mouse PNECs, although *CARTPT* was expressed in many fewer human PNECs (4% vs 18% in mouse) and *POMC* in many more (57% vs 4% in mouse) (Fig. 6A, Table S4). Remarkably, 13 of the 14 human-specific peptidergic genes encode hormones (Fig. 6A, Table S4), including some of the most biomedically significant: erythropoietin/*EPO*, renin/*REN*, five hypthothalamic releasing/inhibitory hormones and pituitary regulators (thyrotropin-releasing hormone/*TRH*, prolactin-releasing hormone/*PRLH*, gonadotropin-releasing hormone 2/*GNRH2*, corticotropin releasing hormone urocortin/*UCN* and inhibitory hormone somatostatin/*SST*), the common subunit of multiple pituitary hormones (glycoprotein hormone subunit A/*CGA*), reproductive organ developmental regulator anti-mullerian hormone/*AMH*, digestive hormones gastrin-releasing peptide/*GRP* and cholecystokinin/*CCK*, and the potent vasoregulators urotensin 2B/*UTS2B* and kininogen/*KNG1*. Of the 17 mouse-specific PNEC peptidergic genes (Table S4),

all except three (*Igf2,* 72% of PNECs; *Ucn2,* 13%; *Iapp,* 7%*)* were detected only in rare PNECs (1-3%) so could also be rare in human PNECs and found on further profiling.

The human lung cell expression patterns of receptors for the 32 PNEC peptidergic signals with known receptors are shown in Fig. S10A, identifying potential direct targets in lung. As for mouse, expression patterns were diverse and almost all lung cell types expressed receptors for one or more signals, implying human PNECs can also transmit pulmonary sensory information throughout the lung. The predicted targets of the conserved PNEC signals were also largely conserved, for example broad stromal and vascular targeting by inhibin, immune cell targeting by CGRP and VGF, and pericyte targeting by angiotensin. Autocrine signaling appears prominent, as PNECs express receptors for ghrelin and erythropoietin, one of the human-specific signals, and almost every PNEC neurotransmitter (GABA, glutamate, dopamine, and epinephrine/norepinephrine) (Fig. S10B). Curiously, the sole PNEC autocrine signal identified in mouse, IGF2, was not detected in any profiled human PNEC (Fig. 6A, Table S4). No receptor expression was detected in lung for nearly half (45%, 15 of 33) the human PNEC peptidergic signals, including 10 of the 14 human-specific signals. While some of these signals may target pulmonary sensory neurons or rare pulmonary cells not captured in our human lung atlas, others may enter circulation and target sites beyond the lung.

Expression of sensory genes (Fig. S11, S12) indicates that, as in mouse, human PNECs are multimodal sensors with almost all cells expressing different combinations of sensor genes for diverse stimuli. These include orthologues of mouse PNEC sensors such as mechanically-activated channels *PIEZO2* and *KCNK10*, thermosensor *TRPV1*, carbonic anhydrase *CA11*, and acid-sensitive channel TASK-1/*KCNK3*, plus human-specific TASK-3/*KCNK9* with a more acidic range (pK 6.0-6.7) (Duprat et al., 1997). Like mouse, human PNECs express 19 different olfactory receptor genes but in a greater proportion of cells (33% vs. 11% for mouse), with individual cells expressing up to five different *OLFR* genes. As in mouse, NADPH-oxidase

complex genes proposed as PNEC oxygen sensors were not detected or not specifically expressed by PNECs (Fig. S12), but two human PNEC OLFR genes (*OR51E1*/*Olfr558,* the most widely-expressed, in 16% of PNECs, and *OR51E2/Olfr78)* are a close family member and orthologue of OLFR78, which is activated by lactate and implicated in acute hypoxia-sensing in the mouse carotid body (Chang et al., 2015). *OR51E1* and *OR51E2* were co-expressed along with acid-sensitive channel *KCNK3* in a single human PNEC (Fig. S11, combination 25), so this cell may be specialized for hypoxia or chemosensing. As in mice, rare human PNECs also expressed a pheromone receptor (*VN1R1)* and opsins (*OPN1SW*, *OPN3)* (Figs. S11, S12), but human PNECs also expressed bitter (*TAS2R10, TAS2R31, TAS2R5*) and sour taste receptors (*PKD2L1, PKD1L3*) and the trace amino acid receptor *TAAR1*.

Thus, as in mouse, almost every human PNEC is equipped to perceive multiple diverse stimuli and expresses a large number of peptidergic signals and neurotransmitters that can be received by a variety of cells within the lung and potentially throughout the body, like the signals from PNEC-derived tumors described below.

scRNA-seq profile of a human lung carcinoid: amplification of a rare PNEC

PNECs are the presumed origin of a variety of human lung neuroendocrine tumors (Travis et al., 2015) that can cause diverse ectopic hormone syndromes such as classic carcinoid syndrome (wheezing, flushing, diarrhea, increased heart rate), Cushing's syndrome (Arioglu et al., 1998), and acromegaly (Athanassiadi et al., 2004). To determine the full sensory and signaling potential of a PNEC tumor and their relationship to those of normal PNECs, we obtained scRNA-seq profiles of NE tumor cells of a lung carcinoid from a 51 year-old female patient with onset of idiopathic hypertension in the year preceding therapeutic lung resection. The profiled tumor cells expressed most general PNEC markers (e.g., *SCG5, CHGB, SCG2, PCSK1N, CHGA, SCG3*) indicating retention of PNEC identity. However, they did not recapitulate the full spectrum of PNEC diversity in their sensory and signaling gene profiles.

Some peptidergic genes such as *CALCA* and *GRP* that are expressed in almost all normal PNECs were expressed in few if any tumor cells, whereas *NPW*, *NMB*, and *CARTPT* that are expressed in only rare PNECs were expressed in most (*NPW*) or many (*NMB*, *CARTPT*) of the tumor cells (Fig. 6D). Likewise, the tumor cells lacked expression of the common PNEC mechanosensor *PIEZO2* but showed broad expression of the acid-sensitive channel *KCNK3* and opsin *OPN1SW*, which is expressed only by rare normal PNECs (Fig. S11A,B). This suggests that proliferating tumor cells retain a "memory" (albeit imperfect) of the expression profile of the PNEC from which they originated. And, because this tumor's unusual expression profile nearly matched that of a rare normal PNEC (#50, Fig. 6B,C), which like the tumor cells expressed *NPW, NMB, CARTPT, KCNK3,* and *OPN1SW* and lacked expression of common PNEC genes *CALCA, GRP,* and *PIEZO2* (Fig. 6D, E, Fig. S11), we suggest that this carcinoid arose by transformation of a cell similar to PNEC #50. Tumor cell expression of *NPW* may have caused the patient's hypertension because the active peptide (NPW-23) is a circulating hormone that regulates vascular tone and has been proposed to play a role in the pathophysiology of hypertension (Ji et al., 2015; Yu et al., 2007).

To determine the peptidergic signaling profiles of other lung carcinoids, we analyzed peptidergic gene expression in 111 human lung carcinoids profiled by bulk RNA-seq (Alcala et al., 2019). This revealed prominent expression of 70% (28 of 40) of the normal human PNEC peptidergic genes, including new genes we identified in rare subpopulations of normal PNECs, in at least some of the tumors (Fig. S13). Ten peptidergic genes expressed by lung carcinoids, including four previously detected in clinical samples (GHRH, PENK, TAC1, VIP), were not found in our normal PNECs; however, all of these were expressed in only rare carcinoids so may be expressed in rare PNECs and if so should be identified on further PNEC profiling. Interestingly, some of the peptidergic genes expressed in only minor subpopulations of normal PNECs (e.g., *NPW, NPPA, SST, CARTPT*) were detected in many carcinoids, whereas the nearly

ubiquitous PNEC peptidergic gene *CALCA* was absent from most carcinoids. This suggests that not all normal PNECs are equally susceptible to carcinoid transformation.

DISCUSSION

By single cell expression profiling of hundreds of the exceedingly rare PNECs in mouse and human, we discovered they express over 40 peptidergic genes, nearly half of all such genes including many classic hormones. Individual PNECs express up to 18 peptidergic genes, with almost every cell expressing a distinct combination. The diversity of expressed signals is further increased by alternative splicing, and by post-translational processing as inferred from expression of prohormone processing genes. These diverse signals can directly target a wide array of cell types in the lung, predicted by expression of the cognate receptors, including almost every cell type across all five tissue compartments: epithelial (including putative autocrine PNEC signals), endothelial, stromal, immune, and neural. The richest targets are pulmonary sensory neurons (PSNs) that innervate PNECs. We confirmed one predicted signal to PSNs, angiotensin (Agt II, mature product of *Agt)*, can directly activate a recently identified PSN subtype (PSN4) that innervates PNECs, expresses its receptor (*Agtr1a*), and projects to the brainstem to regulate respiratory rate (Liu et al., 2021); Diaz de Arce, et al., unpublished data. Hence in addition to its classical role as a circulating vasopressor whose receptor is targeted by major anti-hypertensive drugs, Agt II may serve as a neuromodulator in the breathing circuit. Eighteen other PNEC signals are also classical hormones, but unlike Agt II their receptors are not expressed in any cell types in the lung cell atlases, suggesting that PNECs could contribute to the circulating pool of these hormones. PNECs are thus extraordinarily rich and diverse signaling hubs that produce scores of neuropeptides and peptide hormones that can signal directly to many cells in the lung, to the brain through pulmonary sensory neurons, and potentially to cells throughout the body through the circulation.

PNECs are scattered throughout the airway epithelium and form large clusters at bronchial branchpoints, so they are ideally positioned to serve as sentinels that monitor inhaled air and airway status. Our single cell data imply that each PNEC is in fact a multimodal sensor,

expressing a distinct combination of mechanical, thermal, and acid sensors along with carbonic anhydrases important in CO_2 sensing, diverse chemosensors including olfactory receptors, vomeronasal receptors and taste receptors, and even light-sensing opsins. Different combinations of sensors are co-expressed in individual PNECs along with different combinations of signals (Figs. 5B, 6C). Because the peptidergic signals are packaged in separate vesicles, secretion of each signal could be independently regulated, presumably in response to activation of a different sensor (or sensor combination). In this way PNECs can simultaneously monitor many aspects of airway status at many positions along the airway, and selectively transmit this information to target cells in the lung, the brain, and the rest of the body.

In many ways PNECs resemble enteroendocrine cells (EECs), the sentinels scattered along the gut epithelium to monitor nutrients, microbial products, and other luminal contents and signal that information locally in the gut to coordinate ingestion, absorption, metabolism, and disposal, and throughout the body and brain to regulate mood and appetite (Bai et al., 2019; Bellono et al., 2017; Kaelberer et al., 2018). The enteroendocrine system is commonly called the "gut-brain axis" and is considered the largest endocrine organ because of its many endocrine cells and signals. Our data suggest it may be rivaled or even surpassed by the pulmonary neuroendocrine system, which expresses more than double the ~20 signals produced by EECs (Beumer et al., 2020). Each EEC apparently expresses only one or a few peptidergic signals and neurotransmitters, which define 12 classical EEC subtypes (Worthington et al., 2018), whereas individual PNECs express 5-10 times more and their expression patterns define at least an order of magnitude more molecular subtypes. While PNECs have long been speculated to serve as local signaling centers in the lung and fast conduits of sensory information to the brain through afferent sensory neurons, our data suggest that like EECs they also serve a more global endocrine function. This would explain why the basal surface of some PNECs, where secretory vesicles are densely packed, are apposed to fenestrated capillaries (Lauweryns et al., 1973). Although PNEC

contribution to circulating hormone pools under normal physiological conditions is yet to be demonstrated, they contribute at least under pathological conditions. Indeed, we found that the extensive PNEC signaling repertoire described here including most of newly identified peptidergic genes, were collectively expressed in the 111 available cases of profiled human lung carcinoids. An individual human carcinoid, however, expresses a discrete set of signals resembling that of a normal PNEC, suggesting that each such tumor amplifies the set of signals expressed by the tumor-initiating PNEC, thereby explaining the diversity of carcinoid syndromes (Limper et al., 1992; Pernow and Waldenstrom, Jan, 1957; Shalet et al., 1979). PNECs may comprise a second global signaling axis we dub "the lung-brain axis".

This new understanding of PNEC function and their extraordinary diversity, including many sensor and signaling genes detected in only a single profiled cell, required pre-enrichment (mouse) or massive profiling (human) to obtain just the first few hundred expression profiles of these exceedingly rare cells. The neuroendocrinology of the lung our study reveals has broad implications for medicine even beyond NE cell tumors (Rudin et al., 2019; Travis et al., 2015; Young et al., 2011), including the many other pulmonary diseases such as asthma (Sui et al., 2018), SIDS (Cutz et al., 2007; Mou et al., 2021), and bronchopulmonary dysplasia (Gillan and Cutz, 1993) that have been associated with PNEC abnormalities and perhaps now extending to diseases outside the lung. The results already have implications for the Covid-19 pandemic because SARS-CoV-2 virions use an angiotensin pathway regulator (ACE2) to enter and destroy lung cells and cripple gas exchange along with the ability of the patient to sense the deficit. Although expression of angiotensinogen and many other PNEC sensory and signaling genes are conserved so their functions can be explored in mice, we also uncovered 13 human-specific PNEC signaling genes encoding classical hormones (e.g., ACTH (*POMC*), GRP, TRH, AMH, CCK). Hence the lung-brain axis may be especially prominent in humans.

ACKNOWLEDGMENTS

We thank Kyle Travaglini and Ahmad Nabhan (preparation of carcinoid tumor from patient 3), Y. Ouadah (mouse Fluidigm scRNA-seq data), and all members of Krasnow lab and Jeffrey Wine (discussions and comments on manuscript); Norma Neff and members of the Quake laboratory (single cell expertise and resources); Cathy Crumpton, Brandon Carter, and Stanford FACS facility (flow sorting); Joseph Shrager and Jalen Benson (IRB protocol and tissue procurement); the patient in this study; Rong Lu (normalization of bulk RNAseq data), and Maria Peterson (figure preparation), the Pediatric Pulmonary Division, and Department of Pediatrics for space and resources. This work was supported by grants from the NIH/NHLBI (K08HL129081, C.S.K), Doris Duke Charitable Foundation (2018105, C.S.K), Howard Hughes Medical Institute (M.A.K.), Ludwig Foundation (M.A.K.), NIH/NCI 5U01CA231851 (M.A.K, C.S.K.) and the Chan Zuckerberg Initiative (M.A.K.). M.A.K. is an investigator of the Howard Hughes Medical Institute.

AUTHOR CONTRIBUTIONS

C.S.K and M.A.K conceived the project. C.S.K and S.D. designed the mouse PNEC scRNA-seq isolation experiments. C.S.K processed tissue and C.S.K and S.D. sorted cells and prepared sequencing libraries. S.D. processed and aligned sequencing data, and performed quality control. C.S.K. and M.A.K. designed bioinformatic analyses, implemented methods, and interpreted results. Y.L. performed scRNA-seq and analysis of pulmonary sensory neurons. A.D.d.A. and C.S.K. designed and performed pulmonary sensory neuron functional imaging experiments. T.W. performed rarefaction analysis. C.S.K. and N.A. performed immunostaining and in situ hybridization. C.S.K. created clinical protocols for human lung tumor and normal tissue acquisition and assessed clinical data. C.S.K, S.R.Q., and M.A.K. provided resources and

supervised the project. C.S.K, and M.A.K. wrote the manuscript, and all authors reviewed and edited the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

FIGURE LEGENDS

Figure 1. Single cell RNA sequencing of mouse PNECs reveals expression of dozens of peptidergic genes. (A) Strategy for labeling, enrichment, and scRNA-seq of exceedingly rare PNECs. Timeline (top) of tamoxifen (Tam) injections (gray arrowheads) of *Ascl1*^{CreERT2};*Rosa26*^{Zsgreen1} mice beginning at embryonic day (E) 13 - E14 to permanently induce ZsGreen expression in pulmonary neuroendocrine cells (PNECs or NE cells). Lungs were dissected at indicated ages (black arrowheads, postnatal day (PN) 21, PN90, and PN120) and mechanically and enzymatically dissociated (dissoc.) into single cells. Endothelial (CD31+) and immune (CD45+) cells were depleted by magnetic-cell sorting (MACS) then PNECs enriched by fluorescence-activated cell sorting (FACS) EpCAM^{+}/Zsgreen^{+} double-positive cells. Sorted cells were analyzed by plate-based scRNA-seq using SmartSeq2 protocol. (B) Most sensitive and specific PNEC markers identified by scRNA-seq, ranked by ratio of the natural logs of the average expression (ln (counts per million, CPM +1)) of the marker in PNECs (NE cells) vs. non-PNEC (non-NE) epithelial cells. *, previously known PNEC marker. (C) Violin plots showing expression of five new markers (*Resp18, Pcsk1, Scg5*, *Chgb, Sez6l2*) and three previously known markers (*; *Calca, Syp, Chga)* across 40 cell types from combined mouse lung cell atlas (Travaglini et al., 2020) From left to right (x-axis): (1) neuroendocrine (NE, PNEC), (2) club, (3) multiciliated, (4) basal, (5) goblet, (6) alveolar type 1, (7) alveolar type 2, (8) glial, (9) smooth muscle, (10) myofibroblast, (11) adventitial fibroblast, (12) alveolar fibroblast, (13) pericyte, (14) mesothelial, (15) chondrocyte, (16) artery, (17) vein, (18) capillary aerocyte, (19) capillary-general, (20) lymphatic, (21) B, (22) Zbtb32+ B cells, (23) plasma, (24) CD8+ T, (25) CD4+ T, (26) regulatory T, (27) Ly6g5bt+ T, (28) natural killer, (29) Alox5+ lymphocyte, (30) neutrophil, (31) basophil, (32) alveolar macrophage, (33) interstitial macrophage, (34) plasmacytoid dendritic, (35) myeloid dendritic type 1, (36) myeloid dendritic type 2, (37) Ccr7+ dendritic, (38) classic monocyte, (39) nonclassical monocyte, (40) intermediate monocyte. (D)

Close-up of neuroendocrine body (NEB) in PN155 wild type (C57BL/6NJ) mouse lung probed by multiplex single molecule RNA fluorescence in situ hybridization (smFISH) to detect expression of indicated PNEC markers, with DAPI nuclear counterstain. Dashed circles, individual PNECs (numbered); dashed line (basement membrane). Scale bar, 10 □m. Quantification (right) of clustered PNECs that express indicated markers (n=76 cells scored in left lobe and right lower lobe). Note classic marker *Calca* (CGRP) was not detected in 6% of *Resp18*$^{+}$*Scg5*$^{+}$ double-positive PNECs. (E) Quantification of peptidergic gene expression in PNECs by scRNA-seq. Bars show percent of profiled PNECs (NE cells, n=176) with detected expression of the 43 peptidergic genes indicated; values above bars are log-transformed mean gene expression (ln (CPM+1)) among expressing cells. Black dots, expression values from a second PNEC dataset (filled circles, n=92 PNECs) in which PNECs were genetically labeled using *Cgrp*^{CreER};*Rosa26*^{Zsgreen} mice, sorted, and isolated on a microfluidic platform (Ouadah et al., 2019). *, previously known mouse PNEC peptidergic genes; blue highlight, classic hormone genes. (F) Micrograph of NEB from PN90 *Ascl1*$^{CreERT2/+}$;*Rosa26*$^{Zsgreen/+}$ mouse lung immunostained for CGRP and newly identified PNEC neuropeptide CARTPT. White arrowheads, CGRP^{+} PNECs; red arrowheads, CARPT^{+} PNECs; *, CGRP^{-} CARPT^{-} double-negative PNEC. Right panel, quantification of CGRP and CARTPT staining in PNECs defined by *Ascl1*-CreER-lineage label (n=237 PNECs scored in three PN60 *Ascl1*$^{CreERT2/+}$;*Rosa26*$^{Zsgreen/+}$ mice). (G) Quantification of PNEC (NE cell) expression of the indicated peptidergic genes by scRNA-seq (black bars, n=176 cells) and multiplex smFISH (grey bars and dashed extensions, n=100 cells scored in NEBs from 2 mice, see Fig. 2). Grey bars, cells with high expression (≥5 puncta/cell); dashed extensions, cells with low expression (1-4 puncta/cell).

Figure 2. PNECs express myriad combinations of peptidergic genes

(A) Peptidergic genes expressed in individual PNECs (n=176) from scRNA-seq. Histogram (top) shows number of PNECs expressing each of the 154 observed combinations of expressed

peptidergic (NP) genes (dots, bottom). Values at right are average expression level of peptidergic genes among expressing cells; values at bottom are number of expressed peptidergic genes for each combination (NPs/cell). Each PNEC expressed 7.2 ±1.9 (mean ±SD) peptidergic genes (range 2-13, median 7, mode 6). (B) Micrograph of NEB from adult PN155 wild type mouse lung probed by multiplex smFISH (RNAscope) for PNEC marker *Resp18* and the indicated peptidergic genes including ones detected by scRNA-seq in most (*Scg5, Calca*), some (*Cartpt, Chga*), or rare (*Agt, Pomc, Nmb, Adcyap1*) PNECs. Epithelial cells are outlined (white dots), basement membrane is indicated by dashed line, and individual PNECs numbered (n = 11) with close-up of the indicated PNEC shown at right and PNEC 1 at bottom. Scale bar, 2 um. Filled colored circles above cells, PNECs with high expression of gene (≥ 5 puncta; filled circles); open circles, PNECs with low expression of gene (1-4 puncta, open circles). Schematic (bottom) shows summary of expression of the eight peptidergic genes (using colored circles as above) in the 11 PNECS in this NEB optical plane. Grey cells, other (non-PNEC) epithelial cells in field of view. Bar, 10 □m. (C) Percent of PNECs expressing each of the 30 observed combinations of the eight peptidergic genes probed by smFISH (n=100 PNECs scored in NEBs from 2 wild type mice). Filled circles, expressed peptidergic gene. (D) Percent of PNECs expressing each of the 24 observed combinations of the same eight peptidergic genes in scRNA-seq dataset (n=176 PNECs).

