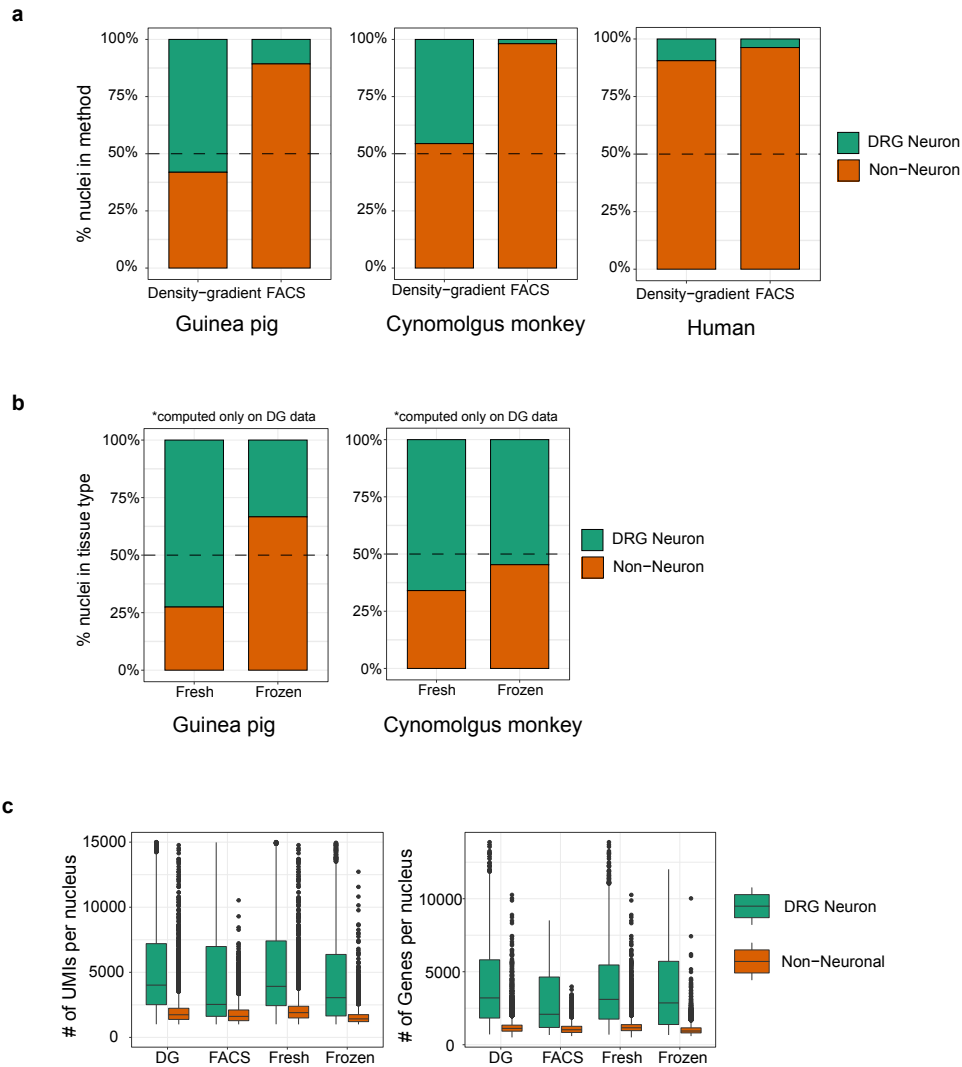


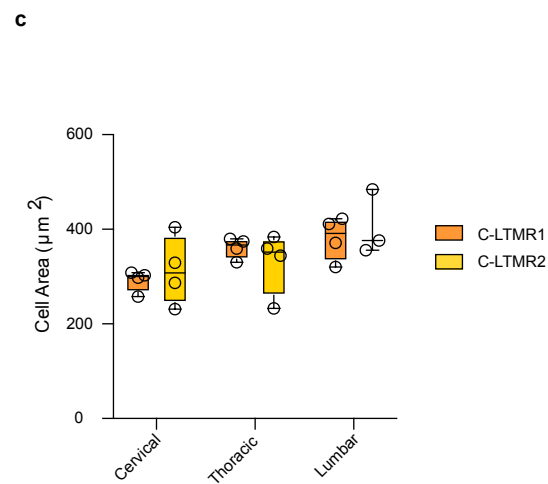
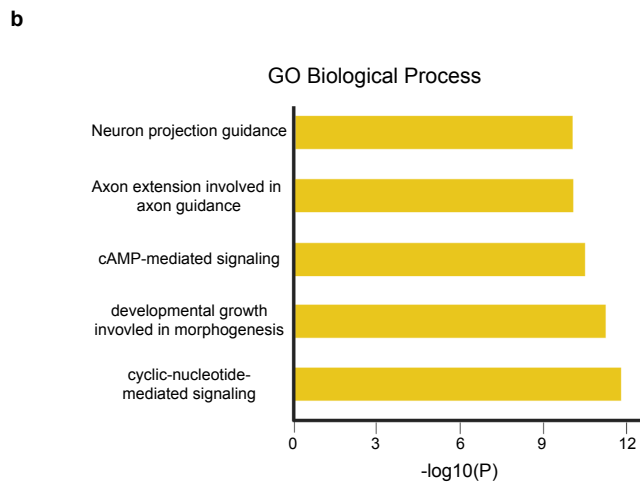
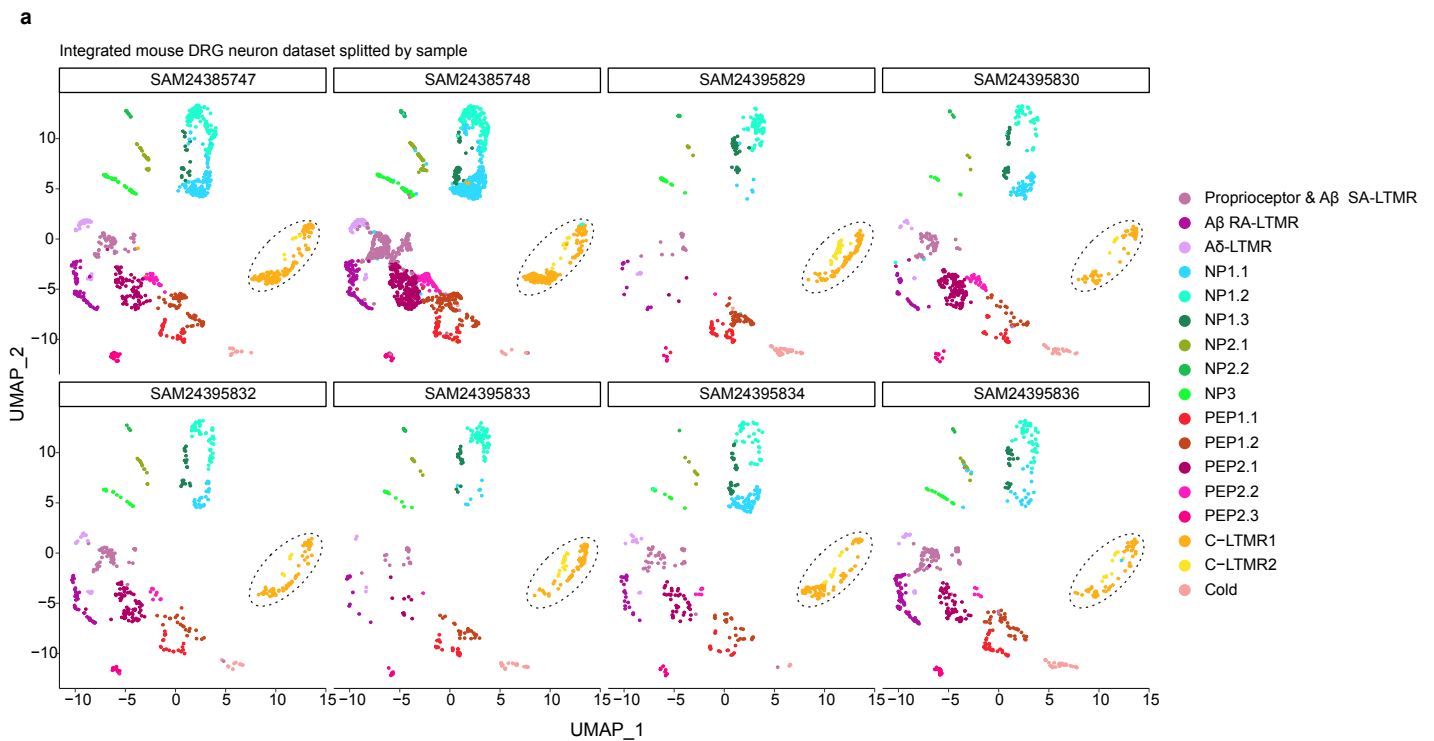
## **Supplementary Information**

