Neuron, Volume *110*

Supplemental information

Human and mouse trigeminal ganglia cell atlas implicates multiple cell types in migraine

Lite Yang, Mengyi Xu, Shamsuddin A. Bhuiyan, Jia Li, Jun Zhao, Randall J. Cohrs, Justin T. Susterich, Sylvia Signorelli, Ursula Green, James R. Stone, Dan Levy, Jochen K. Lennerz, and William Renthal

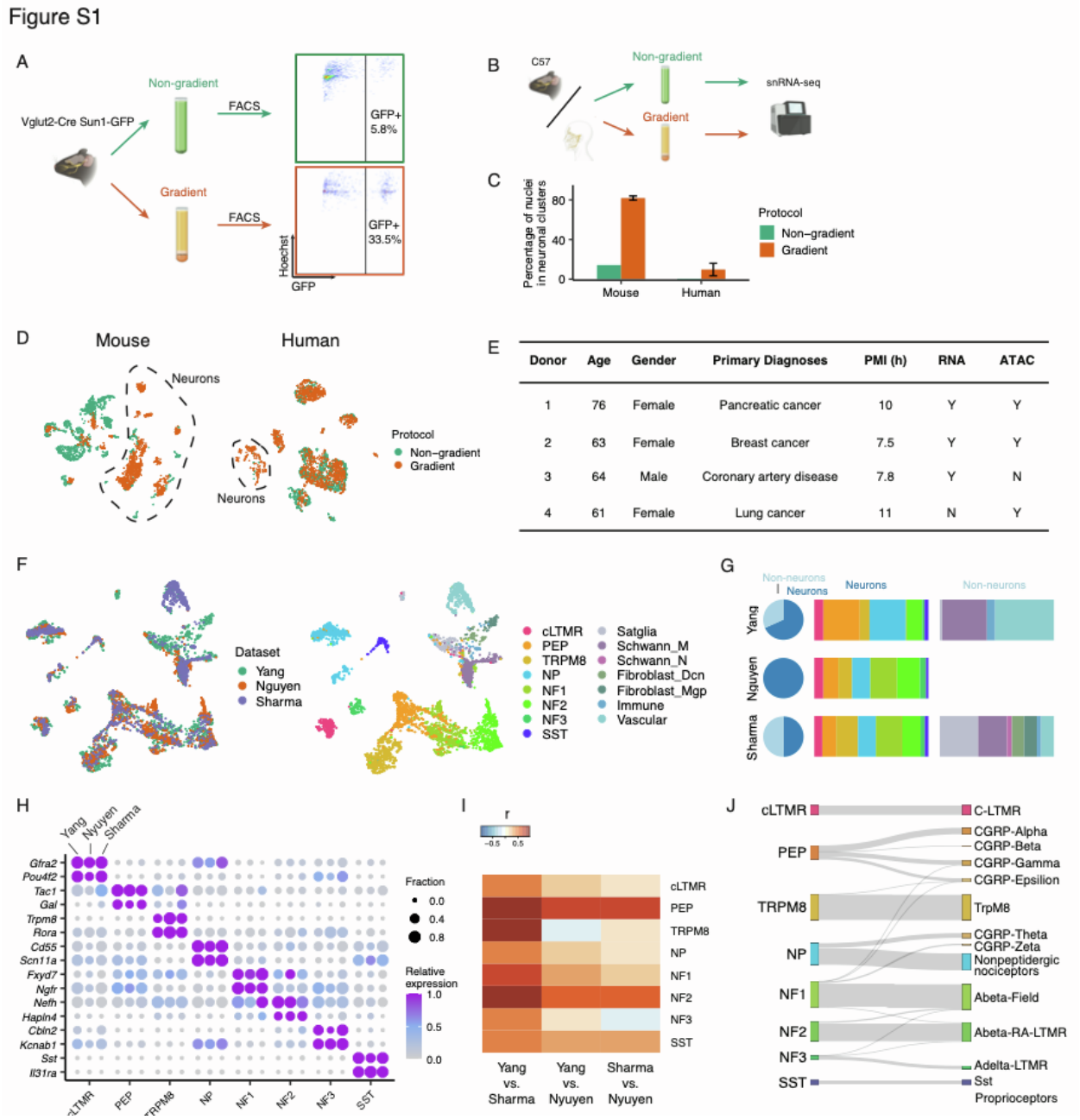

Figure S1. snRNA-seq of human and mouse trigeminal ganglion. Related to Figure 1.

A. TGs from *Vglut2-cre;Sun1-GFP* mice were dissociated into single-nuclei suspension using either the non-gradient or gradient method and the fraction of neurons are quantified (GFP+ events/total Hoechst+ events). Plot is representative of 3 biological replicates.

B. Experimental approach for dissociating TGs from *Vglut2-cre;Sun1-GFP* mice or human into single-nuclei suspension using either the non-gradient or gradient method followed by snRNA-seq.