Figure 3. **Additional PNEC peptidergic diversity from post-transcriptional processing**

(A) Sashimi plots (top) showing mapped scRNA-seq reads for *Calca* gene and deduced alternative splicing patterns for three representative PNECs (cells 2, 9 and 19 in B), and the resultant mRNAs structures (bottom) encoding either CGRP or calcitonin, with exons numbered (black fill, coding exons). Arrowhead, translation start site. Note cell 2 expresses only CGRP, cell 19 expresses only calcitonin mRNA, and cell 9 expresses both. (B) Quantification of alternative splicing of *Calca* mRNA as above for 20 randomly selected *Calca*-expressing PNECs

(sashimi plots in Fig. S3). (C) Fluorescence super-resolution (AiryScan SR) confocal micrograph (left) and close-ups (boxed) of PNECs in adult PN68 wild type mouse lung immunostained for CGRP (green), calcitonin (red), and E-cadherin (Ecad, white) to show epithelial cell boundaries, with DAPI nuclear counterstain (blue). Most PNECs express both peptides but they localize to separate vesicles. Scale bar, 10 □m (inset 2 □m). Right, quantification of immunostaining (n=326 PNECs scored in 3 mice).

(D) Classical post-translational processing scheme for pro-opiomelanocortin (POMC) in anterior pituitary, showing cleavage sites (arrowheads) of endopeptidases proprotein convertase subtilisin/kexin type 1 (PCSK1, orange) and PCSK2 (blue), and carboxypeptidase E (CPE, purple), and modification sites (arrowheads) of peptidyl-glycine–amidating monooxygenase (PAM amidation site, red) and N-acetyltransferase (NAT acetylation site, green) (Harno et al., 2018). ACTH, adrenocorticotrophic hormone; β-LPH, lipotropin, MSH, melanocyte stimulating hormone, Endo, endorphin. Gray box, junctional peptide; red bar, antigen (residues 27-52) of POMC antibody in panel G. In pituitary, other PCSK2 cleavage events produce additional peptides (γ3-MSH, ACTH (1-17), γ-LPH, β-endorphin) although γ–MSH and β-MSH are likely not produced in mouse due to absence of those dibasic cleavage sites. (E) Heatmap of expression of POMC processing genes in individual PNECs (n=176) from scRNA-seq dataset. Note expression of *Pcsk1*, *Cpe*, and *Pam* in most (94%, 65%, 74%), *Pcsk2* in some (16%), and *Nat2* and *Nat3* in rare (1%) PNECs; none expressed *Nat1.* This predicts production of all the major pituitary peptides in PNECs, with individual PNECs producing different sets of POMC peptides due to differential expression of the processing enzymes. (F) Micrograph of NEB in adult PN90 wild type (CD-1) mouse lung immunostained for PCSK1 (red), CGRP (green), and E-cadherin (Ecad, white). Most PNECs co-express PCSK1 and CGRP, although rare PNECs (~10%) express only PCSK1 (asterisks). Scale bar, 10 um. (G) Super-resolution confocal micrograph of

NEB in adult PN155 wild type (CD-1) mouse lung immunostained for CGRP (green), POMC (red), and E-cadherin (Ecad, white), with DAPI counterstain (blue). Bottom panel, close-up of boxed region showing separate POMC and CGRP puncta. Bar, 10 um (inset 2 um).

Figure 4. Predicted lung cell targets of PNEC signals from expression of their receptors. (A) scRNA-seq dot plot of mean expression level (dot heatmap, ratio relative to maximum value for 10 pulmonary sensory neuron (PSN) subtypes, (Liu et al., 2021) and percent of cells in population with detected expression (dot size) in the NEB-innervating PSN subtypes (PSN4, PSN7) of genes encoding receptors for each of the indicated PNEC neurotransmitters. (B) scRNA-seq dot plot of expression in PSN4 and PSN7 (as above) and mean level of expression (dot intensity) and percent of cells in population with detected expression (dot size) in cell types (arranged by tissue compartment) in mouse lung cell atlas (Travaglini et al., 2020) of genes encoding receptors for indicated PNEC peptidergic ligands (shown above corresponding receptor genes). Ligands with expressed receptors are colored; receptors without detectable expression and their corresponding ligands are shown in black. Values (at right), number of PNEC peptidergic signal receptors expressed in at least 10% of the cells of that cell type. CGRP, calcitonin gene-related peptide; Calcit, Calcitonin; Adm, adrenomedullin; Iapp, islet amyloid polypeptide (Amylin); Vgf, neuroendocrine regulatory peptide; Agt, angiotensinogen; α-MSH, alpha melanocyte-stimulating hormone; β-end, β-endorphin; Pnoc, prepro-nociceptin; Edn1, endothelin-1; Edn3, endothelin-3; Nppc, C-type natriuretic peptide; Nmb, neuromedin B; Npff, neuropeptide FF; Crh, corticotropin releasing hormone; Igf1, insulin-like growth factor 1; Igf2, insulin-like growth factor 2; Inhba, inhibin a; Inhbb, inhibin b; Pcsk1n, proprotein convertase substilisin/kexin type 1 inhibitor; Adcyap1, pituitary adenylate cyclase-activating polypeptide; Npw, neuropeptide W; Nppb natriuretic peptide B; Nppa, natriuretic peptide A; Gal, galanin; Uts2, urotensin II; Gnrh1, gonadotropin-releasing hormone 1; Hcrt, hypocretin; Ghrl, ghrelin;

Oxt, oxytocin. Cell types: NE, neuroendocrine (PNEC); AT1, alveolar epithelial cell, type 1; AT2, alveolar epithelial cell, type 2; Sm.M, smooth muscle; MyoF, myofibroblast; AdvF, adventitial fibroblast; AlvF, alveolar fibroblast; Peri, pericyte; Meso, mesothelial; Chondro, chondrocyte; Cap-a, capillary aerocyte; Cap, general capillary (Cap-g); Lym, lymphatic cell; B ZBTB32, B cells (ZBTB32+); Reg T, T cells (regulatory); T LY6G5B, T cells (LY6G5B+); NK, natural killer; T Alox5+, T cells (Alox5+); Neut, neutrophil; Baso, basophil; AlvMP, alveolar macrophage; IntMP, interstitial macrophage; pDC, plasmacytoid dendritic; mDC1, myeloid dendritic, type 1; mDC2, myeloid dendritic, type 2; CCR7+ DC, dendritic cell (Ccr7+); Mono(Class), monocyte (classical); Mono(NC), monocyte (non-classical); Mono(Int), monocyte (intermediate). (C) Photomicrograph of NEB in adult (PN57) *Ascl1*^{CreER}; *Rosa26*^{Zsgreen} mouse lung with PNECs (NE cells) labeled with Zsgreen (green) and probed by smFISH (PLISH) for *Angiotensinogen* (*Agt*) mRNA (red). Dashed white line, basement membrane. In close up of boxed region (right), note localization of *Agt* mRNA near apical (luminal) surface of PNECs. Scale bars, 10 μm (2 μm, right panel). Quantification (right) of *Agt*-expressing PNECs (n=87 PNECs scored in 2 animals). (D) NEB of adult (PN90) *Agtr1a*^{CreER} mouse with Cre-responsive reporter virus (AAV-*flex-tdTomato*) injected into vagal ganglion and co-stained for tdTomato (red, *Agtr1a*-expressing PSNs) and for CGRP (green, PNECs). *Agtr1a*-expressing PSNs penetrate basement membrane (dashed white line) of NEB and terminate at or near its apical surface. Scale bar, 10 μm. (E) Scheme for isolation and functional imaging of *Agtr1a*-labeled pulmonary sensory neuron (PSN) response to Agt II. (Top panel) Wheat germ agglutinin (WGA) was instilled into trachea of *Agtr1a*^{Cre};*Rosa26*$^{LSL\text{-}tdTom}$ mice to retrograde label PSNs and genetically label *Agtr1a*-expressing PSNs (tdTomato) in vagal ganglia. Ganglia were dissected, dissociated into single cells, and cultured overnight before loading with calcium indicator Fluo-4 and addition of Agt II. (Bottom) Diagram of labeled, dissociated vagal ganglion cells on confocal

microscope platform with perfusion chamber to flow in Agt II or KCl (positive control). Following imaging, cells were stained for PSN4 marker TRKA. (F) Neuronal traces showing evoked Ca^{2+} changes measured by Fluo-4 fluorescence (y-axis, fold change in fluorescence relative to baseline) of individual NEB-innervating PSNs (*Agtr1a*^{Cre}-lineage-positive (Agtr1a > Tdt+), TRKA+, WGA+; upper traces) or control ganglion cells from same preparation (lower traces) in response to Agt II (0.5-1 μM for interval indicated) or to KCl (50 mM). Graph scale, 25s (x-axis), 0.5 Fluo-4 ratio (y-axis).

Figure 5. Patterns of sensory gene expression in individual PNECs. (A) Patterns of expression from scRNA-seq profiles of sensory genes (colored dots, expressed genes) in individual PNECs (n=176). Histogram shows number of PNECs observed with the pattern of sensory gene expression below it. Values at right are percent of PNECs (NE) and other (non-NE) profiled cells that express the gene and the mean expression level for expressing PNECs. Sensory genes are grouped by modality: mechanosensing (Mech); thermosensing (□); oxygen sensing (O_2); acid-sensing (Acid); CO_2-sensing, carbonic anhydrases (CO_2), chemosensing (Chemo). Note the 56 different combinations of expressed sensory genes, with most PNECs predicted to be multimodal because they express more than one class of sensor. (B) Schematics of sensor and signal genes expressed by three individual PNECs. Number in center indicates sensory gene combination in panel A. Genes above number are expressed sensory genes, and genes below number are expressed signal genes with arrows indicating lung targets (cells that express receptor) or signals without identified lung targets that may enter circulation (circ) to target other organs. ?, signals without known receptors. Imm, multiple immune populations; PSN, pulmonary sensory neurons.

Figure 6. Expression of sensory and signaling genes in normal human PNECs and a pulmonary carcinoid. (A) Histogram of peptidergic gene expression in normal human PNECs

(n=55 PNECs from 2 subjects) by scRNA-seq. Bars show percent of PNECs with detected expression of the indicated peptidergic gene; values above bars are mean gene expression (ln (CPM+1)) in expressing cells. Peptidergic genes are listed in same order as orthologous mouse genes in Fig. 1E, and ellipses show percent of mouse PNECs that express the orthologous genes (from Fig. 1E, Table S4) for comparison. Among the 41 total peptidergic genes detected in human PNECs, the 14 genes at right ("human-specific") were expressed in human but not mouse PNECs; genes in parentheses were expressed only in mouse PNECs. *, previously known human PNEC peptidergic genes; blue font, classic hormone genes. Box and whisker plots (right) show quantification of peptidergic genes expressed per cell in human (H) and mouse (M) PNECs. Human values: 12.2 ±2.7 (mean ±SD; range 6-18, median 12, mode 12) expressed peptidergic genes. (B) Combinations of peptidergic genes expressed in individual human PNECs (n=55) from scRNA-seq. Histogram (top) shows number of PNECs expressing each of the 54 observed combinations of expressed peptidergic genes (dots, bottom). Values at right are mean expression (ln (CPM+1)) of indicated gene in expressing cells. *, individual PNECs diagrammed in panel C (#38, blue; #50, magenta). (C) Schematics (as in Figure 5) of sensor and signal genes expressed by individual human PNECs (combination #38, cell ID F21_B001235; #50, cell ID D5_B003269). Sensory and secretory genes expressed by cell #50 and by either cell #38 or the composite tumor cell (in panel E) are highlighted in magenta. (D) Histogram as in panel A of peptidergic gene expression in PNEC-like tumor cells of a human bronchial carcinoid (n=330 profiled tumor cells from one subject) from scRNA-seq. Pink bars show percent of tumor cells with detected expression of the indicated peptidergic gene; values above bars are mean gene expression (ln (CPM+1)) in expressing cells. To facilitate comparison, the percent of normal human PNECs that express each peptidergic gene (black ellipses, from panel A) and peptidergic gene expression in the single normal PNEC that most closely matches the tumor cell profile (magenta ellipses, PNEC #50 in panels B, C) are included in plot. Tumor-specific (spec),

peptidergic genes expressed in tumor but not normal PNECs. (E) Schematic (as in panel C) of sensor and signal genes expressed by a "composite" carcinoid tumor cell (T) from data in D. Genes shown are those expressed in >15% of the profiled tumor cells, plus *POMC* (14% of tumor cells), *INHBB* (14%), and *OPNSW1* (6%). Genes expressed by both "composite" tumor cell and normal PNEC #50 (panel C) are highlighted in magenta.

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. FACS enrichment strategy for PNECs. (A) Fluorescence-Activated Cell Sorting (FACS) gating path of adult (PN120) mouse lung cells with PNECs genetically lineage-labeled with ZsGreen and with immune (CD45+) and endothelial cells (CD31+) depleted by MACS prior to flow sorting as shown in Figure 1A. Red boxes, cell population selected at each gate; percentage values, percent of cells within the indicated gate. Gate 1, side scatter (SSC) and forward scatter (FSC) to separate cells (from debris). Gate 2, forward scatter width (FSC-W) and height (FSC-H) to separate single cells (from doublets and larger aggregates). Gate 3, DAPI to select viable cells, and APC to exclude endothelial (CD31+) and immune (CD45+) cells. Gate 4 (Q2), ZsGreen+ (*Ascl1* genetic lineage label) and EpCAM+ (epithelial) to enrich for neuroendocrine (NE) cells; alternative gates were also used at this step to sort control cells (Q1, $ZsGreen^{-}EpCAM^{+}$; Q4, $ZsGreen^{+}$ $EpCAM^{-}$). (B) Results of an analytic sort of 100,000 cells using above scheme. Count, number of cells within each gate. % of parent, % of analyzed cells at that sort; % of total, % of total analyzed cells (100,000). (C) Computational clustering using t-distributed stochastic neighbor embedding (t-SNE) of the expression profiles of 534 experimental and control cells sorted as above and analyzed by scRNA-seq. Identity of the cells (dots) in each cluster (colors) was assigned by cluster-selective expression of the canonical lung cell type markers indicated. (D) Number and fraction of cells obtained by each gating strategy. NE, neuroendocrine (PNECs); AT2, alveolar type 2; AT1, alveolar type 1; Endo, endothelial. (E) Violin plots comparing expression of top marker genes in common (core NE markers) between pulmonary neuroendocrine cells (PNECs) and tracheal neuroendocrine cells (TNECs), and the TNEC-selective markers. Cell types are indicated by tic mark on x-axis from left to right: NE, club, glial, stromal 1, stromal 2, AT2, basal, multi-ciliated, endothelial, AT1, and endothelial.

Figure S2. Peptidergic gene expression in PNECs. (A) Clustering of the rare POMC-expressing PNECs in a NEB. Fluorescence confocal micrographs (left) and schematics (right) of a NEB in adult PN68 wild type (CD-1) mouse lung immunostained for CGRP (green) and POMC (red). Two optical sections (separated by 65 μm) of the same NEB are shown. Dotted circles, individual PNECs; dashed lines, basement membrane. In schematic at right, green fill circles are CGRP+, POMC- PNECs, yellow fill circles are CGRP+ POMC+ double-positive PNECs, and gray fill circles are non-CGRP-expressing cells in cluster. Scale bar, 10 μm. **(B)** Integrated rarefaction (interpolation, solid line) and prediction (extrapolation to 500 PNECs, dashed extension) curve showing cumulative number of distinct peptidergic genes expressed by PNECs, as determined by bootstrapping with 200 re-samplings of the scRNA-seq profiles of 268 individual PNECs (176 flow-sorted and 92 captured by microfluidics). Grey shading, 95% confidence intervals. Rarefaction curve is based on detected expression but not magnitude of expression of each peptidergic gene in each profiled PNEC (Figure 2).

Figure S3. *Calca* alternative splicing in mouse PNECs. (A) Sashimi plots (top) as in Figure 3A showing mapped scRNA-seq reads for *Calca* gene and deduced alternative splicing patterns for the 20 PNECs in Figure 3B. Schematics (bottom) show structures of the resultant mRNAs encoding either CGRP or calcitonin, with exons (boxes) numbered and translational start site (arrowhead) and coding exons (black fill) indicated. Aqua, PNECs expressing only CGRP mRNA; magenta, PNECs expressing only calcitonin; blue, PNECs expressing both. (B) Fluorescence confocal micrograph of a NEB in adult (PN68) mouse lung immunostained for CGRP (green), calcitonin (red), E-cadherin (Ecad, white) to show cell boundaries, with DAPI nuclear counterstain (blue). Individual PNECs are numbered. Note that most PNECs in this NEB

express both CGRP and calcitonin, but PNEC 3 expresses only CGRP whereas PNECs 1 and 2 express calcitonin but little or no CGRP. Bar, 5 μm.

Figure S4. Mouse lung cell expression of receptors for PNEC neurotransmitters.

(A) scRNA-seq dot plot (reproduced from Figure 4 to facilitate comparison with panel B) of mean level of expression (dot heat map, ratio relative to maximum value for 10 pulmonary sensory neuron (PSN) subtypes, (Liu et al., 2021)) and percent of cells in population with detected expression (dot size) in the NEB-innervating PSN subtypes (PSN4, PSN7) of genes encoding receptors for each of the indicated PNEC neurotransmitters. (B) scRNA-seq dot plot of mean level of expression (dot intensity, ln (CPM+1)) and percent of cells in population with detected expression (dot size) in cell types (arranged by tissue compartment) in mouse lung cell atlas (Travaglini et al., 2020) of genes encoding receptors and associated subunits for indicated PNEC neurotransmitters (top). Cell types: NE, neuroendocrine (PNECs); AT1, alveolar epithelial cell, type 1; AT2, alveolar epithelial cell, type 2; Sm.M, smooth muscle; MyoF, myofibroblast; AdvF, adventitial fibroblast; AlvF, alveolar fibroblast; Peri, pericyte; Meso, mesothelial; Chondro, chondrocyte; Cap-a, capillary aerocyte; Cap, general capillary (Cap-g); Lym, lymphatic cell; B ZBTB32, B cells (ZBTB32+); Reg T, T cells (regulatory); T LY6G5B, T cells (LY6G5B+); NK, natural killer; T Alox5+, T cells (Alox5+); Neut, neutrophil; Baso, basophil; AlvMP, alveolar macrophage; IntMP, interstitial macrophage; pDC, plasmacytoid dendritic; mDC1, myeloid dendritic, type 1; mDC2, myeloid dendritic, type 2; CCR7+ DC, dendritic cell (Ccr7+); Mono(Class), monocyte (classical); Mono(NC), monocyte (non-classical); Mono(Int), monocyte (intermediate).

Figure S5. Mouse lung cell expression of renin-angiotensinogen pathway genes.

(A) Schematic of classical renin-angiotensinogen pathway showing sites of expression (tissue origin) of pro-hormone and processing enzymes, and hormone receptors (target sites) of the mature hormones (red). Some clinically significant anti-hypertensive drugs (blue) and the proteins they target are indicated. Renin (*Ren1*) catalyzes the first step in converting the pro-peptide angiotensinogen (*Agt*) to angiotensin I. In the lung, angiotensin converting enzyme (ACE, *Ace1*) converts angiotensin I to the active fragment, angiotensin II (Agt II). (B) scRNA-seq dot plots showing expression levels (dot intensity) and percent of cells in population with detected expression (dot size) in mouse lung cells (Travaglini et al., 2020) of indicated genes in the angiotensin pathway.

Figure S6. Mouse lung cell expression of sensory genes. (A) Heatmaps showing expression levels (ln (CPM+1)) in individual PNECs (Neuroendocrine, n=176) and other lung cell types as indicated (see Figure S1). Sensory genes are grouped by sensory modality; only olfactory receptor genes (Olfr) expressed in PNECs are shown in heatmap. Genes of the NADPH oxidase complex previously proposed to participate in PNEC hypoxic response: *Cybb*/p91phos, *Cyba*/p22phox, *Ncf2*/p67phox, *Rac2*/Rho GTPase. *Mcoln3*, a TRP family channel, is also expressed in auditory hair cell stereocilia, but role in mechanosensation not established. (B) Heatmaps as above showing expression of voltage and ligand-gated potassium and calcium channels. K2P, potassium 2-pore channel; Kv, voltage-gated potassium channel; CNG-gated, cyclic nucleotide-gated channel; Ca2-act, calcium-activated channel. Note the many PNEC-specific ion channel genes identified including *Kcnc3*, *Kcnk16, Kcnc1, Kcnc2, Kcnb1, Kcnv1, Kcnf1, Kcnq5, Kcnh2, Kcnh6,* and *Kcnh7.* (C) Heatmap as above showing expression of opsin genes. MC, multiciliated cell; AT2, alveolar epithelial type 2 cell; AT1, alveolar epithelial type 1 cell; Stromal 1 and Stromal 2, stromal cell populations. Endo, endothelial cells; K2P, potassium

2-pore channels; Kv, voltage-gated potassium channels, CNG-gated, cyclic nucleotide-gated channels; Ca^{2+} act, calcium-activated channels.