### **Cross-species transcriptomic atlas of dorsal root ganglia reveals species-specific programs for sensory function**

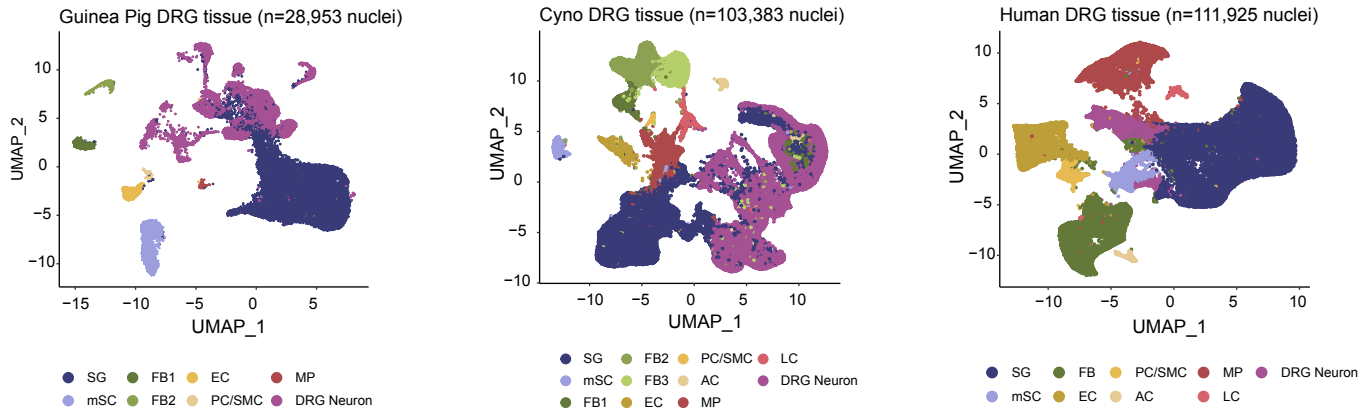
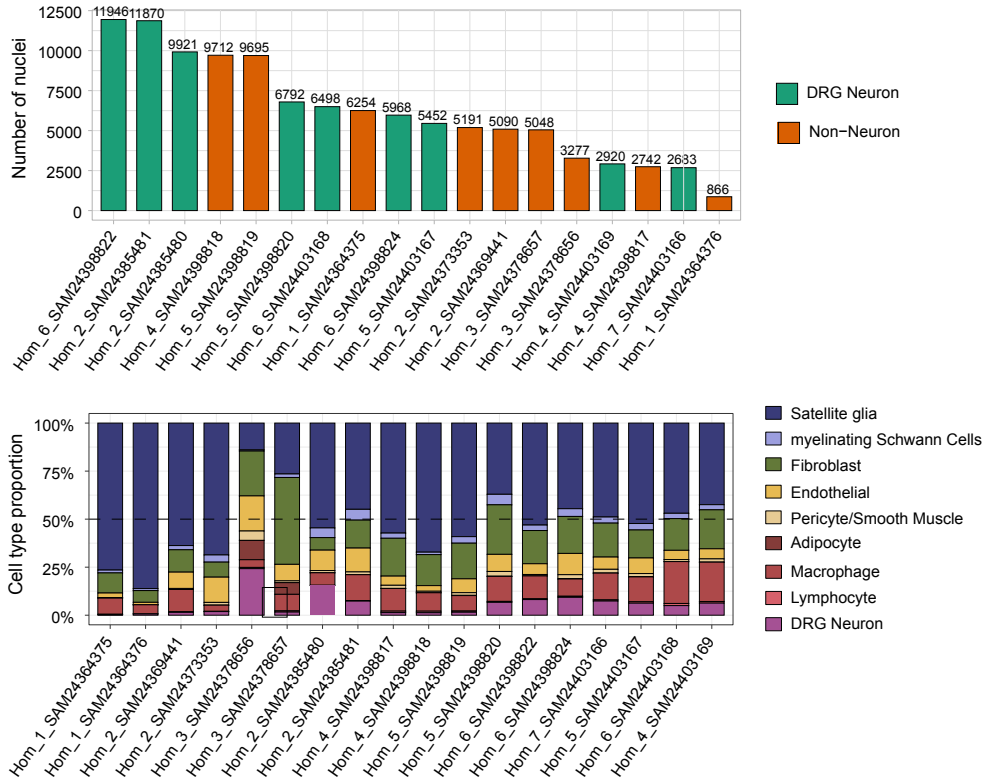
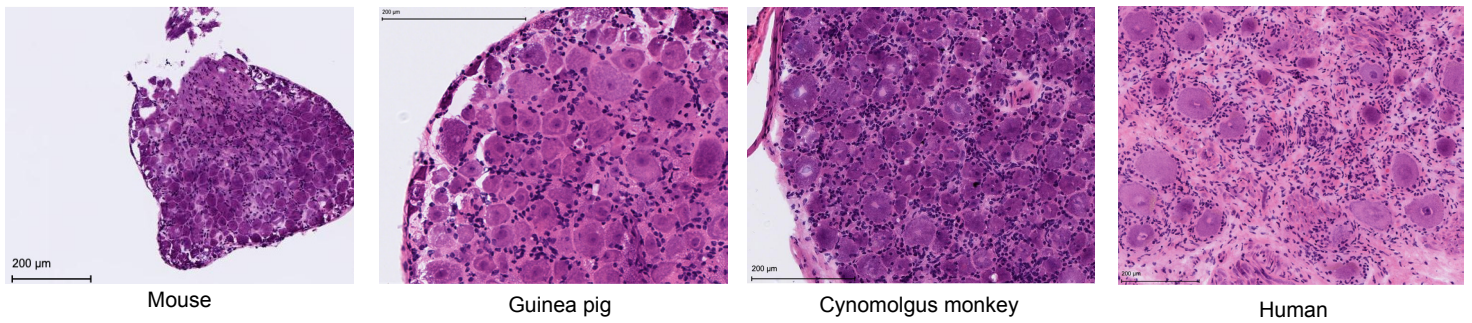
Jung & Dourado et al.