C. Quantification of the percent neuronal nuclei obtained using the non-gradient or gradient method as measured by snRNA-seq of TG from C57Bl/6 mice or human. See methods for details on classification of TG cell types. Note that degree of neuronal enrichment observed by snRNA-seq can differ from that observed by GFP+ FACS events of *Vglut2-Cre;Sun1-GFP* mice based on analysis methodology because neurons and non-neuronal cells vary in their mRNA complexity. Thus, the minimum number of genes detected per cell and/or doublet removal strategy can affect the final fraction of neurons obtained by snRNA-seq.

D. UMAP plot of snRNA-seq data from 18,606 naïve male mouse TG nuclei (7,164 gradient and 11,442 non-gradient) downsampled to display 3,000 nuclei (left) or 43,493 human TG nuclei (38,023 gradient, 5,465 non-gradient) downsampled to display 3,000 nuclei (right). Nuclei were isolated using either the non-gradient or gradient protocol (see methods). Colors represent nuclear isolation protocol.

E. Table of human TG donor characteristics including the donor number that corresponds to the snRNA-seq data, age, gender, primary diagnosis, and post-mortem interval.

F. UMAP plots of RNA-seq data from 29,119 mouse TG nuclei (Yang – naïve male TGs from this study, Nguyen) or cells (Sharma), downsampled to display 1,500 nuclei from each study. UMAP plots are colored by the study (left) or cell type (right).

G. Fraction of neuronal and non-neuronal cell types and subtypes in each mouse TG snRNA-seq (Yang, Nguyen) or single-cell RNA-seq study (Sharma). Pie charts indicate fraction of neurons (dark blue) and non-neurons (light blue) in each study. Horizontal bars indicate the fraction of neuronal or non-neuronal subtypes that comprise the respective cell type. Colors of bar chart indicate subtype (see S1F).

H. Dot plot of cell-type-specific marker gene (rows) expression in each TG cell type (columns) across this study (Yang et al. [left column]) and two other TG studies (Nguyen et al. [middle column], Sharma et al. [right column]). The fraction of nuclei expressing a marker gene is calculated as the number of nuclei in each cell type that express a gene (>0 counts) divided by the total number of naive nuclei in the respective cell type. Expression in each cell type is calculated as mean expression of a gene relative to the highest mean expression of that gene across all cell types.

I. Heatmap of pairwise Pearson's correlation of marker gene expression in each cell type across the three TG single-nucleus/single-cell RNA-seq studies (Yang, Nguyen, Sharma). Marker genes are defined as differentially expressed genes with Log2FC > 1 and FDR < 0.05 in one cell type compared to all other cells in each study. Marker genes identified in either of the compared studies are included in the correlation.

J. Bioinformatic classification of cell types from mouse TG snRNA-seq data generated in this study. Gray lines connect the classifications used in this study (left, direct clustering [see methods]) and the classifications obtained by anchoring nuclei to Sharma et al.'s dorsal root ganglion scRNA-seq (right).