Figure S7. Mouse lung cell expression of TNF family ligand and receptor genes.

(A) scRNA-seq dot plots showing level of expression (dot intensity, ln (CPM+1)) and percent of cells in population with detected expression (dot size) of TNF family receptor genes. Note PNEC (NE cell) expression of *Tnfrsf12a* (orange highlight) and *Tnfrsf21* (blue). (B) scRNA-seq dot plots showing expression as above of TNF ligands. Note the expression of cognate ligands for the TNF receptors expressed in PNECs, *Tnfsf12* (orange, ligand of Tnfrsf12a) and *App* (blue, ligand of Tnfrsf21). Cell types: NE, neuroendocrine (PNECs); AT1, alveolar epithelial cell, type 1; AT2, alveolar epithelial cell, type 2; Sm.M, smooth muscle; MyoF, myofibroblast; AdvF, adventitial fibroblast; AlvF, alveolar fibroblast; Peri, pericyte; Meso, mesothelial; Chondro, chondrocyte; Cap-a, capillary aerocyte; Cap, general capillary (Cap-g); Lym, lymphatic cell; B ZBTB32, B cells (ZBTB32+); Reg T, T cells (regulatory); T LY6G5B, T cells (LY6G5B+); NK, natural killer; T Alox5+, T cells (Alox5+); Neut, neutrophil; Baso, basophil; AlvMP, alveolar macrophage; IntMP, interstitial macrophage; pDC, plasmacytoid dendritic; mDC1, myeloid dendritic, type 1; mDC2, myeloid dendritic, type 2; CCR7+ DC, dendritic cell (Ccr7+); Mono(Class), monocyte (classical); Mono(NC), monocyte (non-classical); Mono(Int), monocyte (intermediate)

Figure S8. Human PNEC markers. (A) Most sensitive PNEC-specific markers identified by scRNA-seq (SS2) of human PNECs. Markers are ranked by ratio of their expression (ln (CPM+1)) in PNECs (NE cells) to that of all other epithelial cells in human lung cell atlas (Travaglini et al., 2020). *, previously known PNEC marker. Red highlight, human-specific PNEC markers. (B) Violin plots comparing expression of most sensitive newly identified marker

genes (left, *GRP, SCGN, SCG5, BEX1*) to the four PNEC markers used most often clinically (right). Note Secretogogin (*SCGN*) is a human-specific and highly sensitive and specific PNEC marker. *GRP,* gastrin-releasing peptide; *SCGN*, secretogogin; *SCG5*, secretogranin 5; *BEX1*, brain-expressed 1; *CHGA,* chromogranin A; *SYP*, synaptophysin; *INSM1*, insulinoma-associated 1; *ASCL1*, achaete-scute like homolog-1. Log-transformed expression values, ln (CPM + 1).

Figure S9. Alternative splicing of CALCA transcripts in human PNECs. (A) Sashimi plots as in Figure 3 for 20 randomly-selected *CALCA*-expressing human PNECs. Schematics (bottom) show structures of human reference mRNAs encoding either CGRP or calcitonin, with exons (boxes) numbered, and translational start site (arrowhead) and coding exons (blue fill) indicated. PNECs 1-7 expressed only calcitonin mRNAs and PNECs 8-20 expressed both calcitonin and CGRP mRNAs. (B) Quantification of *CALCA* alternative splicing in A. Note none of the human PNECs expressed exclusively CGRP mRNA, in contrast to mouse where most cells express both CGRP and calcitonin mRNA or exclusively CGRP (Fig. 3).

Figure S10. Human lung cell expression of receptors for human PNEC neuropeptides, peptide hormones, and neurotransmitters. (A) scRNA-seq dot plots showing expression level (dot intensity) and percent of indicated human lung cell types with detected expression (dot size) of indicated receptor genes for PNEC peptidergic signals and neurotransmitters. Expression values from human lung cell atlas (Travaglini et al., 2020). The corresponding ligands for each receptor are shown above the receptor gene and color matched. Some receptor genes are subunits for more than one hormone (gray lettering) or not expressed in any cells in the human lung cell atlas (black lettering). No expression was detected for 17 of the 33 peptidergic PNEC signals including 10 of the 14 human-specific peptidergic signals (asterisks). Values (at right), number of PNEC peptidergic signal receptors expressed in at least 10% of the cells of that cell

type. (B) Dotplots of receptor genes for neurotransmitters predicted to be made by human PNECs. Histamine receptors (grey) are also included because histamine is predicted to be made by mouse (though not human) PNECs (see Table S7). Top row (highlighted pink) shows neurotransmitter receptors expressed in human PNECs, predicting autocrine signaling. Cell types: NE, neuroendocrine (PNECs); Cil, ciliated; dBas, differentiating basal; Ion, ionocyte; AT1, type 1 alveolar epithelial cell; AT2, type 2 alveolar epithelial cell; AT2-s, AT2 signaling; ASM, airway smooth muscle, VSM, vascular smooth muscle; MyoF, myofibroblast; FibroM, fibromyocyte; AdvF, adventitial fibroblast; AlvF, alveolar fibroblast; LipF, lipofibroblast; Peri, pericyte; Art, artery; Cap-a, aerocyte; Cap, general capillary (Cap-g); Cap-i1, capillary intermediate 1 cell; Bro1, bronchial vessel 1 cell; Lym, lymphocytes; B, B cells; CD8 M/E, CD8 + memory/effector T; CD8 Na, CD8+ Naïve T; CD4 M/E, CD4+ memory/effector T; CD4 Na, CD4+ Naïve T; NKT, natural killer T cell; NK/T-p, proliferating natural killer/T cells; Neut, neutrophils; Mast Ba 1; MP, macrophage; pDC, plasmacytoid dendritic cell; mDC2, myeloid dendritic cell 2; DC IGSF21, IGSF21+ dendritic cell; Mono Class, classical monocyte; Mono NC, non-classical monocyte; Mono Int, intermediate monocyte.

Figure S11. Patterns of sensory gene expression in human PNECs. (A) Patterns of expression of sensory ion channels and other sensory genes (dots, expressed genes) in individual PNECs from scRNA-seq profiles (n=55 PNECs). Columns show the 40 identified combinations of expressed sensory genes, and histogram (top) shows number of profiled PNECs expressing each combination. Note sensory gene combination 25 (highlighted in magenta, cell ID, N1_B00124) is a PNEC that appears specialized for oxygen sensing because it expresses *OR51E1* and *OR51E2*, a close family member and ortholog of Olfr78 implicated in acute hypoxia-sensing in mouse carotid body (Chang et al., 2015), and acid-sensitive channel *KCNK3* but no mechanosensory channels. Values at right are percent of PNECs (NE) and other (non-NE)

profiled cells that express the gene. Mech, mechanosensory; Therm, thermosensing; O_2, oxygen-sensing; Acid, acid-sensing; CO2, carbonic anhydrases; Chemo, chemosensory genes. (B) Heatmaps showing log-transformed expression values (ln (CPM+1)) for the sensory genes in panel A in the 55 PNECs (columns).

Figure S12. Human lung cell expression of sensory genes. (A) Heatmaps showing log-transformed expression values (ln (CPM+1)) for selected genes of the indicated sensory modalities in PNECs and other human lung epithelial cells (Travaglini et al., 2020). Cell types: NE, neuroendocrine (PNECs); B, basal; G, goblet; Multiciliated, multiciliated; AT2, alveolar type 2 cells; AT2-s, alveolar type 2 signaling; AT1, alveolar type 1; Mechano, mechanosensory; O2, oxygen-sensing; Acid, acid-sensitive channel; CO_2, carbonic anhydrases. Oxygen sensory genes include genes of the NADPH-oxidase complex previously implicated in PNEC oxygen sensing. See Table S7 for full set of gene names and annotations.

Figure S13. Expression of peptidergic genes in human lung carcinoids profiled by bulk RNA-seq. Heatmap showing log-transformed, quantile-normalized expression levels of all human peptidergic genes, and for comparison the top PNEC markers indicated, in 74 histopathologically "typical" and 37 "atypical" human lung carcinoids transcriptomically profiled by bulk RNA-seq (patient identifiers at right) (Alcala et al., 2019; Fernandez-Cuesta et al., 2014; Laddha et al., 2019). Histogram shows fraction of normal human PNECs that express each peptidergic gene (from Fig. 6A). Magenta font, peptidergic genes with prominent expression in at least some patients; gray font, uniform but low/moderate expression of peptidergic genes, which could derive from tumor-infiltrating cells rather than the tumor cells themselves (e.g., *Apln* could derive from tumor-infiltrating capillary cells (Gillich et al., 2020)). *, peptidergic

genes previously reported in lung carcinoids. **, peptidergic genes associated with established endocrinopathies of lung carcinoids. Prev., previously

METHODS

RESOURCE AVAILABILITY

Data and code availability

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead author Christin Kuo (ckuo@stanford.edu) or lead contact Mark Krasnow (krasnow@stanford.edu). All raw and processed data with accompanying metadata from scRNA-seq of mouse and human PNECs will be submitted to the Gene Expression Omnibus (GEO) database www.ncbi.nlm.nih.gov/geo. Raw data consists of fastq files corresponding to paired-end reads for each cell, and processed data is in the form of a raw gene counts matrix. All code used to generate data objects and plots is available at github: https://github.com/sdarmanis/Neuroendocrine_scRNA-seq

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals

Mouse lines used were: wild-type mouse strains CD-1 and C57BL/6NJ, aged 2-4 months, tamoxifen-inducible, PNEC-specific knock-in Cre recombinase driver *Ascl1*^{CreERT2} (Kim et al., 2011a) and Cre-dependent fluorescent reporter *Rosa26*^{Zsgreen1} (Madisen et al., 2009), and knock-in Cre recombinase drivers *Agtr1a*^{Cre} (de Kloet et al., 2017), and *Agtr1a-2A-Cre* (Leib et al., 2017). Mice were maintained in 12h light/dark cycle with food and water provide ad libitum. Genotyping was performed on tail clips with oligonucleotide primers as described (Kim et al., 2011a). All animal husbandry, maintenance, and experiments were performed in accordance with Stanford University's IACUC-approved protocols (APLAC 9780, 32092).

Mouse PNEC labeling, enrichment, and isolation

Mouse PNECs were permanently labeled with ZsGreen by intaperitoneal injection of PN15 (n=4 mice, gender-balanced), PN85 (n=2), and PN115 (n=2) *Ascl1*$^{CreERT2/+}$;*Rosa26*^{Zsgreen1} mice with 3

mg tamoxifen (MilliporeSigma T5648; prepared at 20 mg/ml in corn oil, stored at -20°C) to induce Cre-ERT2, and 3-7 days later (PN21, PN90, and PN120) euthanized by CO_2 inhalation and cervical dislocation. For the PN21 time point, a second dose of tamoxifen (3 mg) was provided to the dam at E13.5 when PNEC progenitors robustly express *Ascl1* (Kuo and Krasnow, 2015). Immediately after euthanasia, lungs (unperfused) were micro-dissected en bloc at room temperature, and the trachea and peripheral regions of each lobe removed with a razor blade. The remaining bronchiolar regions (where PNECs reside) of individual lobes were minced with a razor blade and then digested in DMEM/F12 media containing dispase at 1 U/ml (StemCell Technologies, 07923), type 4 collagenase at 10 U/ml (Worthington Biochemical, CLS-4), and elastase at 3 U/ml (Worthington Biochemical) at 37ºC for 30-60 minutes. Samples were manually triturated with a 1000 ul micropipet tip every 10 minutes during the incubation to generate a uniform cell suspension. An equal volume of cold (4°C) FACS buffer [phosphate buffered saline (PBS, 137 mM NaCl, 2.7 mM KCl, 2 mM ethylenediaminetetraacetic acid (EDTA), 8 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.4) with 10% fetal bovine serum (Thermo Fisher 10082147)] was added to quench the enzymatic reactions. All subsequent steps were carried out at 4°C. DnaseI (Stem Cell Technologies, 07900) was added to a final concentration of 5 ug/ml with mixing by tube inversion ten times during the 5 minute incubation. Cell suspensions were then filtered sequentially through 100 um (Corning 431752) and 40 um mesh filters (Corning 352340) to remove cell aggregates, then centrifuged at 400xg for 5 minutes. Cell pellets were resuspended in 2 ml of 1x RBC lysis buffer (BD Biosciences 555899) and incubated for 5 minutes to deplete red blood cells (RBCs). FACS buffer (8 ml) was added to terminate the reaction, and the suspension centrifuged at 400xg for 5 minutes, and cell pellets were resuspended in FACS buffer to 10^5 - 10^6 cells/ml. To deplete endothelial cells and monocytes, 10 ul anti-CD31 MicroBeads (Miltenyi Biotec, 130-097-418) and 10 ul anti-CD45 MicroBeads (Miltenyi Biotec, 130-052-301) were added to 1 ml of cell suspension, incubated for 15 min, and

loaded on an LD MACS column (Miltenyi Biotec, 130-042-901) pre-equilibrated with FACS buffer, according to manufacturer's instructions. The column flow-through was centrifuged at 400xg for 5 minutes, and the cell pellet was resuspended in 0.8 - 1 ml of FACS buffer to 10^6 -10^7 cells/ml.

The cell suspensions were incubated with the following antibodies: allophycocyanin (APC)–conjugated anti-mouse CD31 (BioLegend, 102409, 1:800 dilution), APC-anti-mouse CD45 (BioLegend, 103111, 1:800), APC-anti-mouse F4/80 (BioLegend 123115, clone BM8, 1:800 dilution), phycoerythrin and cyanin 7 (PE-Cy7)-conjugated anti-mouse CD326 (anti-EpCam) (Thermo Fisher Scientific, 25-5791-80, clone G8.8, 1:400 dilution). After 15 min, cells were centrifuged at 300xg x 5 minutes, the cell pellet was resuspended in 1 ml FACS buffer, and this wash step was repeated twice. The final cell suspension was flow-sorted in a FACSAria^{TM}II (Becton-Dickinson) using the indicated sorting gates (Fig. S1). Cells were collected in 96-well plates (BioRad, HSP9631) containing 4 ul per well of cell lysis buffer containing 4U RNase Inhibitor (Takara Bio, 2313A), 0.05% Triton X100 (Thermo Fisher), 2.5mM dNTPs (Thermo Scientific) and 2.5uM oligo-dT_{30}VN (AAGCAGTGGTATCAACGCAGAGTACT30 VN-3′) (IDT) as previously described for Smart-seq2 (Picelli et al., 2014), where ‘V’ represents A,C,G, ‘N’ represents A,T,C, or G, and the synthesized product contains a mix of all possible combinations in approximately equal proportions (variance up to 10%). Plates with sorted cells were sealed with microplate sealing film, vortexed 3-5 seconds, centrifuged at 1000xg for 1 minute, and immediately placed on dry ice and stored at -80^{o}C until complementary DNA (cDNA) generation and sequencing.

Single cell mRNA sequencing

RNA from individual sorted cells was reverse transcribed to cDNA amplified, and Illumina sequencing libraries prepared as previously described (Darmanis et al., 2017). Briefly, 96-well plates containing single-cell lysates were thawed on ice, heated to 72^{o}C for 3 minutes and

immediately put back on ice. For cDNA synthesis, 6 ul of reverse transcriptase mix (1X First-Strand Buffer (Takara Bio, 639538) with 100U SMARTScribe™ Reverse Transcriptase (Takara Bio, 639538), 10U Recombinant RNase Inhibitor (Takara, 2313A) , 8.5mM DTT (Invitrogen, P2325), 0.4mM Betaine (Sigma, B0300-5VL), 10mM $MgCl_2$ (Invitrogen, AM9530G) and 1.6uM template switching oligonucleotide containing one locked nucleic acid-modified guanosine (+G) at 3'end (5′-AAGCAGTGGTATCAACGCAGAGTACATrGrG+G) (Exiqon)) were added to each well, and the reactions were incubated at 42ºC for 90 minutes followed by 70ºC for 5 minutes. For PCR amplification of cDNA, 15 ul of PCR mix (1x KAPA HiFi HotStart ReadyMix (Kapa Biosystems, KK2602) with 0.16uM 1SPCR (one step PCR) oligonucleotide (5′-AAGCAGTGGTATCAACGCAGAGT) (IDT) and 0.56 U of Lambda Exonuclease (New England Biolabs, M0262L)) was added to each well, followed by thermal-cycling at: (i) 37ºC for 30 min, (ii) 95ºC for 3 min, (iii) 21 cycles of 98ºC for 20s, 67ºC for 15s and 72ºC for 4 min, and (iv) 72ºC for 5 min. Amplified cDNA was purified using 0.7x AMPure XP beads (Beckman Coulter, A63880) then analyzed by capillary electrophoresis on a Fragment Analyzer (Advanced Analytical Technologies) and the concentration of cDNA (in fragment size range 500 to 5000 bp) adjusted and Nextera DNA sequencing libraries prepared as described (Darmanis et al., 2015). Libraries from wells on each plate were pooled using a Mosquito liquid handler (SPT Labtech), purified twice using 0.7x AMPure XP beads, and library pool quality assessed on a Fragment Analyzer. Libraries from 679 single cells were sequenced (75 bp paired-end reads) on a NextSeq 500 (Illumina) using High-output v2 kits (Illumina). Raw sequence reads were demultiplexed using bcl2fastq (v1.8.4, Illumina), and remaining sequences aligned to the mouse reference genome (GRCm38-mm10, UCSC, supplemented with Zsgreen1 sequence) with STAR (v2.5.2b, default parameters except Stranded set to false and Mode set to intersection-nonempty), and the number of reads that mapped to each annotated gene (gene counts) determined with HTSEQ (v0.6.1p1, default parameters except Stranded set to false and Mode set to intersection-

nonempty) (Anders et al., 2015). Cells with less than 50,000 mapped reads or less than 1000 detected genes were excluded as a quality metric, leaving 534 cell expression profiles for further analysis.

Analysis of CALCA alternative RNA splicing in PNECs

Sequence reads from both mouse and human PNEC scRNA-seq datasets were aligned to mouse mm10 and human gh38 reference genomes, respectively, using STAR, and the BAM (binary compressed version of sequence alignment map) output visualized by sashimi plots of the Calca (mouse) and CALCA (human) genomic loci using Integrative Genomics Viewer (IGV v2.4.14) (Katz et al., 2010; Robinson et al., 2011) and the mouse CGRP (RefSeq ID: NM_001289444) and calcitonin (NM_001305616), and human CGRP (NM_001033953) and calcitonin (NM_001741) reference mRNA sequences.

Validation set of mouse PNEC scRNA-seq profiles

A second set of adult mouse PNECs that were lineage-labeled, purified, and profiled by scRNA-seq as described (Ouadah et al., 2019) was used as validation set. Briefly, PNECs were lineage-labeled by tamoxifen administration to adult (age 2-3 months) *Ascl1*^{CreERT2};*Rosa26*$^{LSL\text{-}ZsGreen}$ or *CGRP*^{CreERT2};*Rosa26*$^{LSL\text{-}ZsGreen}$ mice, and whole lungs (excluding trachea) were processed into a single cell suspension. Red blood cells were lysed, and endothelial and immune cells depleted using MACS. Lineage-labeled epithelial cells (ZsGreen^{+} EpCam^{+}) were sorted by FACS into a single collection tube, and individual cells captured and cDNA generated using an integrated microfluidic platform (Fluidigm C1). cDNA sequencing libraries were prepared in 96 well format and sequenced on a NextSeq 500 (Illumina) device, and obtained sequences were demultiplexed, processed, aligned to individual genes, and quantified to define gene expression levels in each cell. PCA analysis was performed on cell expression patterns using highly-variably expressed genes, and identities of cell clusters with related expression patterns assigned based on enriched expression of canonical lung cell type markers. Of the 100 PNEC expression profiles

obtained, the eight with "transitional" profiles were excluded from our analysis. A threshold of 5 transcripts/million (TPM) was used for determining if a gene was expressed.