**Supplementary Fig. 1 Comparison of nuclei isolation method and tissue types for dorsal root ganglia single-nucleus RNA-seq. a,b** Relative cell-type composition for the nuclei isolation methods or tissue types in human, cynomolgus monkey, and guinea pig DRG data. **c** Number of UMIs and genes detected per nucleus for different nuclei isolation methods and tissue types in mouse DRG data.  $n=37,384$  nuclei examined over 10 independent experiments. The lower and upper hinges of the boxes correspond to the 25th to 75th percentile with the line in the middle depicting the median. The whiskers are set at the minimum and maximum values of the dataset. Density-gradient is abbreviated as DG.

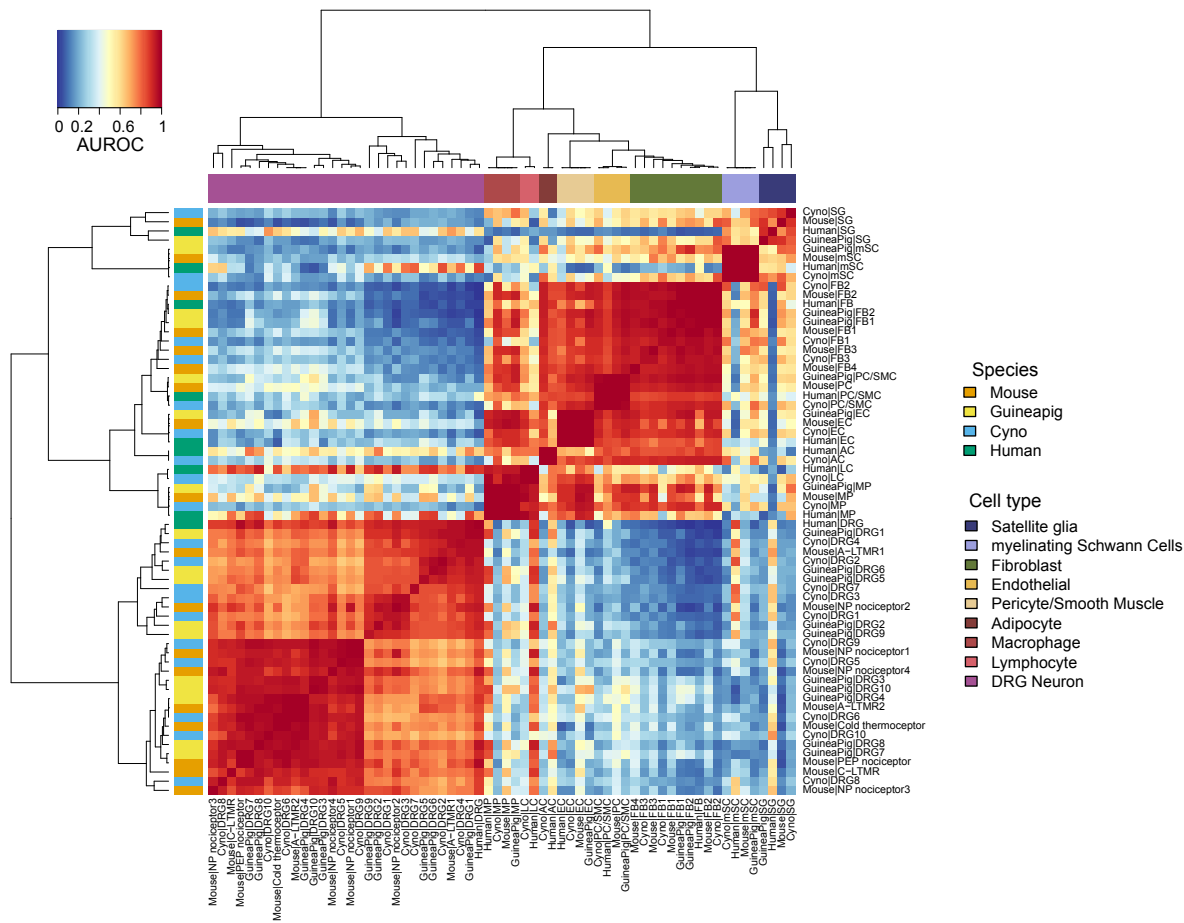


**Supplementary Fig. 2 single-nucleus RNA-seq analysis of mouse C-LTMRs** **a** Integrated mouse DRG snRNA-seq data faceted by samples. Dotted circles indicated C-LTMR clusters in each sample. NP denotes non-peptidergic; PEP denotes peptidergic. **b** Gene ontology analysis of genes enriched in the C-LTMR2 subtype. **c** Cell size distribution of C-LTMR subtypes by mouse DRG levels. N = 4 mice. The lower and upper hinges of the boxes correspond to the 25th to 75th percentile with the line in the middle depicting the median. The whiskers are set at the minimum and maximum values of the dataset. Source data are provided as a Source Data file.