Human PNECs and carcinoid scRNA-seq analysis

The expression profiles of human PNECs characterized here are from our scRNA-seq analysis of cells from histologically normal lung tissue obtained from therapeutic lobectomies and matched blood from three patients with focal lung tumors; these profiles were used to construct our comprehensive molecular cell atlas of the human lung (Travaglini et al., 2020). Among the profiled cells, a cluster of 66 PNECs was identified by their selective expression of classical markers CALCA and ASCL1. For our PNEC analysis, we excluded all 11 human PNEC profiles obtained by droplet-based 10X scRNA-seq, which had less extensive expression profiles than the ones profiled by SS2, the plate-based method used here to profile mouse PNECSs. We also excluded one SS2-profiled cell that was designated a PNEC (cell ID: C7_B002464.gencode.vH29) but was an outlier in the original PNEC cluster (Travaglini et al., 2020); we found it expressed only one PNEC neuropeptide (DBI) but not any classic PNEC markers (CALCA, CHGA, ASCL1, GRP) or our newly identified PNEC markers (SCGN, PCKS1N, SCG3, SCG5), so it is likely a related but distinct and extremely rare lung cell type. We included in our analysis one SS2-profiled cell that was not originally designated a PNEC (Cell ID: H4.B002460.gencode.vH29), which we found expressed both classic and new PNEC markers (ASCL1, GRP, CHGB, SCGN, SCG2, SCG3, SCG5). In total, our analysis included 55 PNECS, 50 from patient 1 and 5 from patient 3. The SS2 scRNA-seq sequencing reads from these 55 PNECs were re-aligned to the primary assembly of human reference genome GRCH38 (and further analyzed as above), to exclude an alternative contig at the CHGA locus (contig KI270847.1) in the reference genome used in the original analysis (GRCH38.p12) that caused vast undercount of CHGA expression.

The scRNA-seq expression profiles of human carcinoid cells characterized here are from a parallel analysis of one of the tumors in the above study, a typical carcinoid (1.3 x 0.9 cm) in the left bronchus of the resected left lower lung lobe of patient 3, a 51 year-old female mild adult-onset asthma and recent worsening hypertension. The tumor sample was processed and profiled in parallel with the accompanying normal tissue from this patient, and the 330 cells described here are sorted cells from the epithelial compartment of the tumor sample that were analyzed by SS2 and identified as carcinoid cells by their abundance in the tumor sample and expression of many classic PNEC markers and peptidergic genes, consistent with the underlying clinical and pathological diagnosis. A full description of the carcinoid tumor expression data will be provided elsewhere.

Peptidergic and sensory genes

The comprehensive list of mouse and human neuropeptide and peptide hormones and their genes ("peptidergic genes") (Table S4) and receptors were compiled from the literature (Kim et al., 2011b; Secher et al., 2016) and an online database (www.neuropeptides.nl), then verified and updated with newly identified receptors by PubMed literature searches (through July 2020) for each included neuropeptide and peptide hormone.

The comprehensive list of sensory genes (Tables S6, S7) was curated from literature reviews of each sensory modality including mechanosensors (Clapham, 2003; Ranade et al., 2015), thermosensors (Caterina et al., 1999; McKemy et al., 2002; Peier et al., 2002; Vandewauw et al., 2018), acid sensors (Lin et al., 2004; Tominaga et al., 1998; Waldmann et al., 1997), hypoxia sensors (Chang et al., 2015; Kumar and Prabhakar, 2012), olfactory receptors (Buck and Axel, 1991), pheromone receptors (Dulac and Axel, 1995), trace amine-associated receptors (Zucchi et al., 2009), taste receptors (Chandrashekar et al., 2006), and opsins/light sensors (Blackshaw and Snyder, 1999; Haltaufderhyde et al., 2014; Terakita, 2005), and includes all genes with biochemical, genetic, or functional data to support their role as sensors plus related

members of the gene family including full families of ion channels (Yu and Catterall, 2004). We included all genes previously implicated in PNEC sensory functions including hypoxia-sensing (Buttigieg et al., 2012; Fu et al., 2000) and mechanosensory (Lembrechts et al., 2012, 2013) pathways and genes.

Mouse lung immunohistochemistry and in situ hybridization

For immunohistochemistry, adult wild type CD-1 or C57BL/6NJ mice as indicated were euthanized as above, and lungs were perfused with room temperature PBS and then inflated with 2% low-melting point agarose (ThermoFisher, UltraPure 16520050). Individual lobes were isolated, fixed at 4°C for 18 - 24 hrs in 4% paraformaldehyde (PFA) in PBS, cryoprotected in 30% sucrose/PBS solution, transferred to cryomold blocks (22 x 40 x 20 mm, VWR) and embedded as entire lobes in Optimal Cutting Temperature (O.C.T.) Compound (Tissue Tek), and stored at -80°C until sectioning. Frozen tissue blocks were sectioned with a cryostat (Leica Biosystems) and the sections (20-50 μm) washed in PBS with 0.1% Tween-20 (PBST) and then incubated with blocking solution (5% goat serum, 0.3% Triton X-100 in PBS) for 1 - 5 hrs, then washed in PBST. Washed sections were incubated with primary antibodies overnight at 4°C, washed with PBST, and then incubated with secondary antibodies at room temperature for 45 minutes followed by counterstaining with DAPI at 0.1 μg/ml in PBS to mark nuclei. Primary antibodies were: anti-PC1/Pcsk1 (rabbit, Abcam ab3532, used at 1:750 dilution), anti-CGRP (rabbit, Sigma C8198, 1:500), anti-POMC (rabbit, Phoenix Pharmaceuticals H-029-30, 1:500), anti-Cartpt (rabbit, Phoenix Pharmaceuticals H-003-62, 1:2000). Secondary antibodies were directly conjugated to Alexa-488, -555, or 633 (Invitrogen) or to Alexa 647 (donkey anti-rat, Jackson ImmunoResearch) and used at 1:250 dilution.

For multiplex single molecule FISH (smFISH), wild type mouse lungs were perfused, inflated, fixed, imbedded in O.C.T. Compound, and stored as above. Cryosections (12 μm) were

probed by RNAscope Hiplex12 technology (Advanced Cell Diagnostics, 324140) according to manufacturer's instructions. The proprietary RNAscope probes were: Mm-Agt-T1 (426941-T1), Mm-Nmb-T2 (459931-T2), Mm-Adcyap1-T3 (405911-T3), Mm-Cartpt-T5 (432001-T5), Mm-Pomc-T6 (314081-T6), Mm-Chga-T9 (447851-T9), Mm-Resp18-T11 (493871-T11), Mm-Calca-T7 (custom probe), and Mm-Scg5-T10 (custom probe). Probed sections were imaged by confocal fluorescence microscopy (Zeiss LSM 880, Airyscan mode), and images were aligned using RNAscope HiPlex Registration software and processed with Zen software (Zeiss). To resolve secretory vesicles immunostained for peptides (Figs. 3C, G), confocal images were acquired in super-resolution mode.

In vitro imaging of mouse pulmonary sensory neuron response to angiotensin

Pulmonary sensory neurons (PSNs) were prepared from adult (PN120) *Agtr1a*$^{Cre/+}$;*Rosa26*$^{LSL\text{-}tdTom/+}$ mice that selectively label the two types of PSNs that innervate NEBs (Liu et al, unpublished data). Three to five days prior to PSN isolation, 50 ul (1 mg/mL) of a fluorescent wheat germ agglutinin (WGA-647, Thermo Fisher, W32466) was instilled into the trachea to retrograde label PSNs. After 3-5 days to allow for WGA uptake by PSNs and retrograde transport to their cell bodies in the tenth (vagus) cranial nerve ganglia, mice were euthanized as above and the vagal ganglia were dissected and immediately placed in cold-buffered Hanks Balanced Salt Solution without calcium or magnesium (HBSS, ThermoFisher, 14190144). Ganglia were digested with 60 U papain (Worthington Biochemical, LS003126) in 1 ml HBSS (with 10 mM HEPES pH 7.4, 0.5 mM EDTA, and 0.4 mg/mL L-cysteine) for 10 minutes at 37°C. The papain solution was then replaced with 3 ml of a second enzymatic digestion solution (1.5 mg/ml collagenase IV (Worthington, LS004186) and 1 mg/ml dispase (Worthington, LS02109) in HBSS with 10 mM HEPES) and the incubation continued at 37°C for 30 minutes, with tube inversion 5 times every 10 minutes. The sample was centrifuged for 4 minutes at 400g, and the pelleted cells were resuspended in 1 ml L-15 medium (Gibco 11415) with 10 mM

HEPES (pH 7.4) by 3 sequential rounds of manual trituration using custom pulled glass micropipettes (Sutter Instrument Company, Model P-87) of successively finer tip diameter (to final range 0.1-0.12 mm). Tips were pre-coated with complete L-15 medium with 10% FBS (Gemini Bio Products, 100-50, diluted in 10 mM HEPES, pH 7.4) to limit neuronal loss. The cell suspension was gently layered on 5 ml of 20% Percoll (Sigma, P4937) in L-15 medium (Gibco, 11415) and centrifuged for 9 min at 400xg to separate dissociated neurons from lower density connective tissue and smaller cells. The cell pellet was resuspended in 2 ml of L15 medium with 10 mM HEPES (pH 7.4), giving a typical yield of ~1000 cells. Cells were transported to the imaging facility at room temperature and centrifuged for 3 minutes at 750xg. Cells were re-suspended in 100 ul of warm CO_2-equilbrated DMEM/F12 medium (Gibco, 10565018), and 30 - 40 ul of the cell suspension were plated in the center of a laminin-coated inset of a poly-lysine-coated 12 mm circular Nunc glass bottom culture dish (Thermo Fisher, 150680), prepared as described below. Cells were incubated at 37˚C for 60-90 minutes to initiate cell adherence to the inset; although only some neurons adhere during this period, incubations beyond 2 hours caused evaporation of the small volume of medium and decreased neuron viability. DMEM/F12 medium (500 ul) was added to each well, and the cultures were incubated at 37˚C (with 5% CO_2) for 12-16 hours to increase cell adherence and equilibrate cells to the culture environment prior to functional imaging. In healthy preparations, typically ~100 - 150 neurons adhered to the inset and ~10% formed extended projections. [We found that coating of the culture insets with fresh reagents as follows was critical for cell adherence and viability: Poly-lysine coating of clean Nunc glass bottom dishes was done by incubating the dish in a solution of 50 ug/ml poly-D-lysine (Millipore, A-003-E) in HBSS at 37ºC overnight, then washing the dish with HBSS three times and removing residual solution by aspiration; laminin coating of the insets was done by covering insets with a solution of 20 ug/ml laminin (Sigma L2020) in HBSS at 4ºC for at least 45

min (typically 3-4 hours), then carefully removing the laminin solution by aspiration, washing the inset 3 times with HBSS at 4°C, and leaving the inset covered in HBSS until cell plating.]

For functional imaging, cells were loaded with fluorescent calcium indicator Fluo-4 by incubating cells with 10 um Fluo-4 (Invitrogen) in HBSS buffered with 10 mM HEPES (pH 7.4) for 15-20 min. Nunc glass bottom dishes with insets containing cultured neurons were placed on a perfusion chamber platform (Warner Instruments, RC-37W) on a Zeiss 880 LSM confocal microscope stage housed within an incubation chamber adjusted to 37°C and 5% CO_2; platform perfusion was by gravity-dependent flow controlled with a stopcock. Cells were continuously perfused with HBSS buffered with 10 mM HEPES (pH 7.4), and fields containing retrograde-labeled pulmonary sensory neurons were identified by WGA-647 fluorescence. Calcium imaging data (Fluo-4 flourescence) were collected every second with 488 nm wavelength excitation and 500-550 nm emission. After baseline recording for 60 seconds, 500 nM Angiotensin II peptide (Sigma-Aldrich, A9525) in HBSS with 10 mM HEPES (pH 7.4) was perfused for 60 seconds, followed by a wash and re-equilibration with HBSS buffered with 10 mM HEPES (pH 7.4) for 3 minutes, and finally a 15 second infusion of 50 mM KCl in HBSS buffered with 10 mM HEPES (pH 7.4) to assess cell excitability/viability. Following the infusions, the dish was removed from the perfusion chamber and cells immediately fixed with 4% PFA at 4°C for 30-60 min for subsequent immunohistochemistry as above to confirm identity of the monitored neurons. For analysis of the time-lapse recordings, Image J software (v2.0) was used to define cell boundaries and determine the mean fluorescence intensity value for each cell in the imaging field at each time point; the obtained fluorescence values were normalized to the average baseline fluorescence value (prior to Angiotensin II exposure) for the cell.

QUANTIFICATION AND STATISTICAL ANALYSIS

Computational clustering and identification of mouse PNEC scRNA-seq profiles

Counts for each gene were normalized across cells, scaled per million and converted to logarithmic scale. Dimensionality reduction was used to compare and cluster the obtained cell expression profiles using Seurat v2.3.4 (Butler et al., 2018). First, genes with highly variable expression across the sample population were identified ('FindVariableGenes'), selecting genes with >1 standard deviation dispersion in mean expression values. Second, the dimensionality of the expression matrix data for the highly variable genes was reduced using principal component analysis (PCA), and the significant principal components (PCs) that captured the majority of variation in the dataset were selected by their standard deviations (PCElbowPlot function) and by examining the top gene loadings in each component as heatmaps. We selected the first fifteen PCs, and the five genes with the highest PC scores along each PC were inspected for biological relevance and for canonical markers of known lung cell types. Then, the relatedness of cell expression profiles was visualized in two dimensional tSNE plots ('RunTSNE for R' with perplexity=30). Third, genes enriched in each cluster of cells with similar expression profiles were identified using Wilcoxon rank sum test with multiple testing correction ('FindAllMarkers'). Cell doublets were identified and removed using Scrublet (Wolock et al., 2019).

The identities of the PNEC cell cluster and the 9 other obtained cell clusters were assigned based on enriched expression of canonical lung cell type markers: PNECs (*Calca, Ascl1*, *Chga*), multiciliated cells (*Foxj1*, *Ccdc153*, *Cdhr3*), basal cells (*Krt5, Trp63, Krt15*), AT1 cells (*Ager*, *Rtkn2*), AT2 cells (*Sftpc*, *Sftpb*), club cells (*Scgb3a2*, *Scgb1a1*), endothelial cells (*Pecam*, *Tie1*), stromal cell populations 1 and 2 (*Col1a1, Col1a2*), and glial cells (*Gfap*, *S100b*, *Plp1*). The expression profiles of the 176 obtained high quality PNECs were used for all analyses.

Analysis of peptidergic gene expression in human carcinoids profiled by bulk RNA-seq

Bulk RNA-seq datasets of 239 lung neuroendocrine neoplasms (Alcala et al., 2019; Fernandez-Cuesta et al., 2014; Laddha et al., 2019) were obtained from <https://nextjournal.com/rarecancersgenomics/a-molecular-map-of-lung-neuroendocrine-neoplasms/>, and the histopathologic classification in the attributes metadata file was used to identify the 111 typical and atypical lung carcinoids whose expression profiles of peptidergic genes and PNEC markers were analyzed here. Log-transformed gene read counts were normalized by quantile normalization (Dillies et al., 2013), and the obtained values were represented as heatmaps (Fig. S13).

Rarefaction analysis to estimate saturation of PNEC neuropeptide diversity

We modeled each PNEC peptidergic expression profile as an incidence sampling of all total possible peptidergic genes expressed by the PNEC population, and estimated the peptidergic diversity using rarefaction and extrapolation analysis, a technique used in ecology to assess species richness (Chao et al., 2014). In this analogy, each 'species' is a peptidergic gene, and a given PNEC that may express any number of distinct peptidergic genes is analogous to a sampling of the total assemblage of species (peptidergic genes). Using the iNEXT package to estimate species richness (Hsieh et al., 2016), we constructed an integrated curve to smoothly link rarefaction (interpolation) and prediction (extrapolation), and the associated 95% confidence intervals, by bootstrapping (N=200). Only the incidence (presence or absence), and not the abundance, of each peptidergic gene RNA was used to estimate the underlying neuropeptide accumulation curve (Chao et al., 2014; Garcia-Ortega and Martinez, 2015).

REFERENCES

Adriaensen, D., Scheuermann, D.W., Gajda, M., Brouns, I., and Timmermans, J.P. (2009). Functional implications of extensive new data on the innervation of pulmonary neuroepithelial bodies. Italian Journal of Anatomy and Embryology *106*, 395–403.

Alcala, N., Leblay, N., Gabriel, A.A.G., Mangiante, L., Hervas, D., Giffon, T., Sertier, A.S., Ferrari, A., Derks, J., Ghantous, A., et al. (2019). Integrative and comparative genomic analyses identify clinically relevant pulmonary carcinoid groups and unveil the supra-carcinoids. Nature Communications *10*, 3407.

Amara, S.G., David, D.N., Rosenfeld, M.G., Roos, B.A., and Evans, R.M. (1980). Characterization of rat calcitonin mRNA. Proceedings of the National Academy of Sciences *77*, 4444–4448.

Amara, S.G., Jonas, V., Rosenfeld, M.G., Ong, E.S., and Evans, R.M. (1982). Alternative RNA processing in calcitonin gene expression generates mRNAs encoding different polypeptide products. Nature *298*, 240–244.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Arioglu, E., Doppman, J., Gomes, M., Kleiner, D., Mauro, D., Barlow, C., and Papanicolaou, D.A. (1998). Cushing's syndrome caused by corticotropin secretion by pulmonary tumorlets. New England Journal of Medicine *339*, 883–886.

Athanassiadi, K., Exarchos, D., Tsagarakis, S., and Bellenis, I. (2004). Acromegaly caused by ectopic growth hormone–releasing hormone secretion by a carcinoid bronchial tumor: a rare entity. The Journal of Thoracic and Cardiovascular Surgery *128*, 631–632.

Bai, L., Mesgarzadeh, S., Ramesh, K.S., Huey, E.L., Liu, Y., Gray, L.A., Aitken, T.J., Chen, Y., Beutler, L.R., Ahn, J.S., et al. (2019). Genetic identification of vagal sensory neurons that control feeding. Cell *179*, 1129–1143.

Bang, H., Kim, Y., and Kim, D. (2000). TREK-2, a new member of the mechanosensitive tandem-pore K +channel family. Journal of Biological Chemistry *275*, 17412–17419.

Barrios, J., Patel, K.R., Aven, L., Achey, R., Minns, M.S., Lee, Y., Trinkaus-Randall, V.E., and Ai, X. (2017). Early life allergen-induced mucus overproduction requires augmented neural stimulation of pulmonary neuroendocrine cell secretion. The FASEB Journal *31*, 4117–4128.

Barrios, J., Kho, A.T., Aven, L., Mitchel, J.A., Park, J.-A., Randell, S.H., Miller, L.A., Tantisira, K.G., and Ai, X. (2019). Pulmonary neuroendocrine cells secrete γ-aminobutyric acid to induce goblet cell hyperplasia in primate models. American Journal of Respiratory Cell and Molecular Biology *60*, 687–694.

Bellono, N.W., Bayrer, J.R., Leitch, D.B., Castro, J., Zhang, C., O'Donnell, T.A., Brierley, S.M., Ingraham, H.A., and Julius, D. (2017). Enterochromaffin cells are gut chemosensors that couple to sensory neural pathways. Cell *170*, 185–198.

Beumer, J., Puschhof, J., Bauzá-Martinez, J., Martínez-Silgado, A., Elmentaite, R., James, K.R., Ross, A., Hendriks, D., Artegiani, B., Busslinger, G.A., et al. (2020). High-resolution mRNA and secretome atlas of human enteroendocrine cells. Cell *181*, 1291–1306.

Blackshaw, S., and Snyder, S.H. (1999). Encephalopsin: a novel mammalian extraretinal opsin discretely localized in the brain. Journal of Neuroscience *19*, 3681–3690.

Boers, J.E., den Brok, J.L., Koudstaal, J., Arends, J.W., and Thunnissen, F.B. (1996). Number and proliferation of neuroendocrine cells in normal human airway epithelium. American Journal of Respiratory and Critical Care Medicine *154*, 758–763.

Buck, L., and Axel, R. (1991). A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. Cell *65*, 175–187.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nature Biotechnology *36*, 411–420.

Buttigieg, J., Pan, J., Yeger, H., and Cutz, E. (2012). NOX2 (gp91 phox) is a predominant O2 sensor in a human airway chemoreceptor cell line: biochemical, molecular, and electrophysiological evidence. American Journal of Physiology. Lung Cellular and Molecular Physiology *303*, L598–L607.

Caterina, M.J., Rosen, T.A., Tominaga, M., Brake, A.J., and Julius, D. (1999). A capsaicin-receptor homologue with a high threshold for noxious heat. Nature *398*, 436–441.

Chandrashekar, J., Hoon, M.A., Ryba, N.J.P., and Zuker, C.S. (2006). The receptors and cells for mammalian taste. Nature *444*, 288–294.

Chang, A.J., Ortega, F.E., Riegler, J., Madison, D.V., and Krasnow, M.A. (2015). Oxygen regulation of breathing through an olfactory receptor activated by lactate. Nature *527*, 240–244.

Chao, A., Gotelli, N.J., Hsieh, T.C., Sander, E.L., Ma, K.H., Colwell, R.K., and Ellison, A.M. (2014). Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. Ecological Monographs *84*, 45–67.

Clapham, D.E. (2003). TRP channels as cellular sensors. Nature *426*, 517–524.

Cutz, E., Perrin, D.G., Pan, J., Haas, E.A., and Krous, H.F. (2007). Pulmonary neuroendocrine cells and neuroepithelial bodies in sudden infant death syndrome: potential markers of airway chemoreceptor dysfunction. Pediatric and Developmental Pathology *10*, 106–116.

Cutz, E., Pan, J., Yeger, H., Domnik, N.J., and Fisher, J.T. (2013). Recent advances and controversies on the role of pulmonary neuroepithelial bodies as airway sensors. Seminars in Cell and Developmental Biology *24*, 40–50.