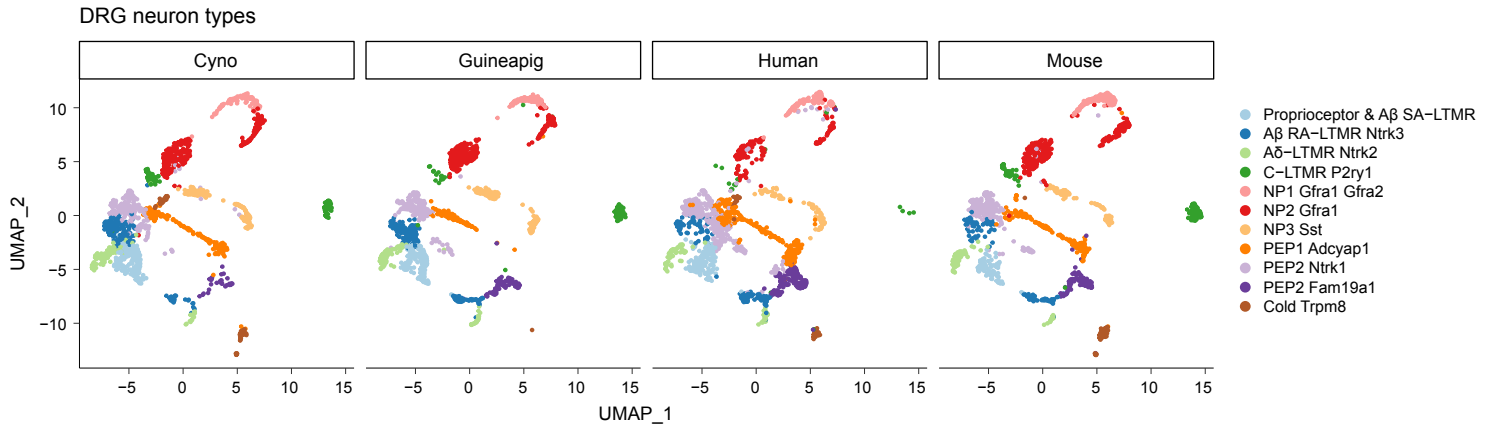
**a****b****c**

**Supplementary Fig. 3 Single-nuclei RNA-seq atlas of dorsal root ganglia from guinea pig, cynomolgus monkey, and human.** **a** UMAP of guinea pig, cynomolgus monkey, and human DRG nuclei colored by cell types and annotated post hoc. We profiled 28,953 nuclei from guinea pig DRGs, 103,383 nuclei from cynomolgus monkey DRGs, and 111,925 nuclei from human DRGs. SC denotes satellite glia; mSC denotes myelinating Schwann Cells; FB denotes fibroblasts; EC denotes endothelial cells; SMC/PC denotes smooth muscle cells and pericytes; AC denotes adipocytes; MP denotes macrophages; LC denotes lymphocytes. **b** Number of nuclei and cell type proportions recovered from human donors. **c** H&E images of DRG tissues from mouse, guinea pig, cynomolgus monkey and human with scale bars. The H&E experiment was repeated with 3 different samples for each species with similar results.

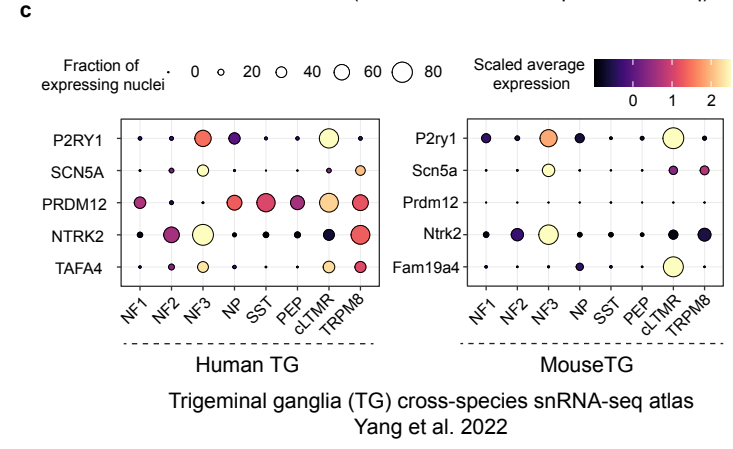
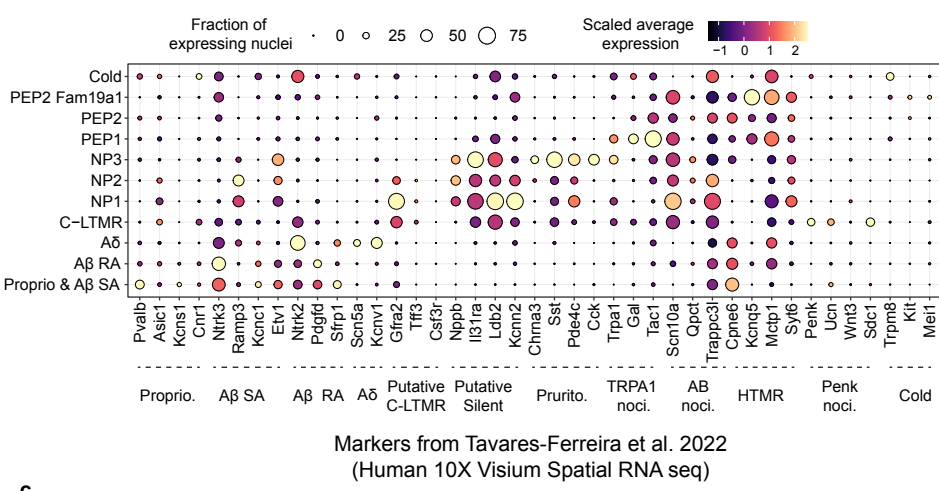
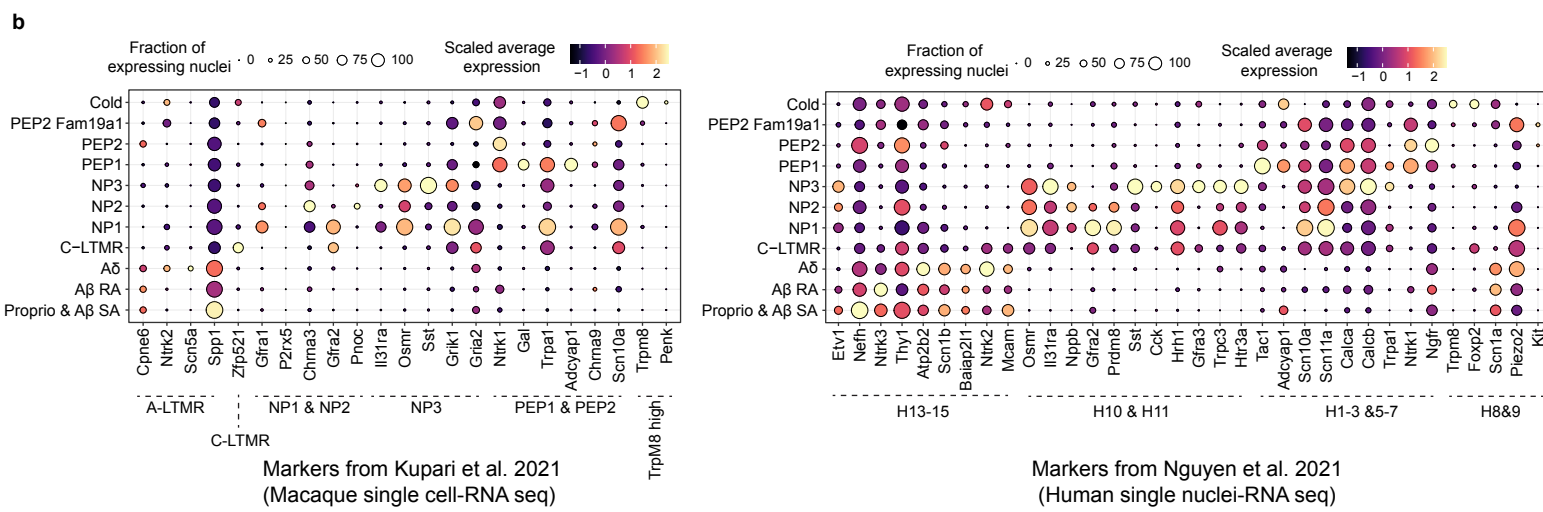
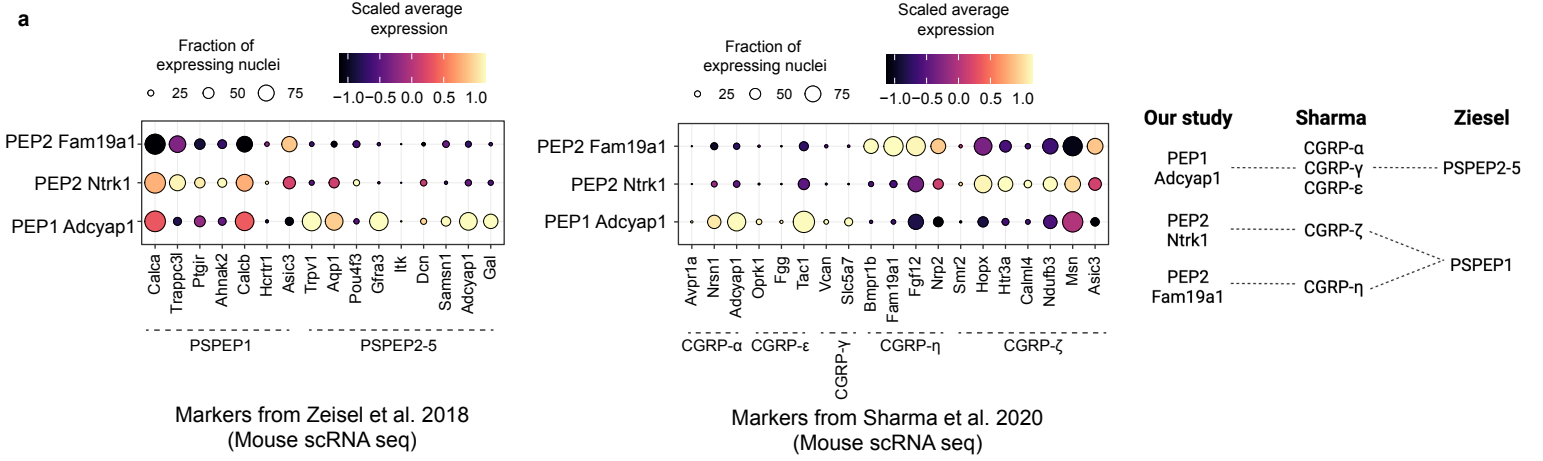
a