Darmanis, S., Sloan, S.A., Zhang, Y., Enge, M., Caneda, C., Shuer, L.M., Hayden Gephart, M.G., Barres, B.A., and Quake, S.R. (2015). A survey of human brain transcriptome diversity at the single cell level. Proceedings of the National Academy of Sciences *112*, 7285–7290.

Darmanis, S., Sloan, S.A., Croote, D., Mignardi, M., Chernikova, S., Samghababi, P., Zhang, Y., Neff, N., Kowarsky, M., Caneda, C., et al. (2017). Single-cell RNA-seq analysis of infiltrating neoplastic cells at the migrating front of human glioblastoma. Cell Reports *21*, 1399–1410.

Davila, D.G., Dunn, W.F., Tazelaar, H.D., and Pairolero, P.C. (1993). Bronchial carcinoid tumors. Mayo Clinic Proceedings *68*, 795–803.

Dillies, M.A., Rau, A., Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., Keime, C., Marot, G., Castel, D., Estelle, J., et al. (2013). A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. Briefings in Bioinformatics *14*, 671–683.

Dulac, C., and Axel, R. (1995). A novel family of genes encoding putative pheromone receptors in mammals. Cell *83*, 195–206.

Duprat, F., Lesage, F., Fink, M., Reyes, R., Heurteaux, C., and Lazdunski, M. (1997). TASK, a human background K+ channel to sense external pH variations near physiological pH. The EMBO Journal *16*, 5464–5471.

Ebina, M., Hoyt, R.F., McNelly, N.A., Sorokin, S.P., and Linnoila, R.I. (1997). Effects of hydrogen and bicarbonate ions on endocrine cells in fetal rat lung organ cultures. American Journal of Physiology *272*, L178-86.

Fernandez-Cuesta, L., Peifer, M., Lu, X., Sun, R., Ozretić, L., Seidel, D., Zander, T., Leenders, F., George, J., Müller, C., et al. (2014). Frequent mutations in chromatin-remodelling genes in pulmonary carcinoids. Nat Commun *5*, 3518.

Fu, X.W., and Cutz, E. (2002). Hypoxia-induced secretion of serotonin from intact pulmonary neuroepithelial bodies in neonatal rabbit. The Journal of Physiology 503–510.

Fu, X.W., Wang, D., Nurse, C.A., Dinauer, M.C., and Cutz, E. (2000). NADPH oxidase is an O2 sensor in airway chemoreceptors: evidence from K+ current modulation in wild-type and oxidase-deficient mice. Proceedings of the National Academy of Sciences *97*, 4374–4379.

Garcia-Ortega, L.F., and Martinez, O. (2015). How many genes are expressed in a transcriptome? Estimation and results for RNA-Seq. PLOS One *10*, e0130262.

Gillan, J.E., and Cutz, E. (1993). Abnormal pulmonary bombesin immunoreactive cells in Wilson-Mikity syndrome (pulmonary dysmaturity) and bronchopulmonary dysplasia. Pediatric Pathology *13*, 165–180.

Gillich, A., Zhang, F., Farmer, C.G., Travaglini, K.J., Tan, S.Y., Gu, M., Zhou, B., Feinstein, J.A., Krasnow, M.A., and Metzger, R.J. (2020). Capillary cell-type specialization in the alveolus. Nature *586*, 785–789.

Gu, X., Karp, P.H., Brody, S.L., Pierce, R.A., Welsh, M.J., Holtzman, M.J., and Ben-Shahar, Y. (2014). Chemosensory functions for pulmonary neuroendocrine cells. American Journal of Respiratory Cell and Molecular Biology *50*, 637–646.

Haltaufderhyde, K., Ozdeslik, R.N., Wicks, N.L., Najera, J.A., and Oancea, E. (2014). Opsin expression in human epidermal skin. Photochemistry and Photobiology *91*, 117–123.

Han, X., Wang, R., Zhou, Y., Fei, L., Sun, H., Lai, S., Saadatpour, A., Zhou, Z., Chen, H., Ye, F., et al. (2018). Mapping the mouse cell atlas by microwell-seq. Cell *172*, 1091–1107.

Harno, E., Gali Ramamoorthy, T., Coll, A.P., and White, A. (2018). POMC: The physiological power of hormone processing. Physiological Reviews *98*, 2381–2430.

Hsieh, T.C., Ma, K.H., and Chao, A. (2016). iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). Methods in Ecology and Evolution *7*, 1451–1456.

Ji, L., Zhu, H., Chen, H., Fan, W., Chen, J., Chen, J., Zhu, G., and Wang, J. (2015). Modulation of CaV1.2 calcium channel by neuropeptide W regulates vascular myogenic tone via G protein-coupled receptor 7. Journal of Hypertension *33*, 2431–2442.

Johnson, M.D., Gray, M.E., and Stahlman, M.T. (1988). Calcitonin gene-related peptide in human fetal lung and in neonatal lung disease. Journal of Histochemistry & Cytochemistry *36*, 199–204.

Kaelberer, M.M., Buchanan, K.L., Klein, M.E., Barth, B.B., Montoya, M.M., Shen, X., and Bohórquez, D.V. (2018). A gut-brain neural circuit for nutrient sensory transduction. Science *361*, eaat5236.

Katz, Y., Wang, E.T., Airoldi, E.M., and Burge, C.B. (2010). Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nature Methods *7*, 1009–1015.

Kim, E.J., Ables, J.L., Dickel, L.K., Eisch, A.J., and Johnson, J.E. (2011a). Ascl1 (Mash1) defines cells with long-term neurogenic potential in subgranular and subventricular zones in adult mouse brain. PLOS One *6*, e18472.

Kim, Y., Bark, S., Hook, V., and Bandeira, N. (2011b). NeuroPedia: neuropeptide database and spectral library. Bioinformatics *27*, 2772–2773.

de Kloet, A.D., Wang, L., Pitra, S., Hiller, H., Smith, J.A., Tan, Y., Nguyen, D., Cahill, K.M., Sumners, C., Stern, J.E., et al. (2017). A unique “angiotensin-sensitive” neuronal population coordinates neuroendocrine, cardiovascular, and behavioral responses to stress. The Journal of Neuroscience *37*, 3478–3490.

Kumar, P., and Prabhakar, N.R. (2012). Peripheral chemoreceptors: function and plasticity of the carotid body. Comprehensive Physiology *2*, 141–219.

Kuo, C.S., and Krasnow, M.A. (2015). Formation of a neurosensory organ by epithelial cell slithering. Cell *163*, 394–405.

Laddha, S.V., da Silva, E.M., Robzyk, K., Untch, B.R., Ke, H., Rekhtman, N., Poirier, J.T., Travis, W.D., Tang, L.H., and Chan, C.S. (2019). Integrative genomic characterization identifies molecular subtypes of lung carcinoids. Cancer Research *79*, 4339–4347.

Lauweryns, J.M., Cokelaere, M., and Theunynck, P. (1972). Neuro-epithelial bodies in the respiratory mucosa of various mammals. A light optical, histochemical and ultrastructural investigation. Zeitschrift Fur Zellforschung Und Mikroskopische Anatomie *135*, 569–592.

Lauweryns, J.M., Cokelaere, J., and Theunynck, P. (1973). Serotonin producing neuroepithelial bodies in rabbit respiratory mucosa. Science *180*, 410–413.

Lauweryns, J.M., De Bock, V., Verhofstad, A.A., and Steinbusch, H.W. (1977). Intrapulmonary neuroepithelial bodies in newborn rabbits. Cell and Tissue Research *182*, 215–223.

Lauweryns, J.M., Cokelaere, M., Lerut, T., and Theunynck, P. (1978). Cross-circulation studies on the influence of hypoxia and hypoxaemia on neuro-epithelial bodies in young rabbits. Cell and Tissue Research *193*, 373–386.

Lauweryns, J.M., van Ranst, L., Lloyd, R.V., and O'Connor, D.T. (1987). Chromogranin in bronchopulmonary neuroendocrine cells. Immunocytochemical detection in human, monkey, and pig respiratory mucosa. Journal of Histochemistry & Cytochemistry *35*, 113–118.

Leib, D.E., Zimmerman, C.A., Poormoghaddam, A., Huey, E.L., Ahn, J.S., Lin, Y.-C., Tan, C.L., Chen, Y., and Knight, Z.A. (2017). The forebrain thirst circuit drives drinking through negative reinforcement. Neuron *96*, 1272–1281.

Lembrechts, R., Brouns, I., Schnorbusch, K., Pintelon, I., Timmermans, J.-P., and Adriaensen, D. (2012). Neuroepithelial bodies as mechanotransducers in the intrapulmonary airway epithelium: involvement of TRPC5. American Journal of Respiratory Cell and Molecular Biology *47*, 315–323.

Lembrechts, R., Brouns, I., Schnorbusch, K., Pintelon, I., Kemp, P.J., Timmermans, J.P., Riccardi, D., and Adriaensen, D. (2013). Functional expression of the multimodal extracellular calcium-sensing receptor in pulmonary neuroendocrine cells. Journal of Cell Science *126*, 4490–4501.

Limper, A.H., Carpenter, Paul C., Scheithauer, Bernd, and Staats, Bruce A. (1992). The Cushing syndrome induced by bronchial carcinoid tumors. Annals of Internal Medicine *117*, 209–214.

Lin, W., Burks, C.A., Hansen, D.R., Kinnamon, S.C., and Gilbertson, T.A. (2004). Taste receptor cells express pH-sensitive leak K +channels. Journal of Neurophysiology *92*, 2909–2919.

Liu, Y., Diaz de Arce, A., and Krasnow, M.A. (2021). Molecular, anatomical, and functional organization of lung interoceptors. bioRxiv 2021.11.10.468116.

Livermore, S., Zhou, Y., Pan, J., Yeger, H., Nurse, C.A., and Cutz, E. (2015). Pulmonary neuroepithelial bodies are polymodal airway sensors: evidence for CO2/H+ sensing. American Journal of Physiology-Lung Cellular and Molecular Physiology *308*, L807-15.

Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., et al. (2009). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nature Neuroscience 1–10.

McKemy, D.D., Neuhausser, W.M., and Julius, D. (2002). Identification of a cold receptor reveals a general role for TRP channels in thermosensation. Nature *416*, 52–58.

Montoro, D.T., Haber, A.L., Biton, M., Vinarsky, V., Lin, B., Birket, S.E., Yuan, F., Chen, S., Leung, H.M., Villoria, J., et al. (2018). A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. Nature *560*, 319–324.

Mou, H., Yang, Y., Riehs, M.A., Barrios, J., Shivaraju, M., Haber, A.L., Montoro, D.T., Gilmore, K., Haas, E.A., Paunovic, B., et al. (2021). Airway basal stem cells generate distinct subpopulations of PNECs. Cell Reports *35*, 109011.

Mulligan, E., Lahiri, S., and Storey, B.T. (1981). Carotid body O2 chemoreception and mitochondrial oxidative phosphorylation. Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology *51*, 438–446.

Nonomura, K., Woo, S.-H., Chang, R.B., Gillich, A., Qiu, Z., Francisco, A.G., Ranade, S.S., Liberles, S.D., and Patapoutian, A. (2017). Piezo2 senses airway stretch and mediates lung inflation-induced apnoea. Nature *541*, 176–181.

Ouadah, Y., Rojas, E.R., Riordan, D.P., Capostagno, S., Kuo, C.S., and Krasnow, M.A. (2019). Rare pulmonary neuroendocrine cells are stem cells regulated by Rb, p53, and Notch. Cell *179*, 403–416.

Park, K.S., Liang, M.-C., Raiser, D.M., Zamponi, R., Roach, R.R., Curtis, S.J., Walton, Z., Schaffer, B.E., Roake, C.M., Zmoos, A.-F., et al. (2011). Characterization of the cell of origin for small cell lung cancer. Cell Cycle *10*, 2806–2815.

Peier, A.M., Moqrich, A., Hergarden, A.C., Reeve, A.J., Andersson, D.A., Story, G.M., Earley, T.J., Dragoni, I., McIntyre, P., Bevan, S., et al. (2002). A TRP channel that senses cold stimuli and menthol. Cell *108*, 705–715.

Pernow, B., and Waldenstrom, Jan (1957). Determination of 5-hydroxytryptamine, 5-hydroxyindole acetic acid and histamine in thirty-three cases of carcinoid tumor (argentaffinoma). American Journal of Medicine 16–25.

Picelli, S., Faridani, O.R., Bjorklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. Nature Protocols *9*, 171–181.

Plasschaert, L.W., Žilionis, R., Choo-Wing, R., Savova, V., Knehr, J., Roma, G., Klein, A.M., and Jaffe, A.B. (2018). A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. Nature *560*, 377–381.

Quintana, A., Zanella, S., Koch, H., Kruse, S.E., Lee, D., Ramirez, J.M., and Palmiter, R.D. (2012). Fatal breathing dysfunction in a mouse model of Leigh syndrome. Journal of Clinical Investigation *122*, 2359–2368.

Ranade, S.S., Syeda, R., and Patapoutian, A. (2015). Mechanically activated ion channels. Neuron *87*, 1162–1179.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nature Biotechnology *29*, 24–26.

Rudin, C.M., Poirier, J.T., Byers, L.A., Dive, C., Dowlati, A., George, J., Heymach, J.V., Johnson, J.E., Lehman, J.M., MacPherson, D., et al. (2019). Molecular subtypes of small cell lung cancer: a synthesis of human and mouse model data. Nature Reviews Cancer *19*, 289–297.

Scheuermann, D.W. (1997). Comparative histology of pulmonary neuroendocrine cell system in mammalian lungs. Microscopy Research and Technique *37*, 31–42.

Scheuermann, D.W., Timmermans, J.P., Adriaensen, D., and De Groodt-Lasseel, M.H. (1987). Immunoreactivity for calcitonin gene-related peptide in neuroepithelial bodies of the newborn cat. Cell and Tissue Research *249*, 337–340.

Schnorbusch, K., Lembrechts, R., Pintelon, I., Timmermans, J.-P., Brouns, I., and Adriaensen, D. (2013). GABAergic signaling in the pulmonary neuroepithelial body microenvironment: functional imaging in GAD67-GFP mice. Histochemistry and Cell Biology *140*, 549–566.

Secher, A., Kelstrup, C.D., Conde-Frieboes, K.W., Pyke, C., Raun, K., Wulff, B.S., and Olsen, J.V. (2016). Analytic framework for peptidomics applied to large-scale neuropeptide identification. Nature Communications *7*, 11436–10.

Shalet, S.M., Beardwell, C.G., MacFarlane, I.A., Ellison, M.L., Norman, C.M., Rees, L.H., and Hughes, M. (1979). Acromegaly due to production of a growth hormone releasing factor by a bronchial carcinoid tumor. Clinical Endocrinology *10*, 61–67.

Shivaraju, M., Chitta, U.K., Grange, R.M.H., Jain, I.H., Capen, D., Liao, L., Xu, J., Ichinose, F., Zapol, W.M., Mootha, V.K., et al. (2021). Airway stem cells sense hypoxia and differentiate into protective solitary neuroendocrine cells. Science *371*, 52–57.

Song, H., Yao, E., Lin, C., Gacayan, R., Chen, M.-H., and Chuang, P.-T. (2012). Functional characterization of pulmonary neuroendocrine cells in lung development, injury, and tumorigenesis. Proceedings of the National Academy of Sciences *109*, 17531–17536.

Stettner, G.M., Viscomi, C., Zeviani, M., Wilichowski, E., and Dutschmann, M. (2011). Hypoxic and hypercapnic challenges unveil respiratory vulnerability of Surf1 knockout mice, an animal model of Leigh syndrome. Mitochondrion *11*, 413–420.

Stevens, T.P., McBride, J.T., Peake, J.L., Pinkerton, K.E., and Stripp, B.R. (1997). Cell proliferation contributes to PNEC hyperplasia after acute airway injury. American Journal of Physiology *272*, L486-93.

Sui, P., Wiesner, D.L., Xu, J., Zhang, Y., Lee, J., Van Dyken, S., Lashua, A., Yu, C., Klein, B.S., Locksley, R.M., et al. (2018). Pulmonary neuroendocrine cells amplify allergic asthma responses. Science *360*, 1086.

Sutherland, K.D., Proost, N., Brouns, I., Adriaensen, D., Song, J.-Y., and Berns, A. (2011). Cell of origin of small cell lung cancer: Inactivation of Trp53 and Rb1 in distinct cell types of adult mouse lung. Cancer Cell *19*, 754–764.

Tabula Muris Consortium (2018). Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. Nature *562*, 367–372.

Terakita, A. (2005). The opsins. Genome Biology *6*, 213.

Tominaga, M., Caterina, M.J., Malmberg, A.B., Rosen, T.A., Gilbert, H., Skinner, K., Raumann, B.E., Basbaum, A.I., and Julius, D. (1998). The cloned capsaicin receptor integrates multiple pain-producing stimuli. Neuron *21*, 531–543.

Travaglini, K.J., Nabhan, A.N., Penland, L., Sinha, R., Gillich, A., Sit, R.V., Chang, S., Conley, S.D., Mori, Y., Seita, J., et al. (2020). A molecular cell atlas of the human lung from single-cell RNA sequencing. Nature *587*, 619–625.

Travis, W.D., Brambilla, E., Nicholson, A.G., Yatabe, Y., Austin, J.H.M., Beasley, M.B., Chirieac, L.R., Dacic, S., Duhig, E., Flieder, D.B., et al. (2015). The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. Journal of Thoracic Oncology *10*, 1243–1260.

Vandewauw, I., De Clercq, K., Mulier, M., Held, K., Pinto, S., Van Ranst, N., Segal, A., Voet, T., Vennekens, R., Zimmermann, K., et al. (2018). A TRP channel trio mediates acute noxious heat sensing. Nature 662–666.

Waldmann, R., Champigny, G., Bassilana, F., Heurteaux, C., and Lazdunski, M. (1997). A proton-gated cation channel involved in acid-sensing. Nature *386*, 173–177.

Weichselbaum, M., Sparrow, M.P., Hamilton, E.J., Thompson, P.J., and Knight, D.A. (2005). A confocal microscopic study of solitary pulmonary neuroendocrine cells in human airway epithelium. Respiratory Research *6*, 311.

Wolock, S.L., Lopez, R., and Klein, A.M. (2019). Scrublet: computational identification of cell doublets in single-cell transcriptomic data. Cell Systems *8*, 281–291.

Worthington, J.J., Reimann, F., and Gribble, F.M. (2018). Enteroendocrine cells-sensory sentinels of the intestinal environment and orchestrators of mucosal immunity. Mucosal Immunology *11*, 3–20.

Xu, J., Yu, H., and Sun, X. (2020). Less is more: rare pulmonary neuroendocrine cells function as critical sensors in lung. Developmental Cell *55*, 123–132.

Young, L.R., Brody, A.S., Inge, T.H., Acton, J.D., Bokulic, R.E., Langston, C., and Deutsch, G.H. (2011). Neuroendocrine cell distribution and frequency distinguish neuroendocrine cell hyperplasia of infancy from other pulmonary disorders. Chest *139*, 1060–1071.

Youngson, C., Nurse, C., Yeger, H., and Cutz, E. (1993). Oxygen sensing in airway chemoreceptors. Nature *365*, 153–155.

Yu, F.H., and Catterall, W.A. (2004). The VGL-chanome: a protein superfamily specialized for electrical signaling and ionic homeostasis. Science's STKE : Signal Transduction Knowledge Environment *2004*, 1–16.

Yu, N., Chu, C., Kunitake, T., Kato, K., Nakazato, M., and Kannan, H. (2007). Cardiovascular actions of central neuropeptide W in conscious rats. Regulatory Peptides *138*, 82–86.

Zucchi, R., Chiellini, G., Scanlan, T.S., and Grandy, D.K. (2009). Trace amine-associated receptors and their ligands. British Journal of Pharmacology *149*, 967–978.