b



**Supplementary Fig. 4 Integrative cross-species single-nuclei RNA-seq analysis of dorsal root ganglia from guinea pig, cynomolgus monkey, and human.** **a** Cross-species dendrogram and heatmap of MetaNeighbor mean AUROC (area under the receiver operator characteristic curve) scores (y-axis in plot) color bar indicates cell types in DRGs (top) and species (left). **b** UMAP of Seurat-integrated DRG neurons across mouse, guinea pig, cynomolgus monkey, and human, faceted by species and colored by DRG neuron subtypes. NP denotes non-peptidergic; PEP denotes peptidergic.



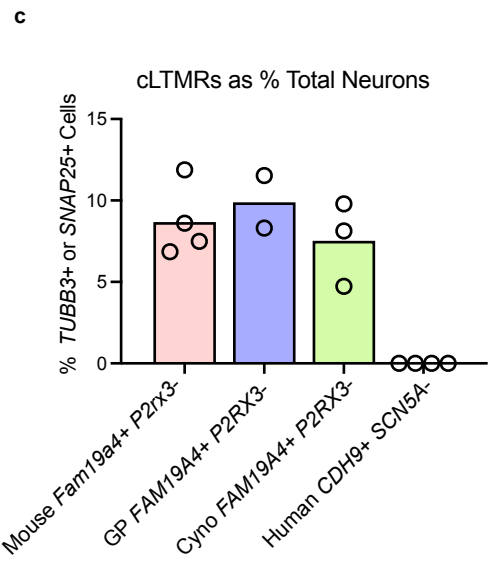
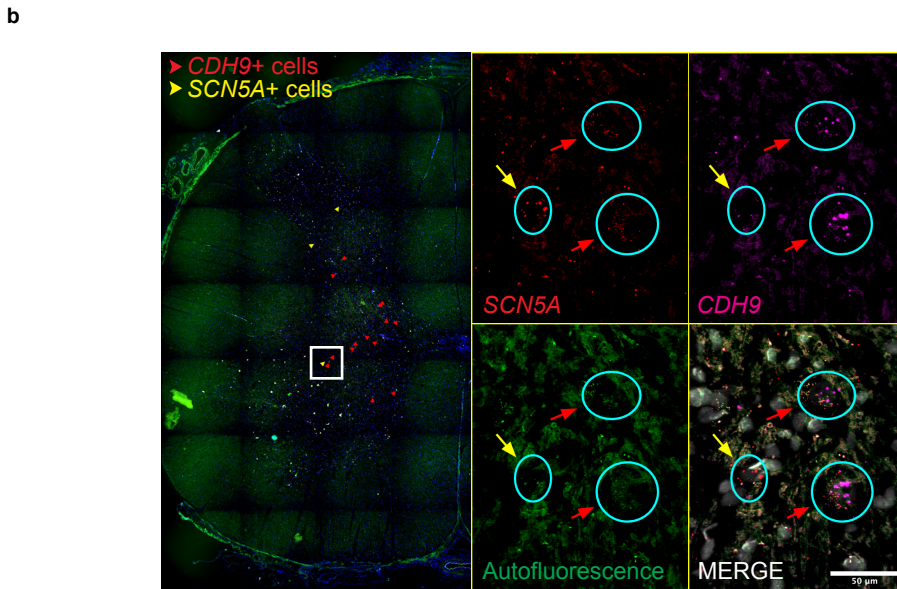
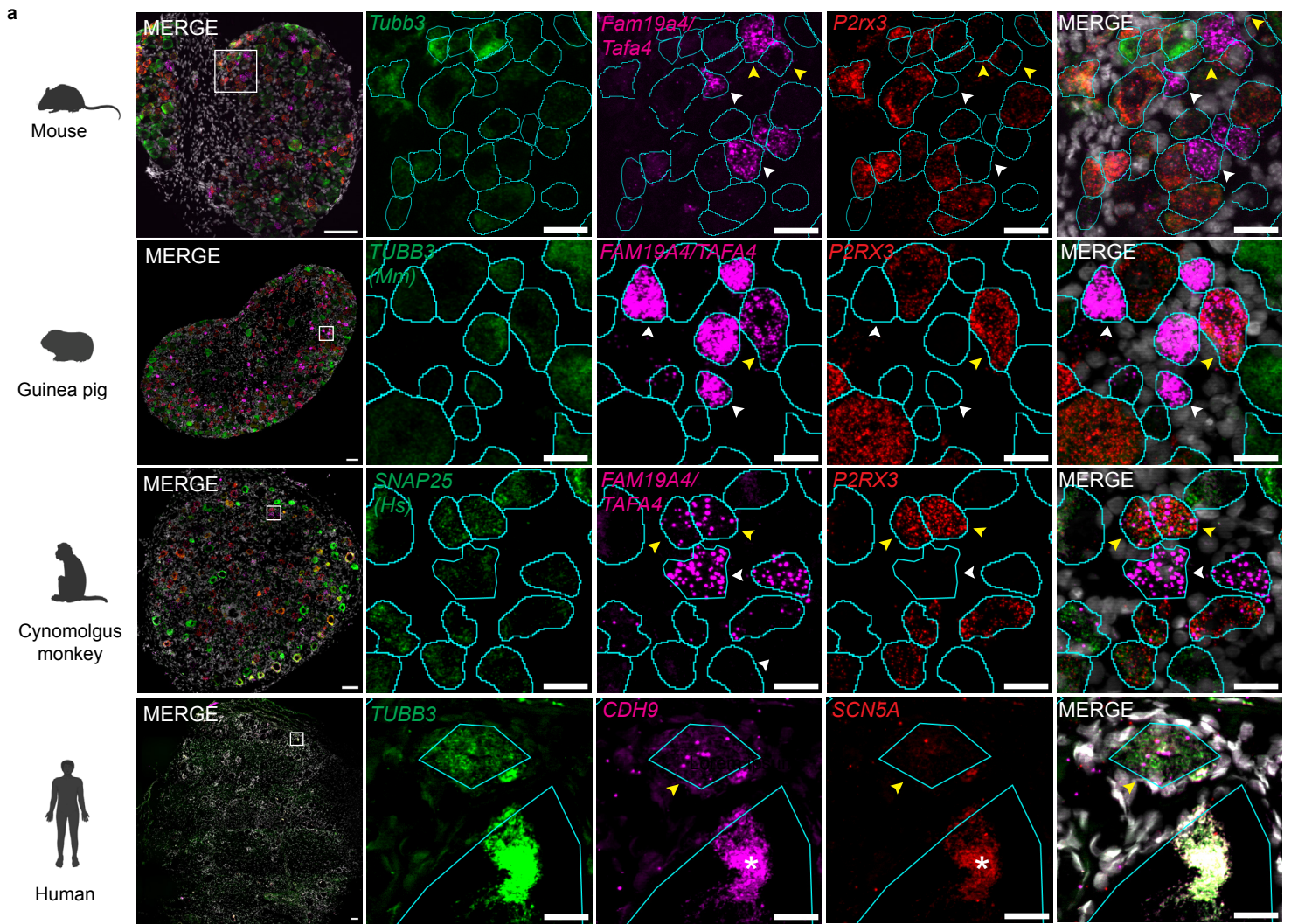
**Supplementary Fig. 5 Expression of markers from external sensory neuron expression data.**