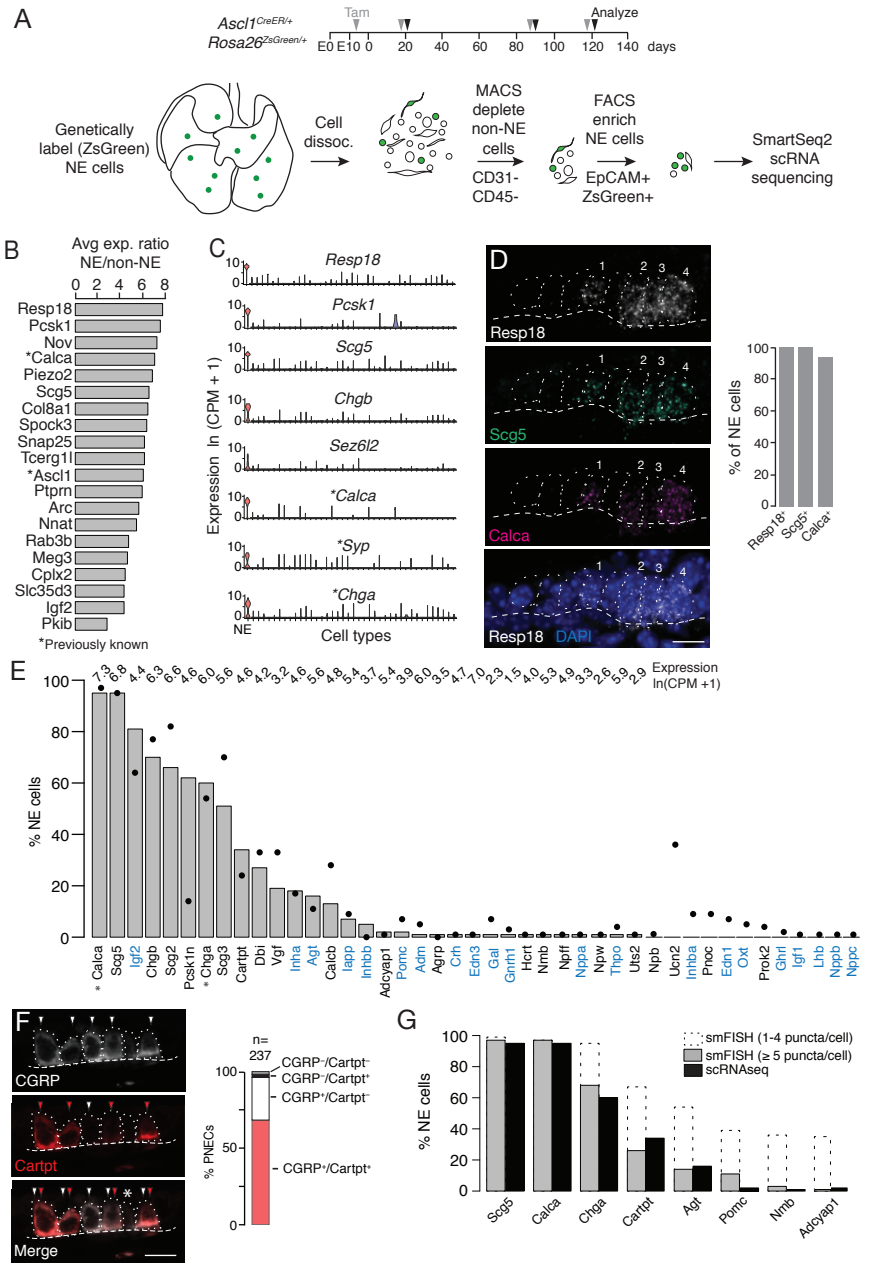


Fig. 2, Kuo et al

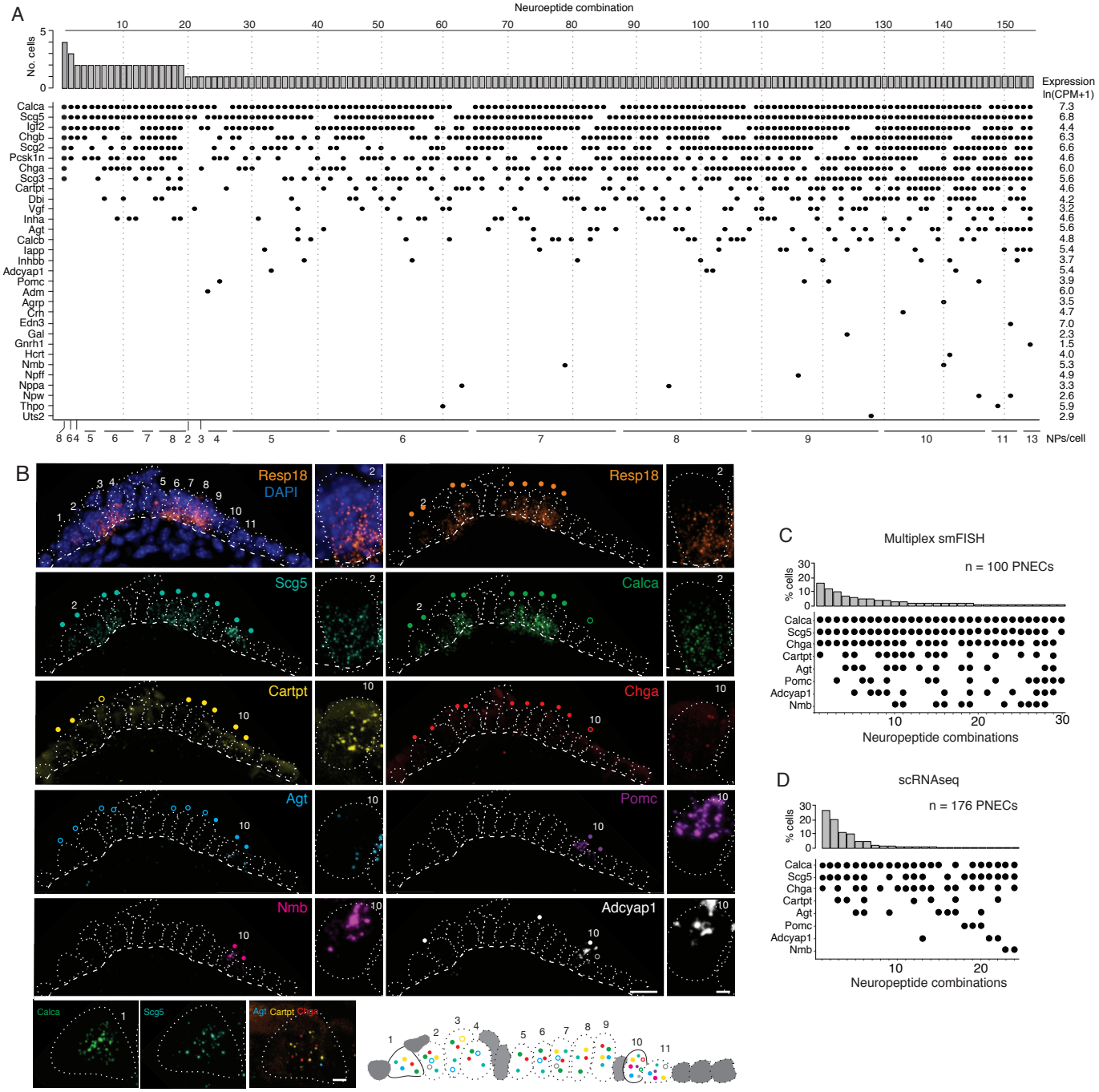
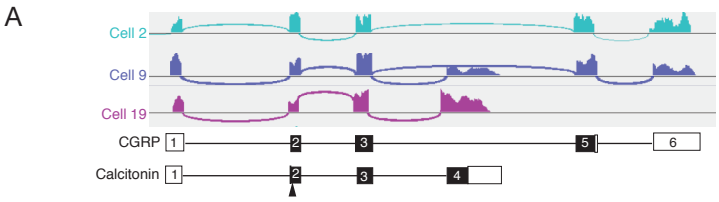


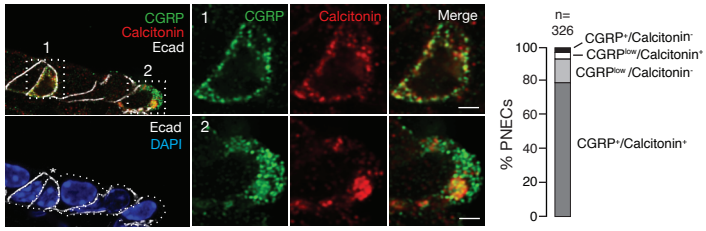
Fig. 3, Kuo et al

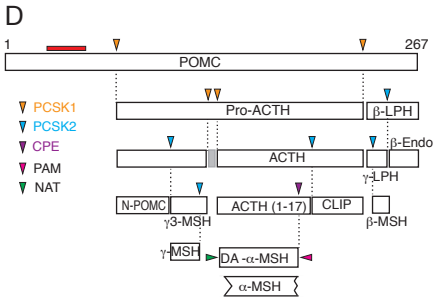


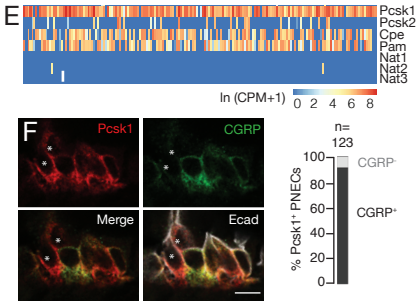
B

Cell ID	CGRP* (%)	Calcit.* (%)	Alt. splice site reads	Total reads	% Alt. splice site reads
1	100	0	66	735865	0.01
2	100	0	10	893719	0.00
3	100	0	242	832403	0.03
4	100	0	16	468633	0.00
5	94	6	77	805934	0.01
6	89	11	1428	678393	0.21
7	85	15	307	716881	0.04
8	68	32	177	629688	0.03
9	66	34	101	681677	0.01
10	65	35	920	710376	0.13
11	64	36	185	582010	0.03
12	58	43	440	655541	0.07
13	54	46	551	523982	0.11
14	45	55	225	695595	0.03
15	45	55	673	524167	0.13
16	40	60	306	440935	0.07
17	37	63	750	716998	0.10
18	18	82	278	655760	0.04
19	0	100	129	892843	0.01
20	0	100	95	751588	0.01

*Percent CGRP or calcitonin (Calcit.) mRNA determined by alternative (Alt.) splice site reads







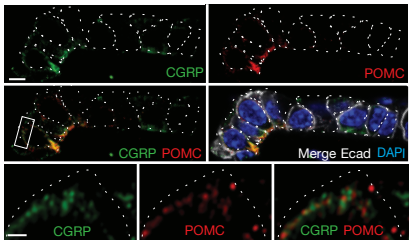


Fig. 4, Kuo et al.

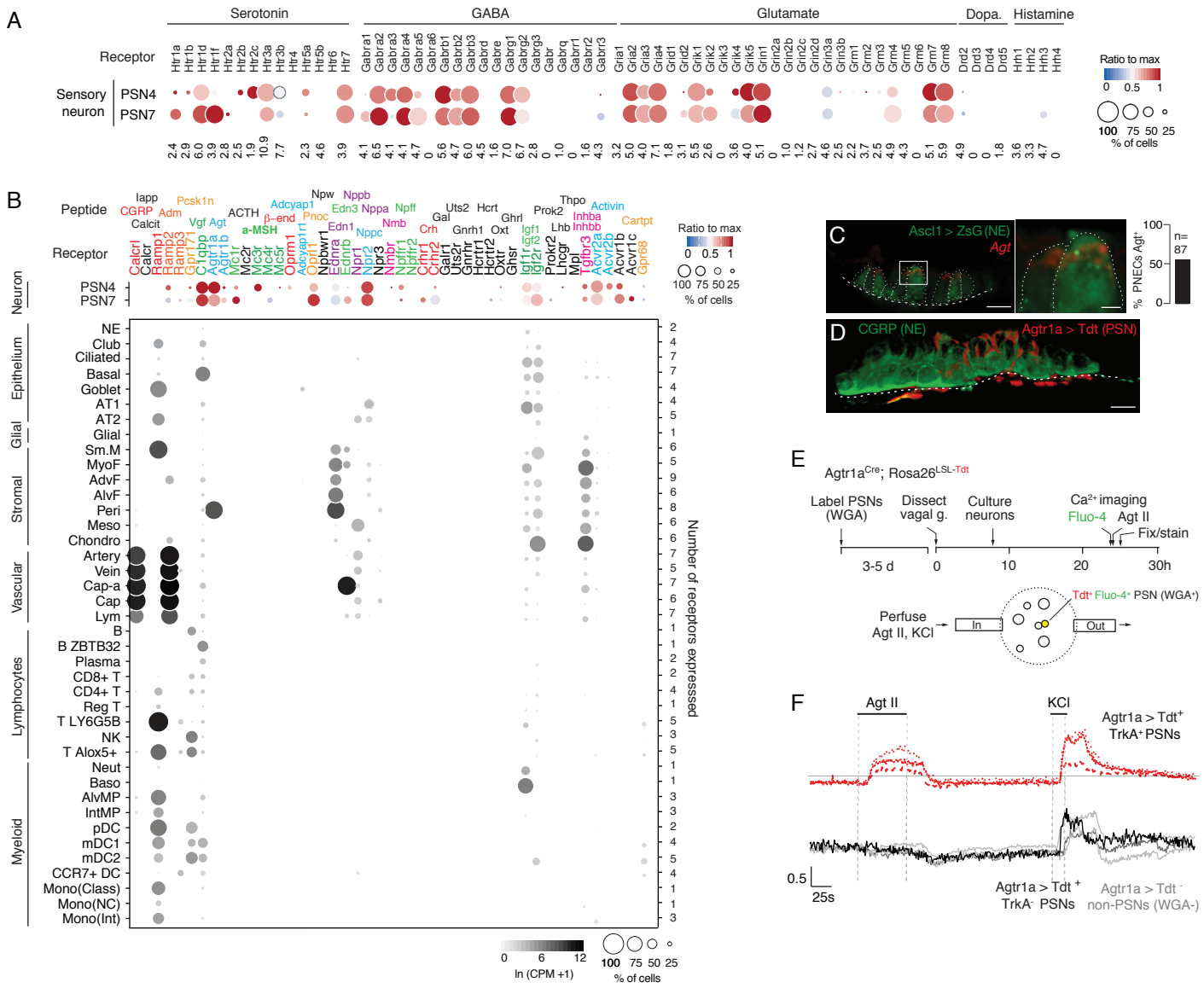


Fig. 5, Kuo et al

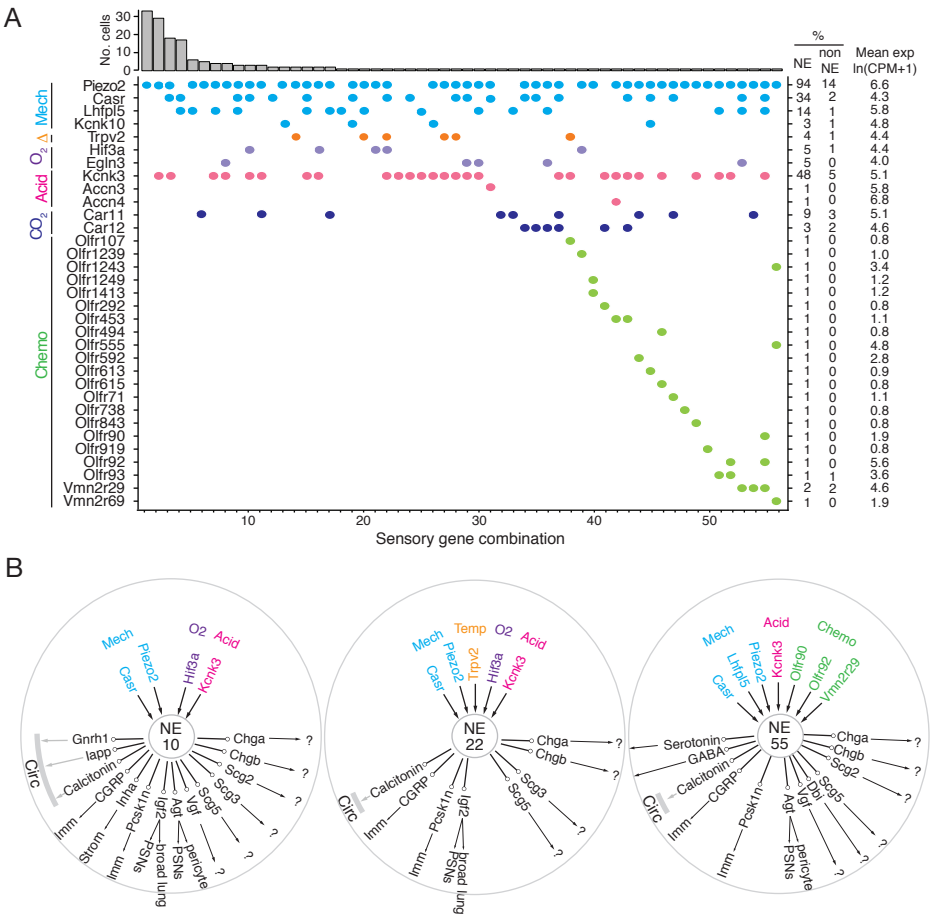
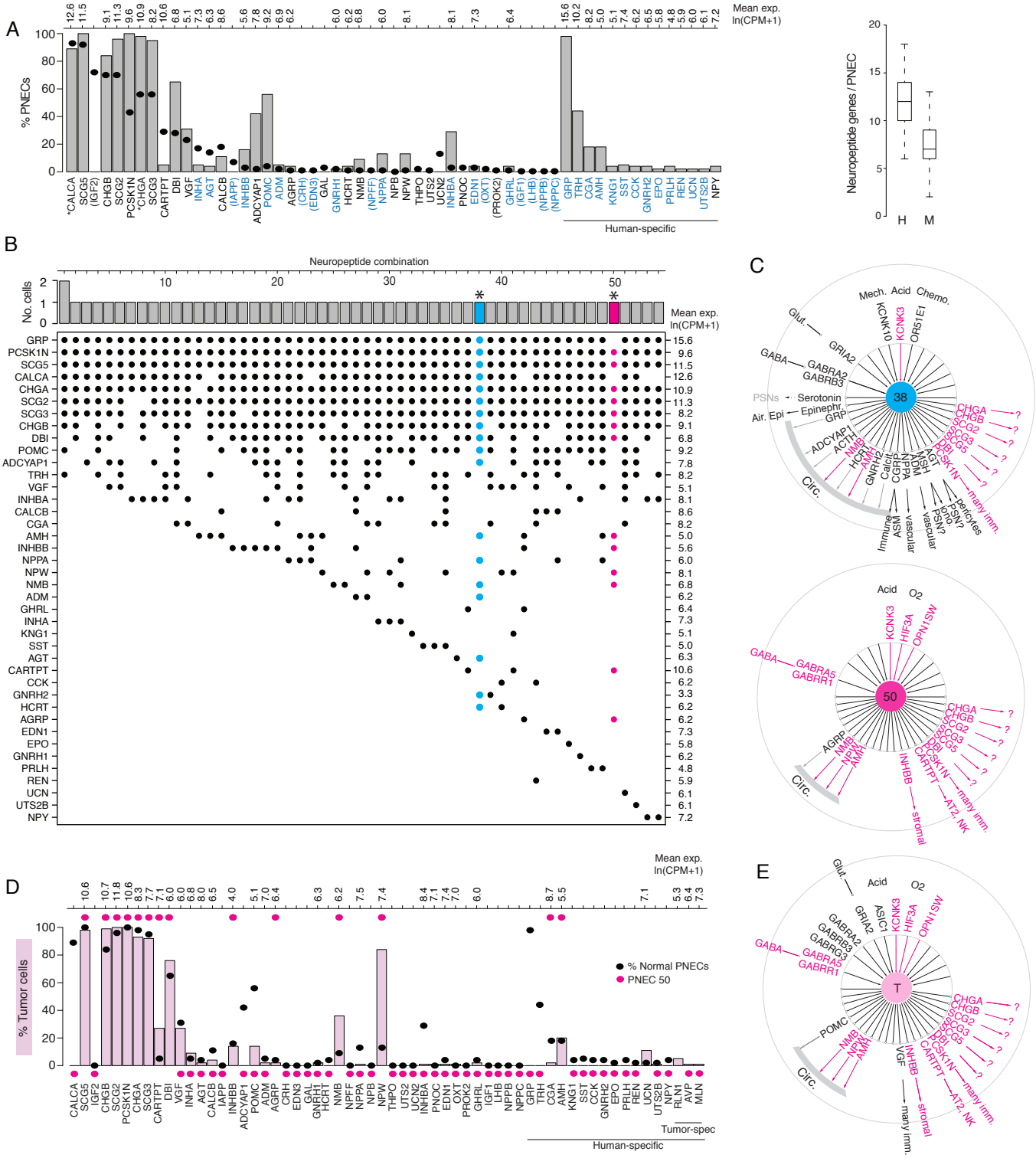
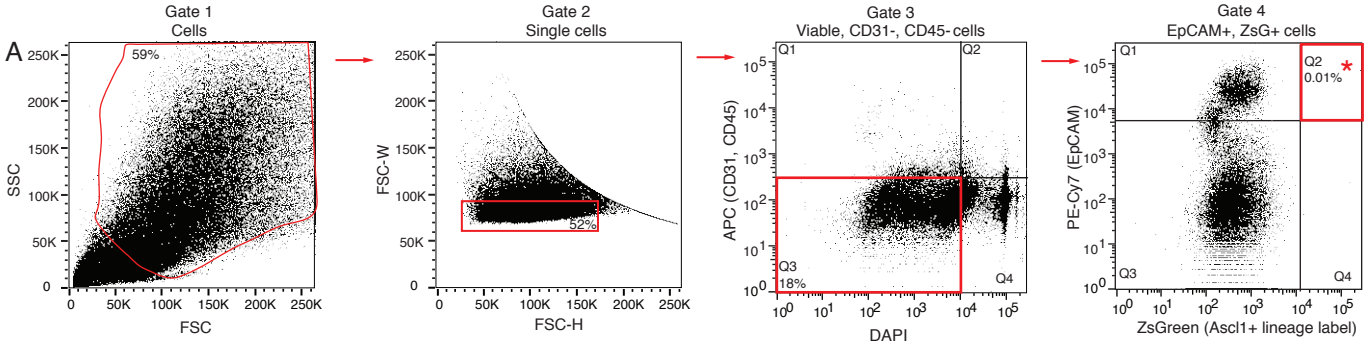


Fig. 6, Kuo et al



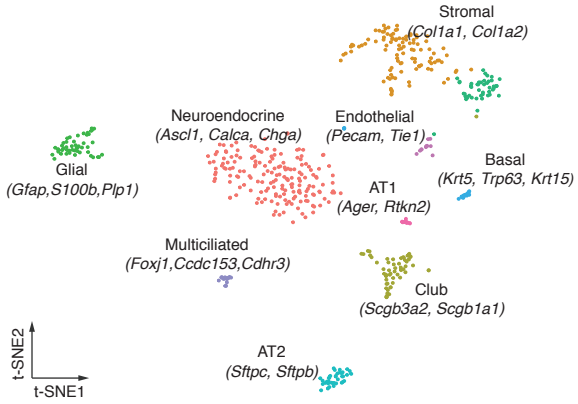
A



B

Full gating path	Count	% of parent	% of total
Ungated	100000	100	100
Gate 1: Cells	59258	59.3	59.3
Gate 2: Single cells	31139	52.5	31.1
Gate 3: Viable (DAPI-), non-endothelial (CD31-), non-immune (CD45-)	18705	60.1	18.7
*Gate 4: Q2. NE (ZsGreen+ , EpCAM+)	2	0.01	0.002
Q1: Epithelial (ZsGreen- , EpCAM+)	3509	18.8	3.5
Q4. Neural (ZsGreen+ , EpCAM-)	2	0.01	0.002
Q3. Non-epithelial, non-neural (ZsGreen- , EpCAM-)	15192	81.2	15.2

C



D

Cell type	EpCAM+ ZsG+ No. of cells (%)	EpCAM- ZsG+ No. of cells (%)	EpCAM+ ZsG- No. of cells (%)	Total by cell type
NE	112 (63)	64 (23)	0 (0)	176
Glial	5 (3)	43 (15)	1 (2)	49
Club	39 (22)	14 (5)	6 (10)	59
MC	6 (3)	3 (1)	8 (47)	17
Basal	0 (0)	5 (2)	13 (72)	18
AT2	3 (2)	3 (1)	31 (84)	37
AT1	0 (0)	1 (1)	12 (92)	13
Stromal	9 (5)	139 (50)	4 (3)	152
Endo.	3 (2)	7 (3)	3 (24)	13
Total	177 (33)	279 (52)	78 (15)	534