**a-c** Fraction of nuclei (dot size) in each neuron subtype expressing markers from external sensory neuron studies (columns) and their scaled average expression level in expressing nuclei (dot color). NP denotes non-peptidergic; PEP denotes peptidergic. **a** Expression of peptidergic marker genes defined by external mouse DRG studies in our mouse peptidergic cell populations data. We identified that: our PEP2 Ntrk1 cluster corresponds to CGRP- $\zeta$  (Sharma) and PSPEP1 (Zeisel) clusters; our PEP2 Fam19a1 cluster corresponds to CGRP- $\eta$  (Sharma) and PSPEP1 (Zeisel) clusters; and our PEP1 Adcyap1 cluster corresponds to CGRP- $\alpha$ ,  $\epsilon$  and  $\gamma$  (Sharma) and PSPEP2-5 (Zeisel).

**b** Expression of marker genes from recently published primate analyses in our primate data.

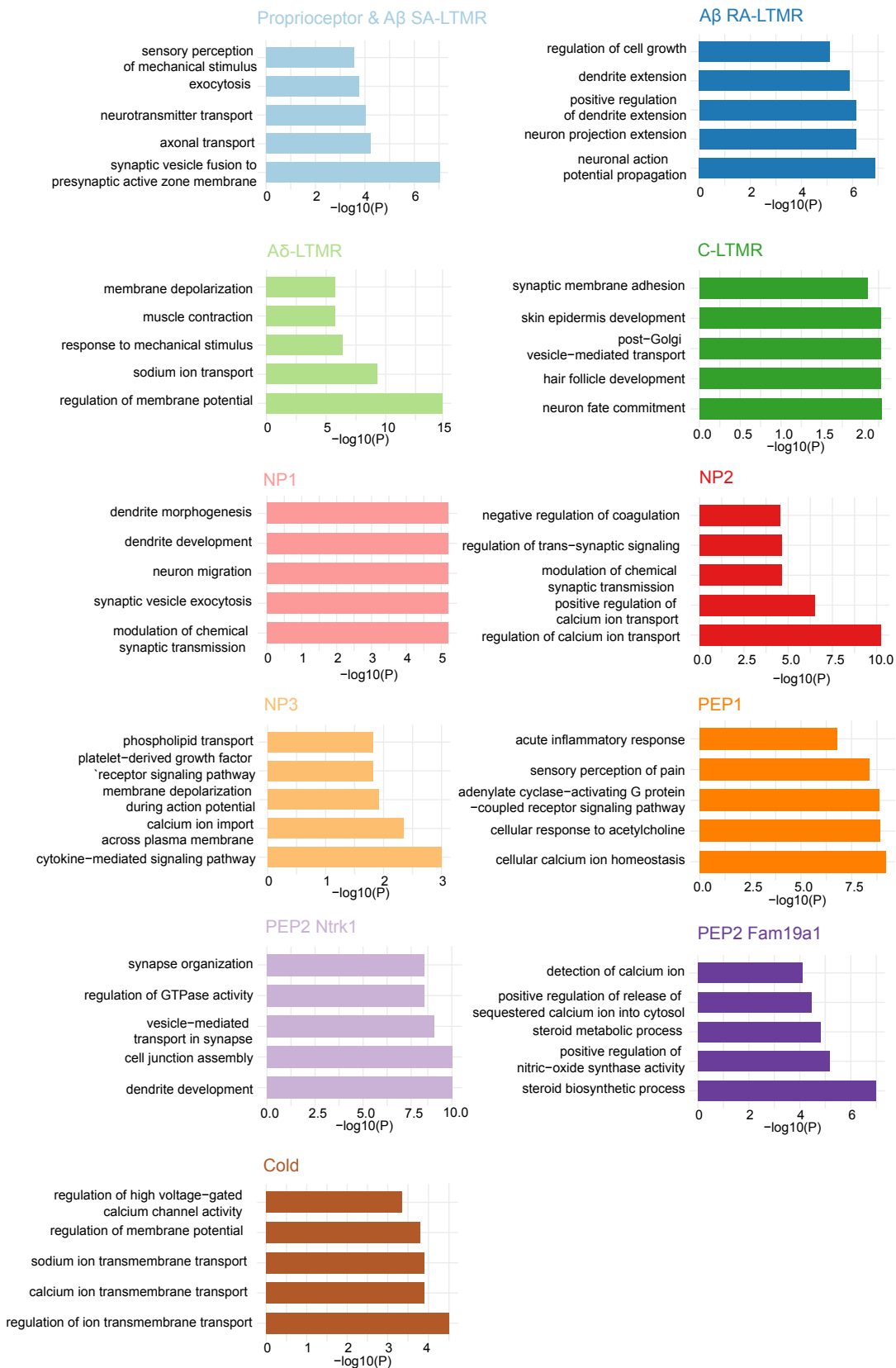
**c** Expression of TAF4A, A $\delta$  and C-LTMR specific markers (as identified in this study) shown in recently published human and mouse trigeminal ganglion single-nuclei RNA-seq atlas.