E

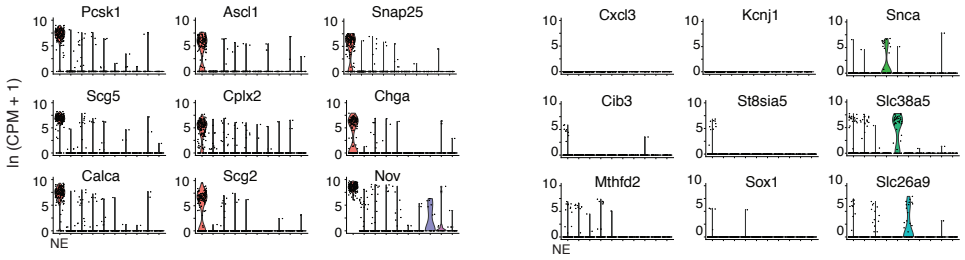


Fig. S2, Kuo et al

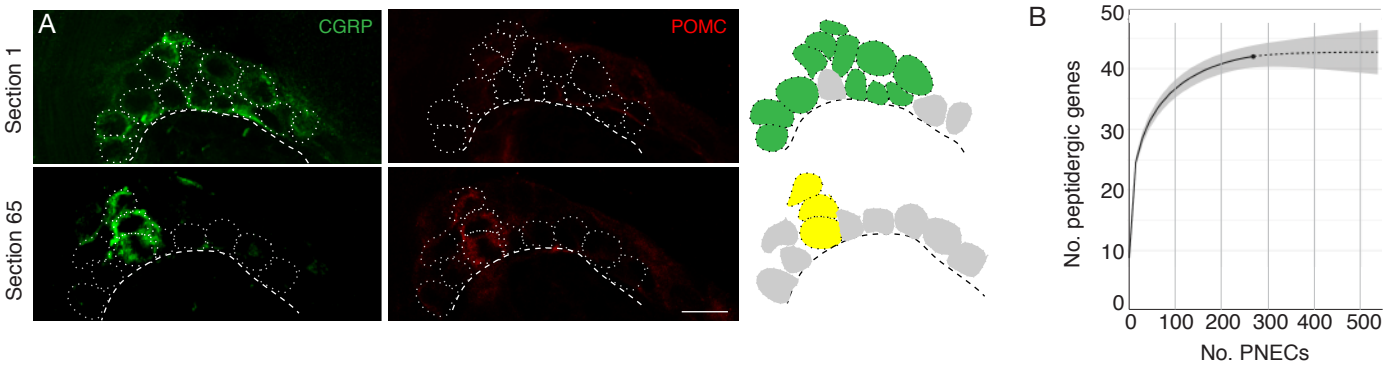


Fig. S3, Kuo et al

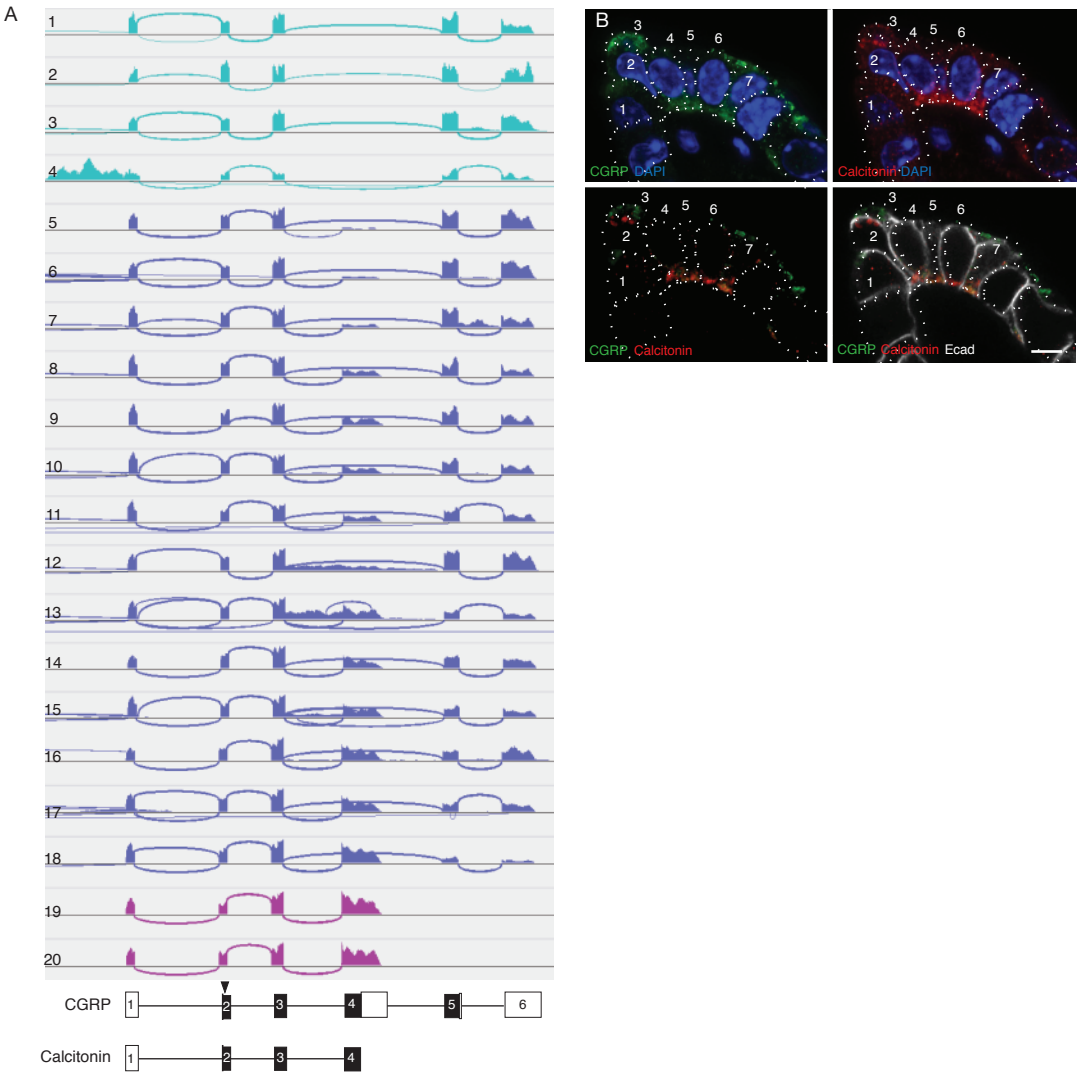


Fig. S4, Kuo et al

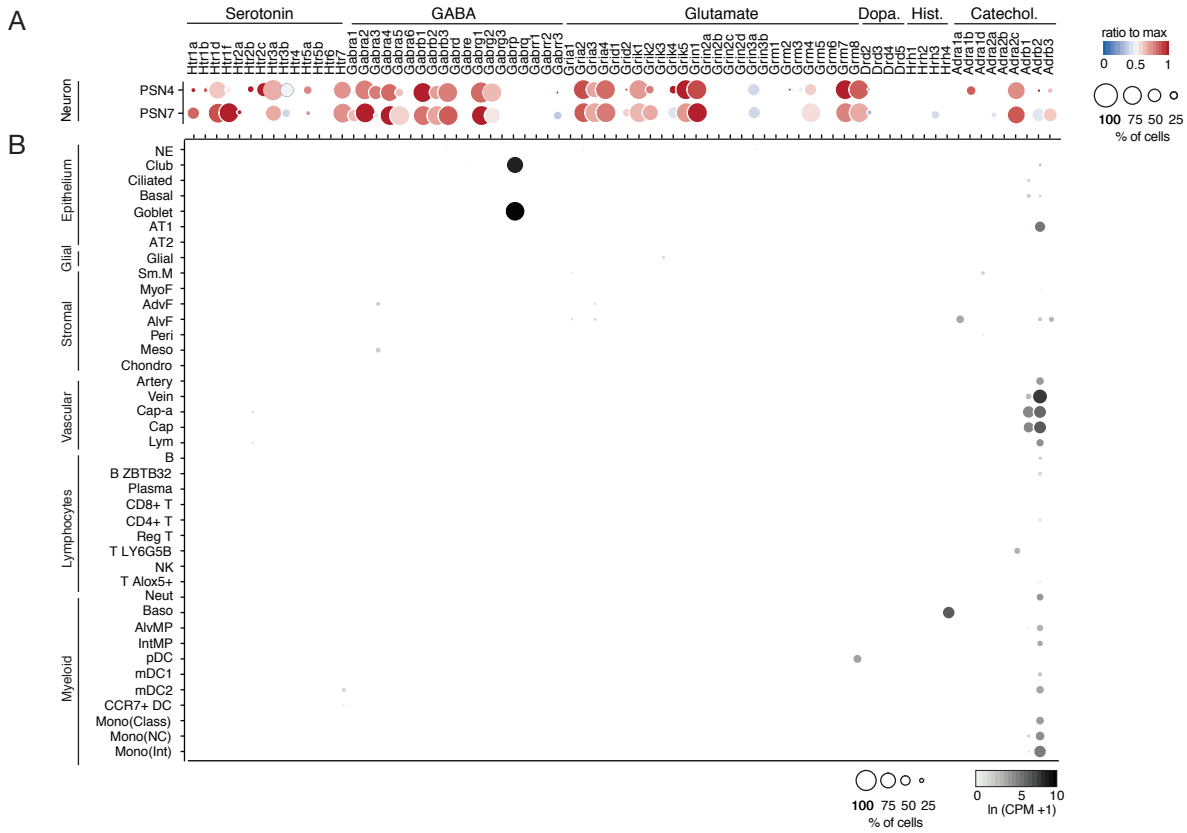


Fig. S5, Kuo et al

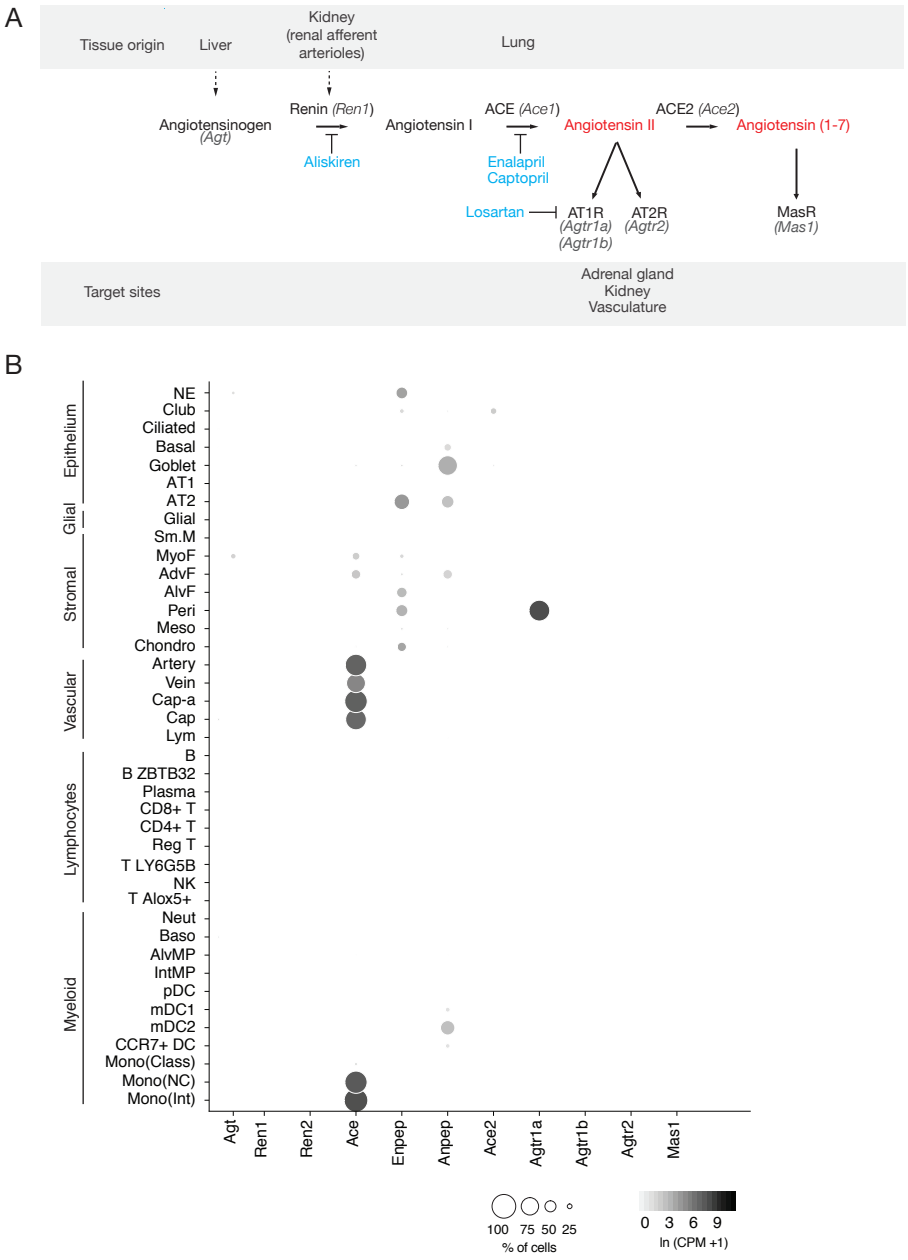


Fig. S6A, Kuo et al

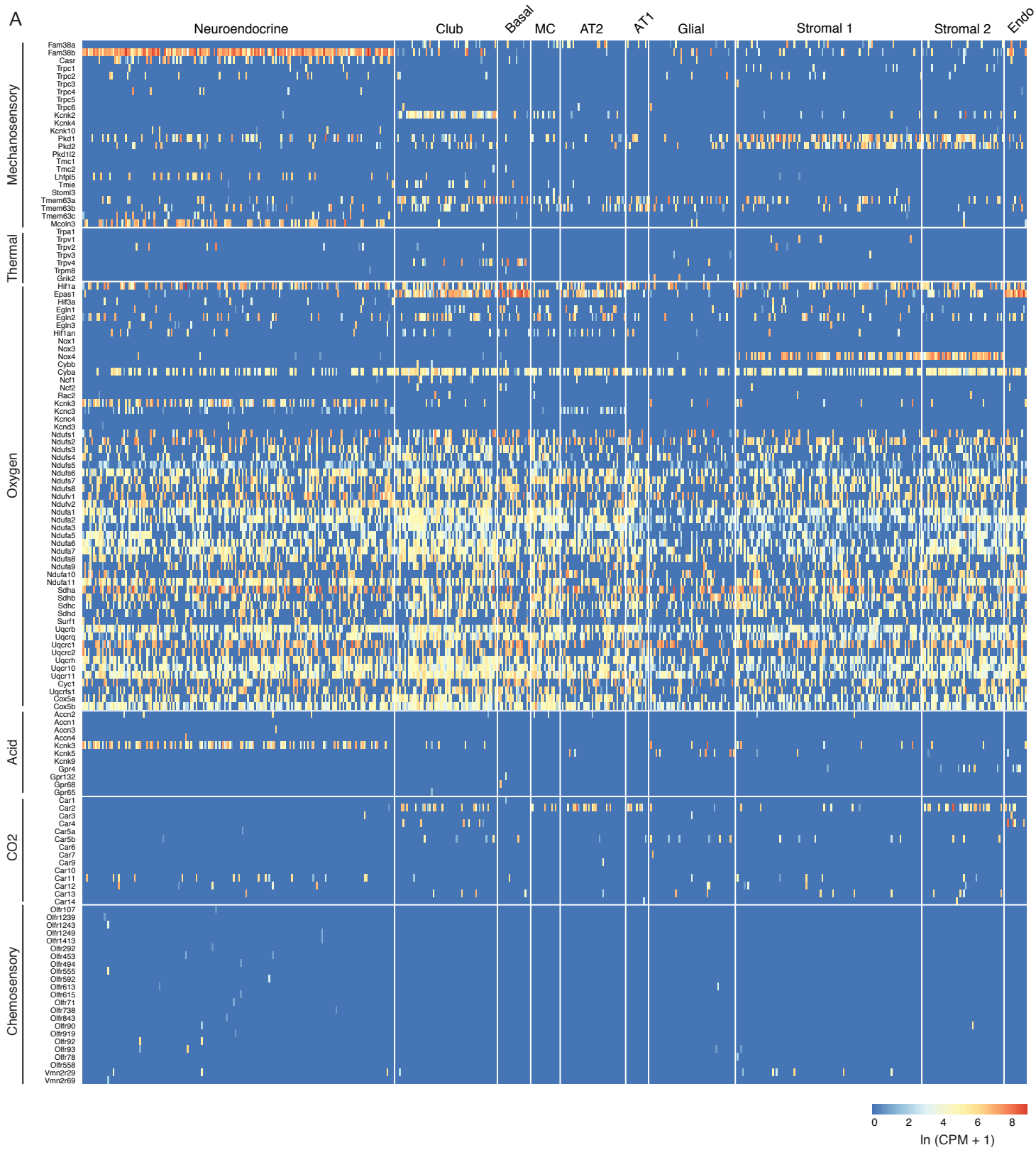


Fig. S6BC Kuo, et al

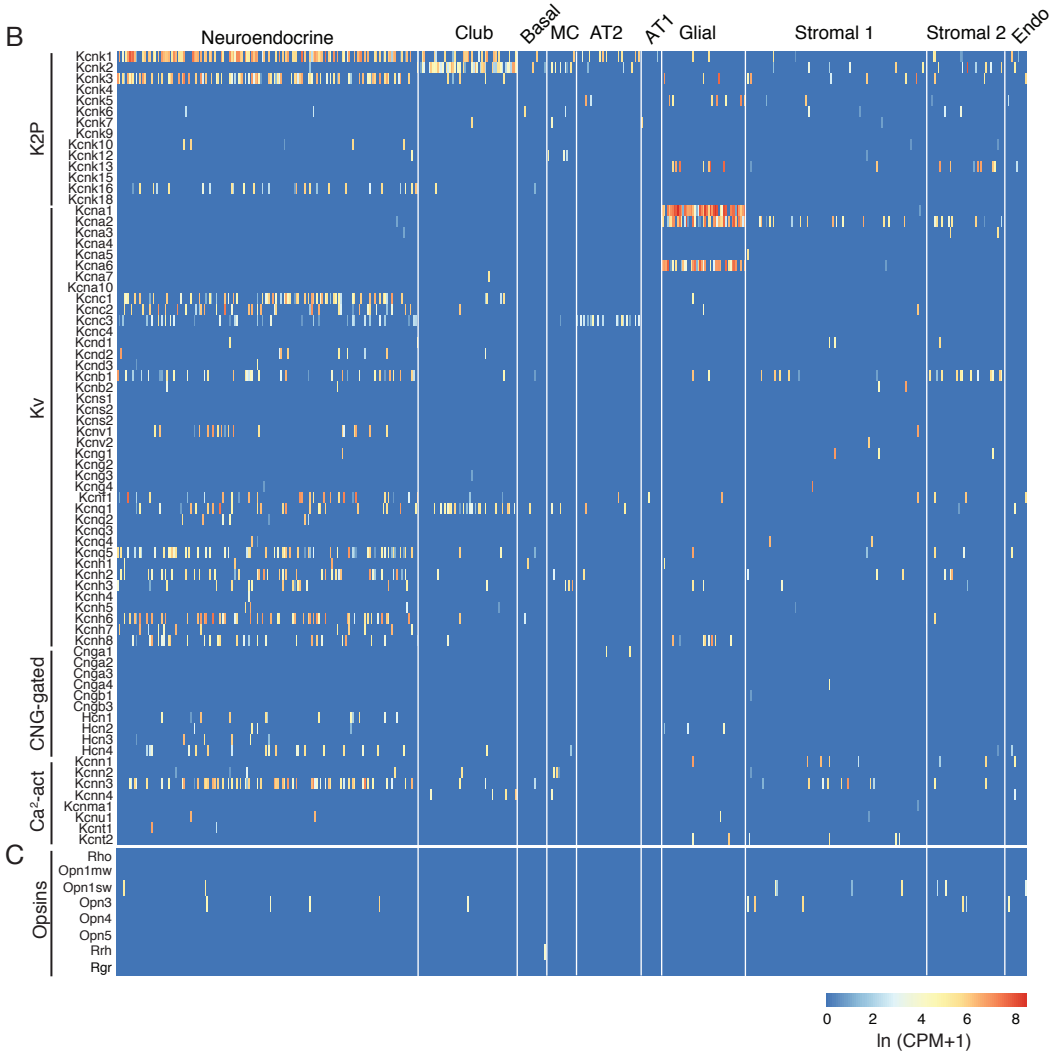


Fig. S7, Kuo et al

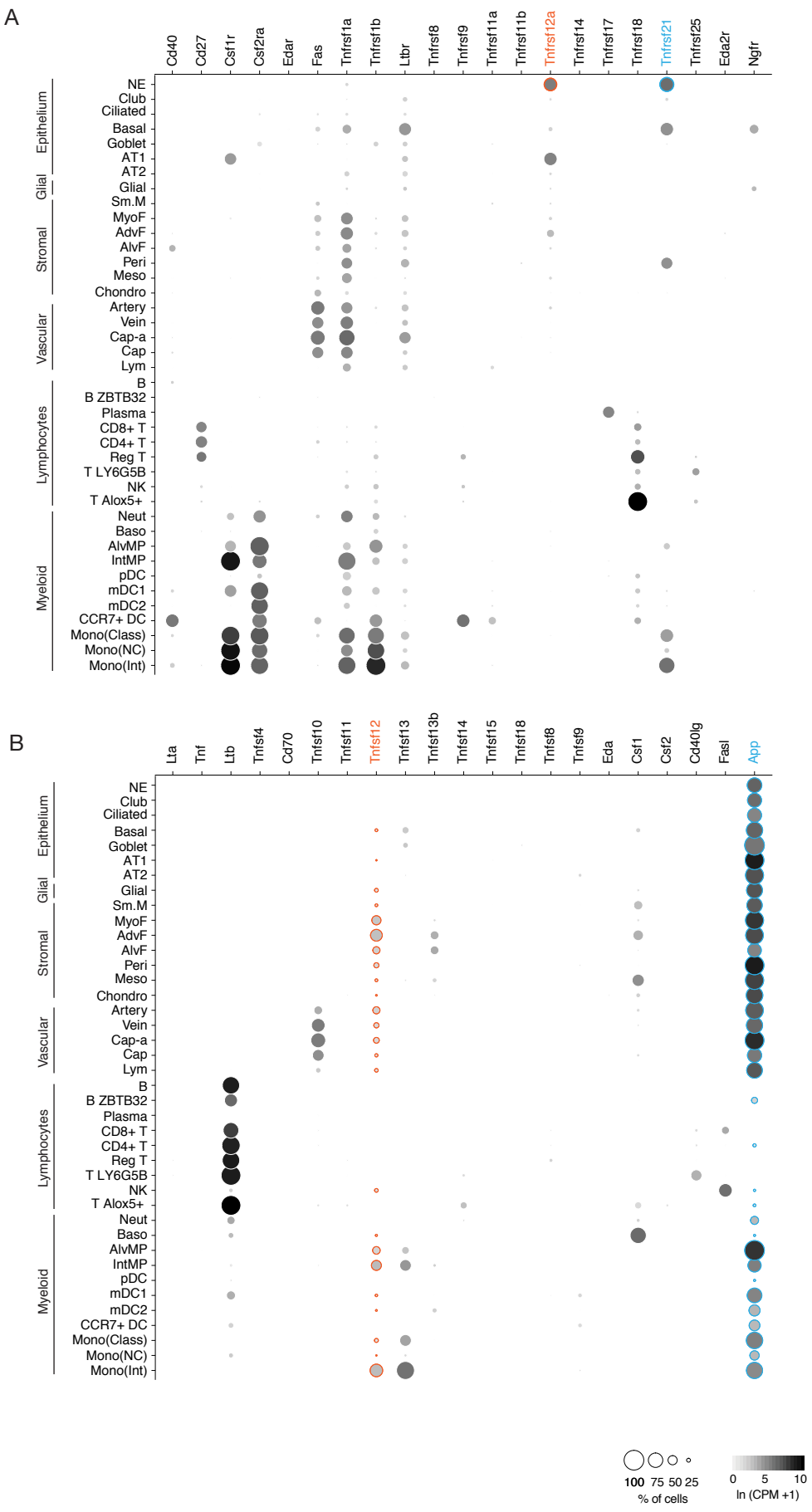


Fig. S8, Kuo et al

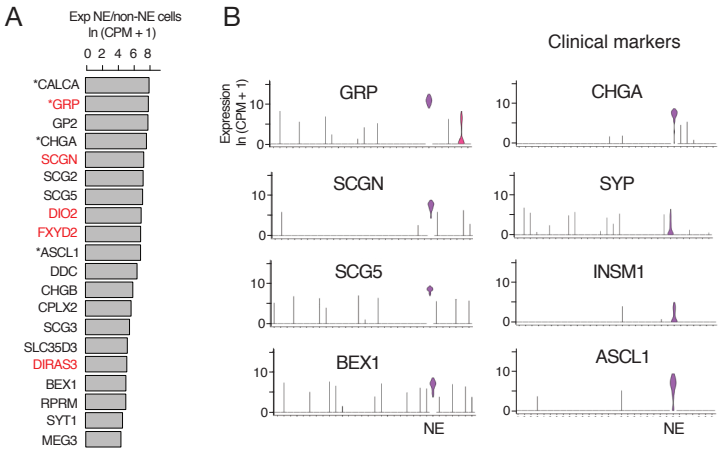
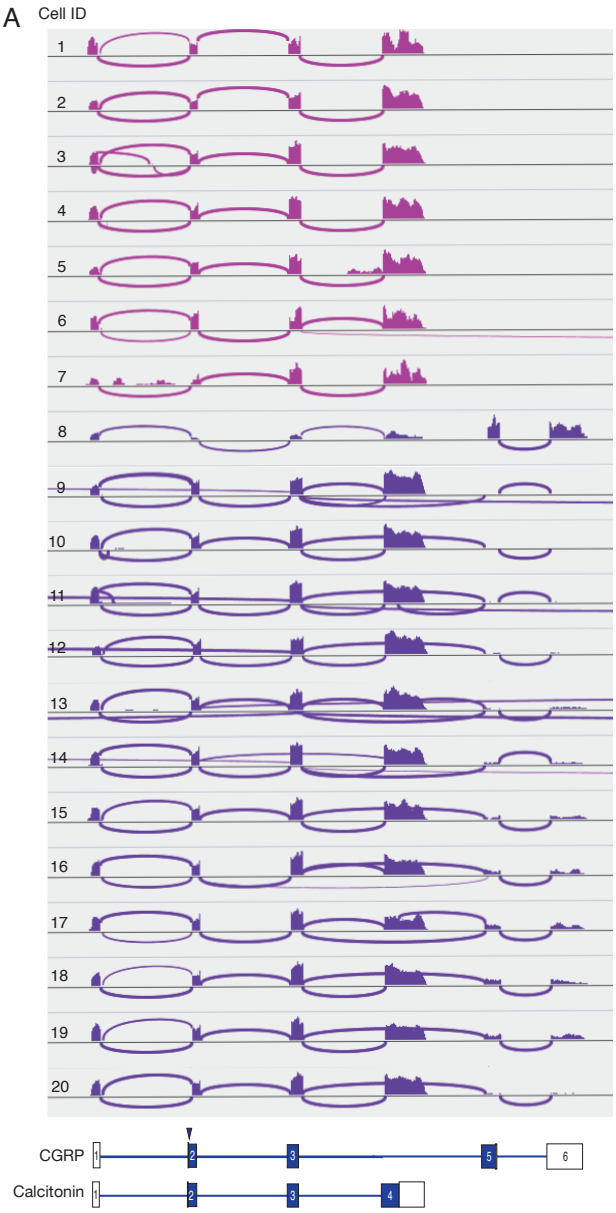


Fig. S9, Kuo et al

A



B

Cell ID	CGRP* (%)	Calcit.* (%)	Alt. splice site reads	Total reads	% Alt. splice site reads
1	0	100	72	249794	0.03
2	0	100	2355	1340758	0.18
3	0	100	391	454161	0.09
4	0	100	267	283545	0.09
5	0	100	7	553757	0.00
6	0	100	145	840597	0.02
7	0	100	67	810614	0.01
8	0**	100	233	318591	0.07
9	1	99	4278	840597	0.51
10	1	99	9634	1325068	0.73
11	2	98	11628	1322355	0.88
12	6	94	2963	839487	0.35
13	6	94	2464	829435	0.30
14	10	90	1497	758444	0.20
15	11	89	1667	661810	0.25
16	11	89	323	1404385	0.02
17	20	80	1780	746385	0.24
18	23	77	343	1245217	0.03
19	25	75	735	413093	0.18
20	29	71	290	52174	0.56

*Percent CGRP or calcitonin (Calcit.) mRNA determined by alternative (Alt.) splice site reads
**Note CGRP exonic reads (but no splice site reads) in panel A

Fig. S10, Kuo et al

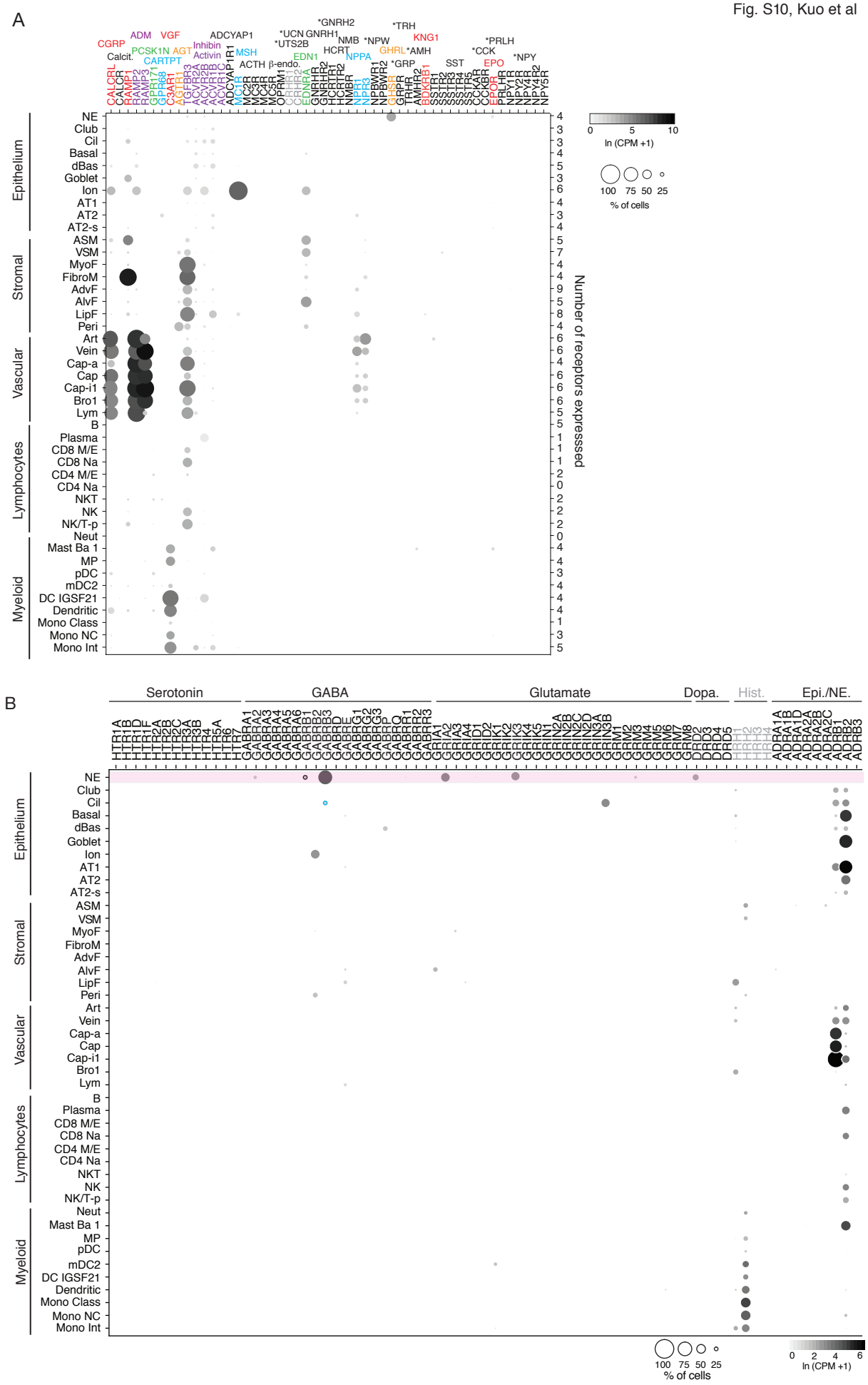


Fig. S11, Kuo et al

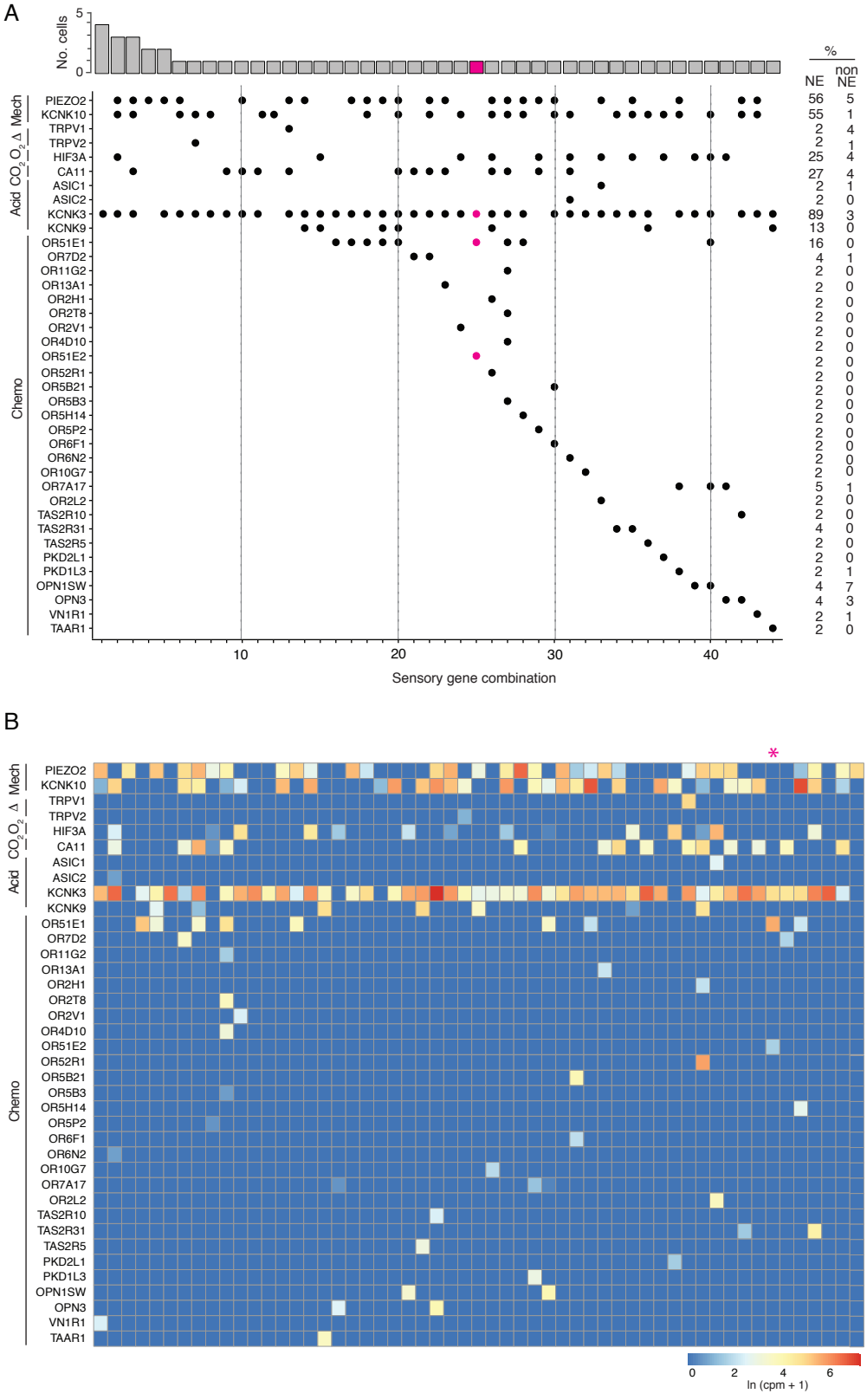


Fig. S12, Kuo et al

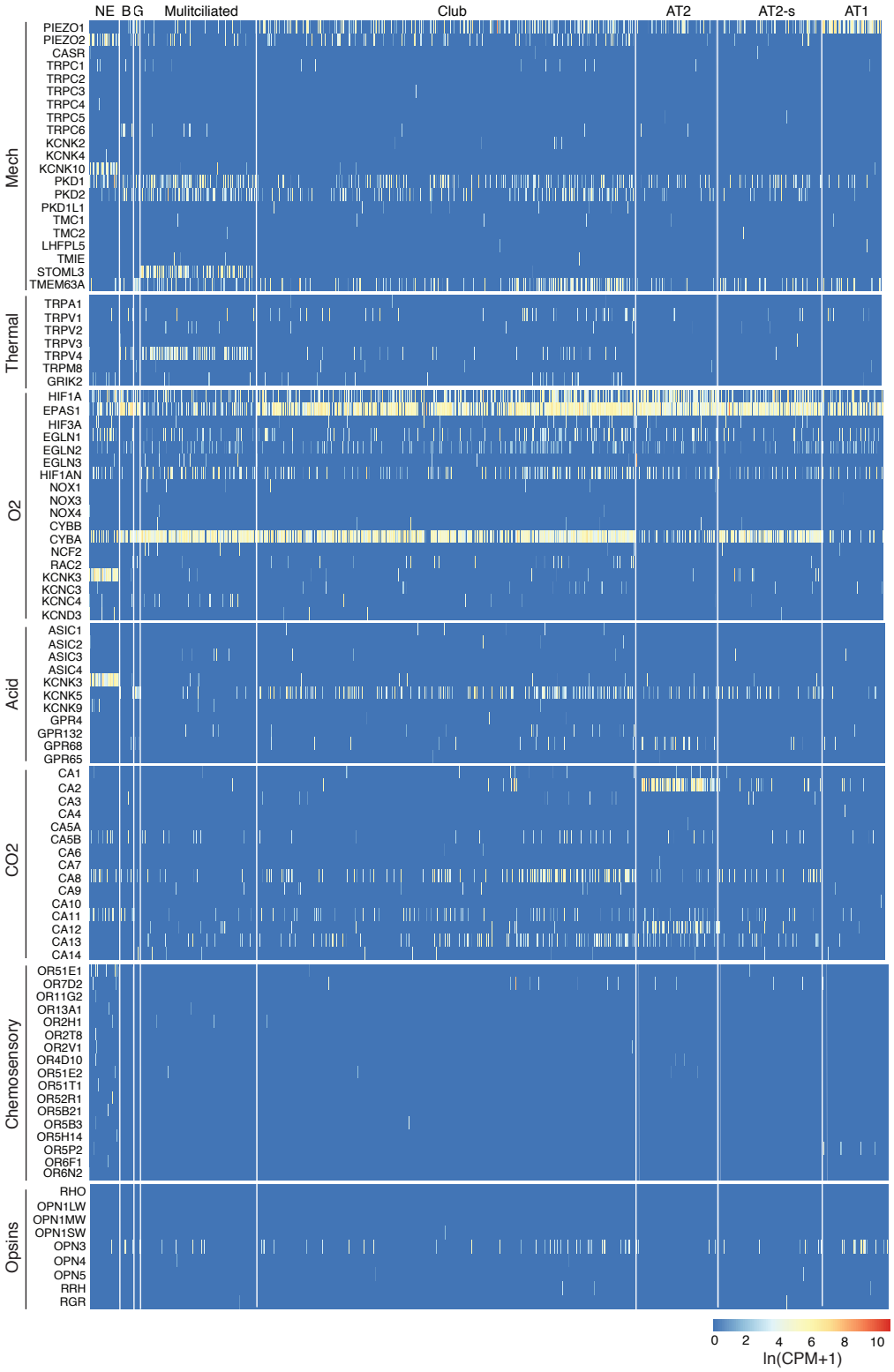


Fig. S13, Kuo et al