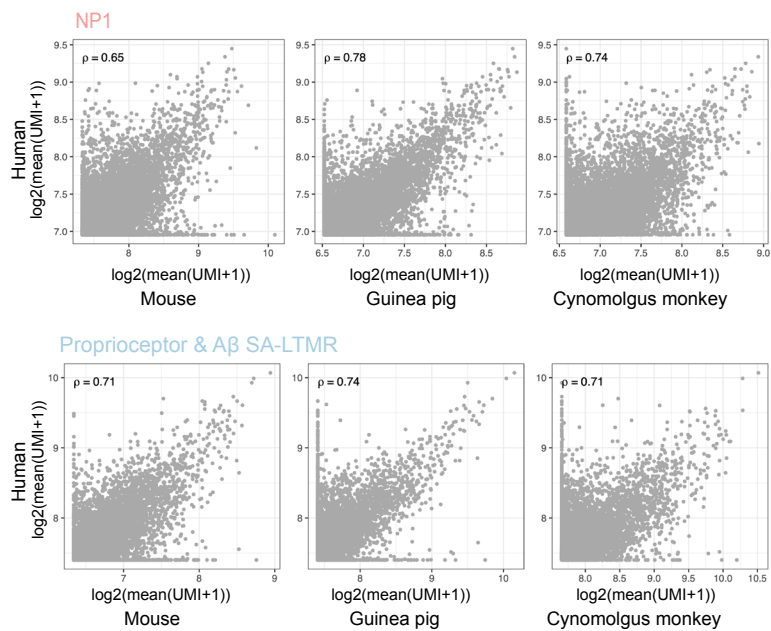
**Supplementary Fig. 6 RNAScope validation of C-LTMRs in mouse, guinea pig, cynomolgus monkey, and human DRGs.** **a** RNAScope validation of *TUBB3* (mouse, guinea pig, and human) or *SNAP25* (cynomolgous monkey), *FAM19A4*, and *P2RX3* (mouse, guinea pig, cynomolgous monkey) or *CDH9* and *SCN5A* (human) in DRGs. The scale bar in the whole DRG image indicates 100µm. Neurons outlined in turquoise. The scale bar in the zoomed images indicate 25µm. Yellow arrowheads indicate *FAM19A4+* and *P2RX3+* or *SCN5A* cells. White arrowheads indicate *FAM19A4+* and *P2RX3-* cells. Asterisks in the human images indicate lipofuscin aggregates. **b** Positive control of *CDH9* probe in human spinal cord tissue. Yellow arrows indicate *SCN5A+* cells and red arrows indicate *CDH9+* cells. Autofluorescence reflects signal from the AF488 channel. The scale bar indicates 50µm. **c** Percentage of *FAM19A4+/P2RX3-* or *CDH9+/SCN5A-* cells (putative C-LTMRs) in mouse, guinea pig, cynomolgus monkey, human. No *CDH9+* cells were observed in human DRGs. Source data are provided as a Source Data file. GP denotes guinea pig and Cyno denotes cynomolgus monkey. Species icons in **a** were created with BioRender.com

a

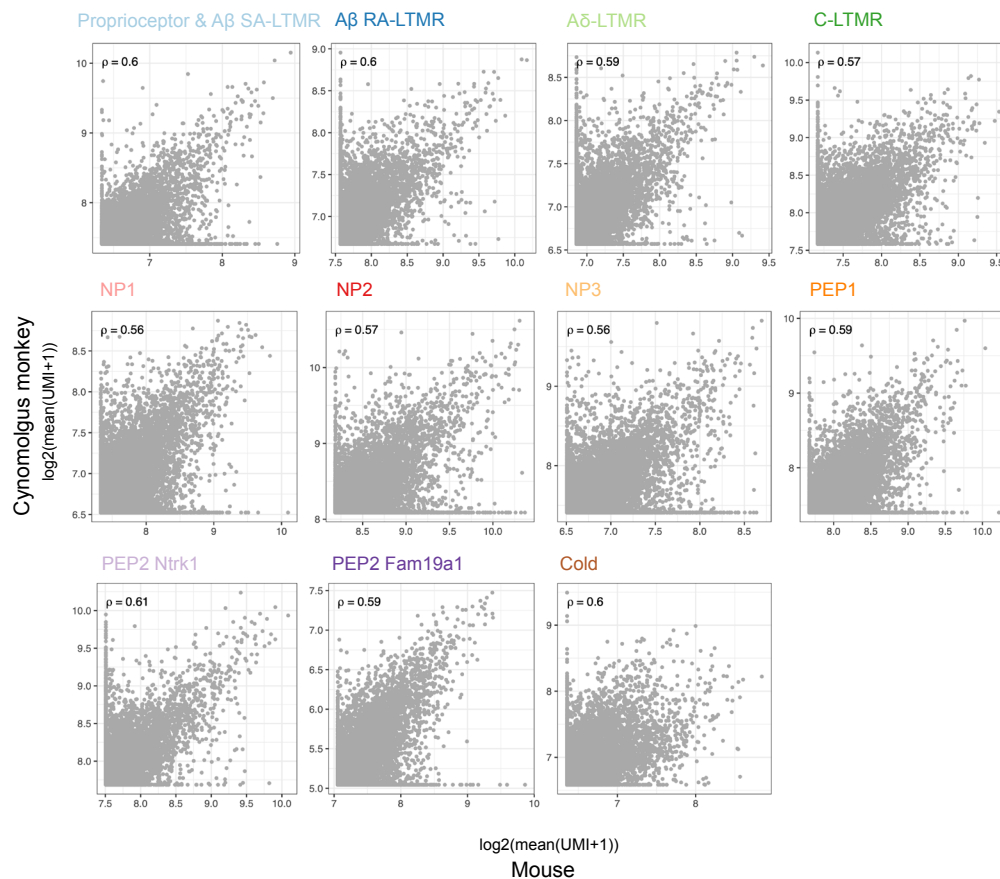


**Supplementary Fig. 7 Analysis of conserved DRG neuron subtype-specific transcriptional programs.** a Gene ontology (Biological processes) analysis of subtype-specific conserved genes. p-values are calculated based on the hypergeometric distribution which corresponds to one-sided Fisher's exact test and corrected for multiple comparisons testing using the Benjamini-Hochberg. NP denotes non-peptidergic; PEP denotes peptidergic.

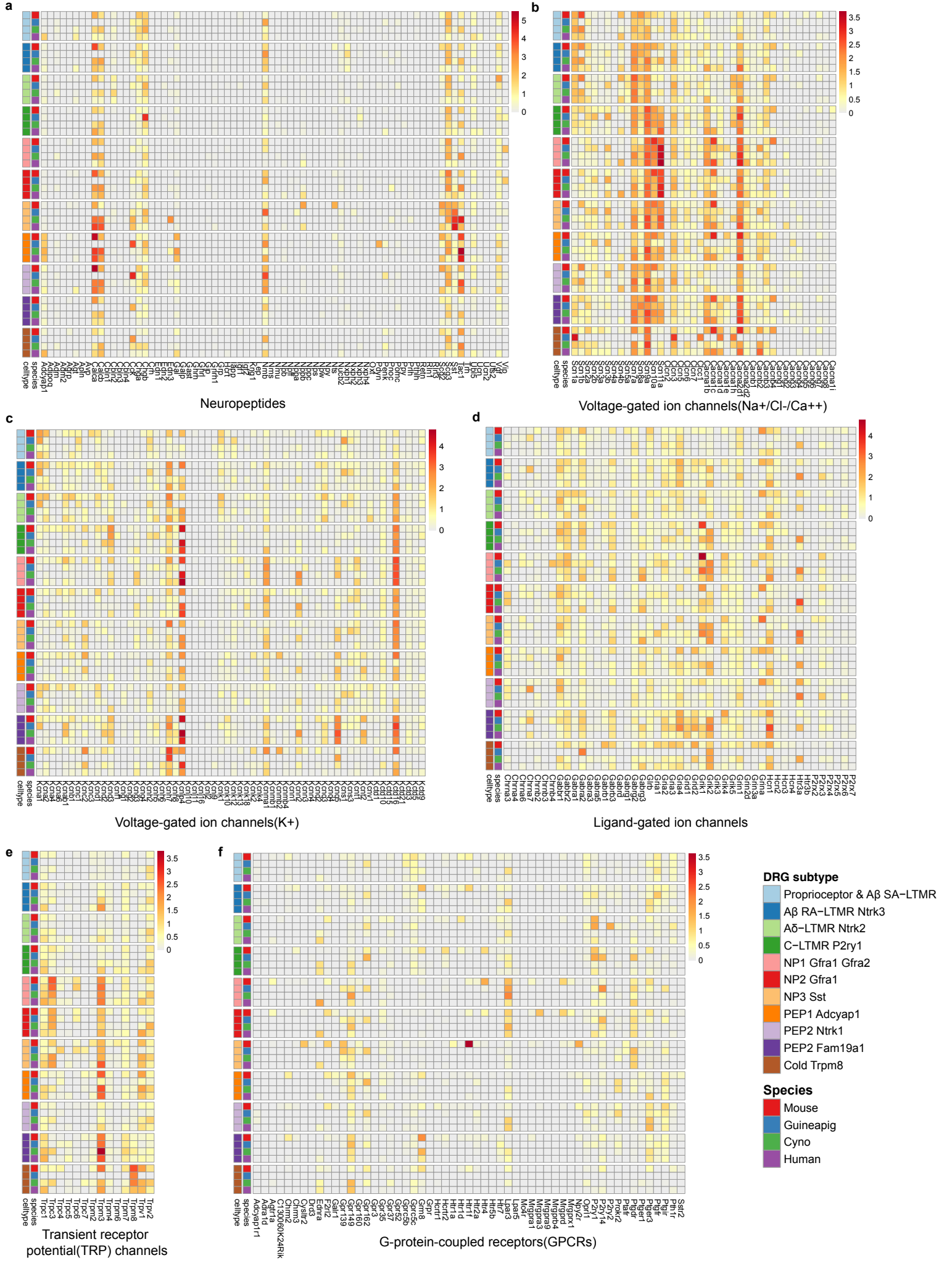
a



b

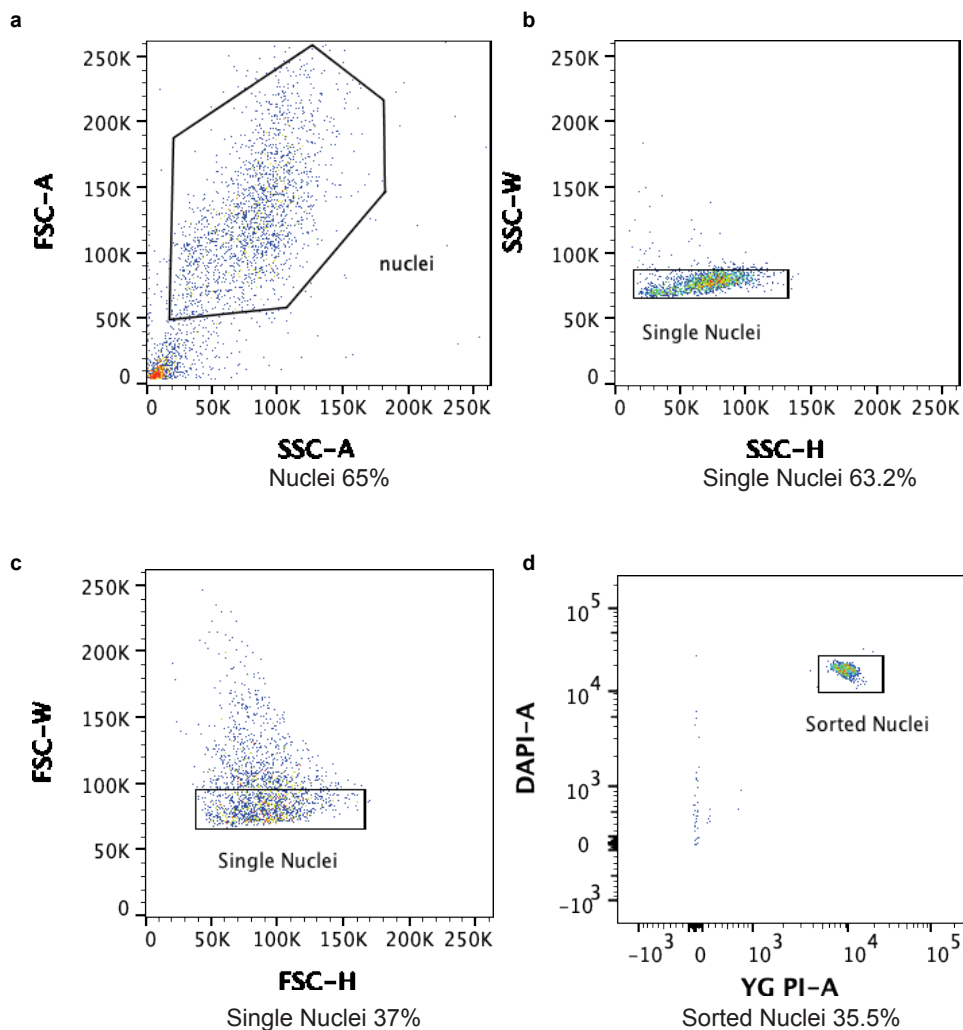


**Supplementary Fig. 8 Analysis of conserved DRG neuron subtype-specific transcriptional programs.** **a** Scatter plots of transcriptome correlation (Spearman) between human and other species for NP1 and Proprioceptor & A $\beta$  SA-LTMRs. NP denotes non-peptidergic **b** Scatter plots of transcriptome correlation (Spearman) between cynomolgus monkey and mouse for each DRG subtype. NP denotes non-peptidergic; PEP denotes peptidergic.



**Supplementary Fig. 9 Expression of molecular determinants of sensory function**

**a-f** Expression heatmap of neuropeptides, voltage-gated ion channels, transient receptor potential channels, ligand-gated ion channels and G-protein coupled receptors in each neuron subtype in mouse, guinea pig, cynomolgus monkey, and human data. NP denotes non-peptidergic; PEP denotes peptidergic.



**Supplementary Fig. 10 Flow cytometry gating strategy for sorted single nuclei.**

Single nuclei were sorted on a BD FACSAria Fusion with BD FACSDiva Software v8.0.3. The instrument was set to sort in "Purity" mode with a system pressure of 70psi on a 70 micron nozzle at 32kHz frequency. The fraction of gated nuclei is listed below each plot as a percentage of total. **a** Primary gate to select nuclei. **b** Secondary gate selecting singlets based on side scatter. **c** Tertiary gate selecting singlets based on forward scatter. **d** Sorted single nuclei selected for sequencing based on DAPI and propidium iodide (PI) labeling.