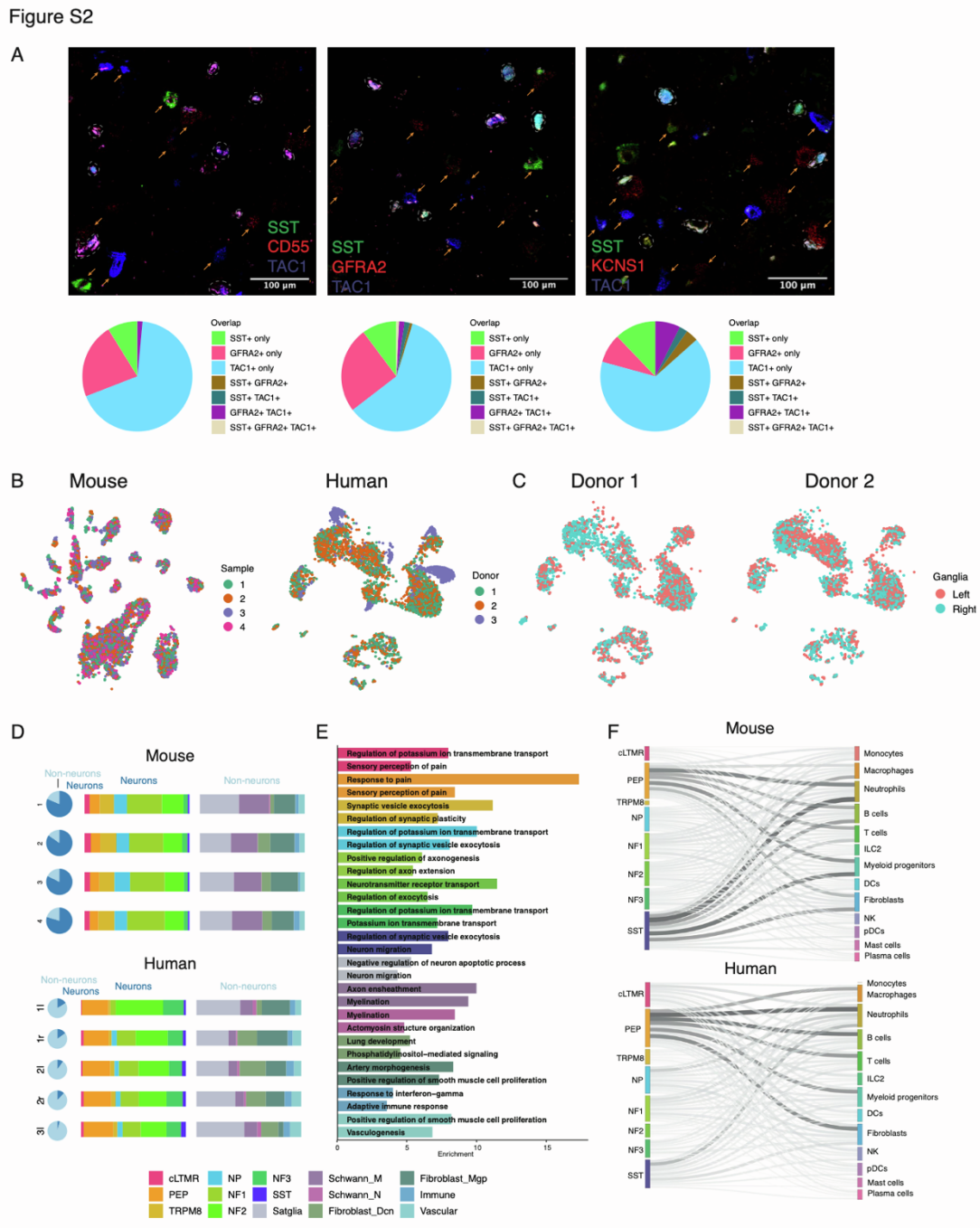

Figure S2. Conservation of cell-type-specific gene expression between human and mouse TG. Related to Figure 2.

A. Representative florescent in situ hybridization images of human TG (donor 3) stained with probes against cell-type-specific marker genes (top). Left) *SST* (SST, green), *CD55* (NP, red), and *TAC1* (PEP, blue); Middle) *SST* (SST, green), *GFRA2* (cLTMR, red), and *TAC1* (PEP, blue); Right) *SST* (SST, green), *KCNS1* (NF, red), and *TAC1* (PEP, blue). Scale bars = 100µm. Arrows point to individual cells, lipofuscin (signal present in all channels) is circled. Pie charts show the

quantification of fraction of cells that express combinations of marker genes (bottom, two sections are analyzed per probe set, average 76 cells are quantified per section).

cLTMR = C-fiber low threshold mechanoreceptor; PEP = peptidergic nociceptor; NP = non-peptidergic nociceptor; NF = *NEFH*+ A-fiber low threshold mechanoreceptors; SST = *SST*+ pruriceptors.

B. UMAP plots of snRNA-seq data from 7,164 mouse TG nuclei from male naive mouse replicates 1 - 4 downsampled to display 6,000 nuclei (left) or 38,028 human TG nuclei downsampled to display 6,000 nuclei (right). Colors indicate biological replicate in mice (left) or the human TG donor (right).

C. UMAP plots of snRNA-seq data from 38,028 human TG nuclei downsampled to display 2,000 nuclei from Donor 1 (left) and 2,000 nuclei from Donor 2 (right). Colors indicate whether nuclei were sequenced from the left or right TG of each donor. Bilateral snRNA-seq was only performed for Donors 1 and 2.

D. Fraction of neuronal and non-neuronal cell types and subtypes in male naïve mouse TG reference snRNA-seq (left) and human TG (right). Pie charts indicate fraction of neurons (dark blue) and non-neurons (light blue) in each sample. Horizontal bars indicate the fraction of neuronal or non-neuronal subtypes that comprise the respective cell type. Colors of bar chart indicate subtype.

E. Gene ontology analysis of cell-type-specific genes (Log2FC > 1, FDR < 0.05, cell type compared to all other TG cell types) identified in both mouse and human TG snRNA-seq. Significant ontology terms were defined as enrichment > 2 and FDR < 0.05 by comparing cell-type-specific genes to all expressed genes in the respective cell type. The two most enriched conserved ontology terms from each cell type are displayed.

F. Putative ligand interactions between mouse (top) and human (bottom) TG neurons and mouse meningeal cell types. Vertical bars are colored by cell type and length of bar depicts the number of ligand-receptor pairs in that cell type. The thickness of connecting lines between cell types is proportional to the number of total ligand-receptor interactions between the two cell types (see methods).

Figure S3

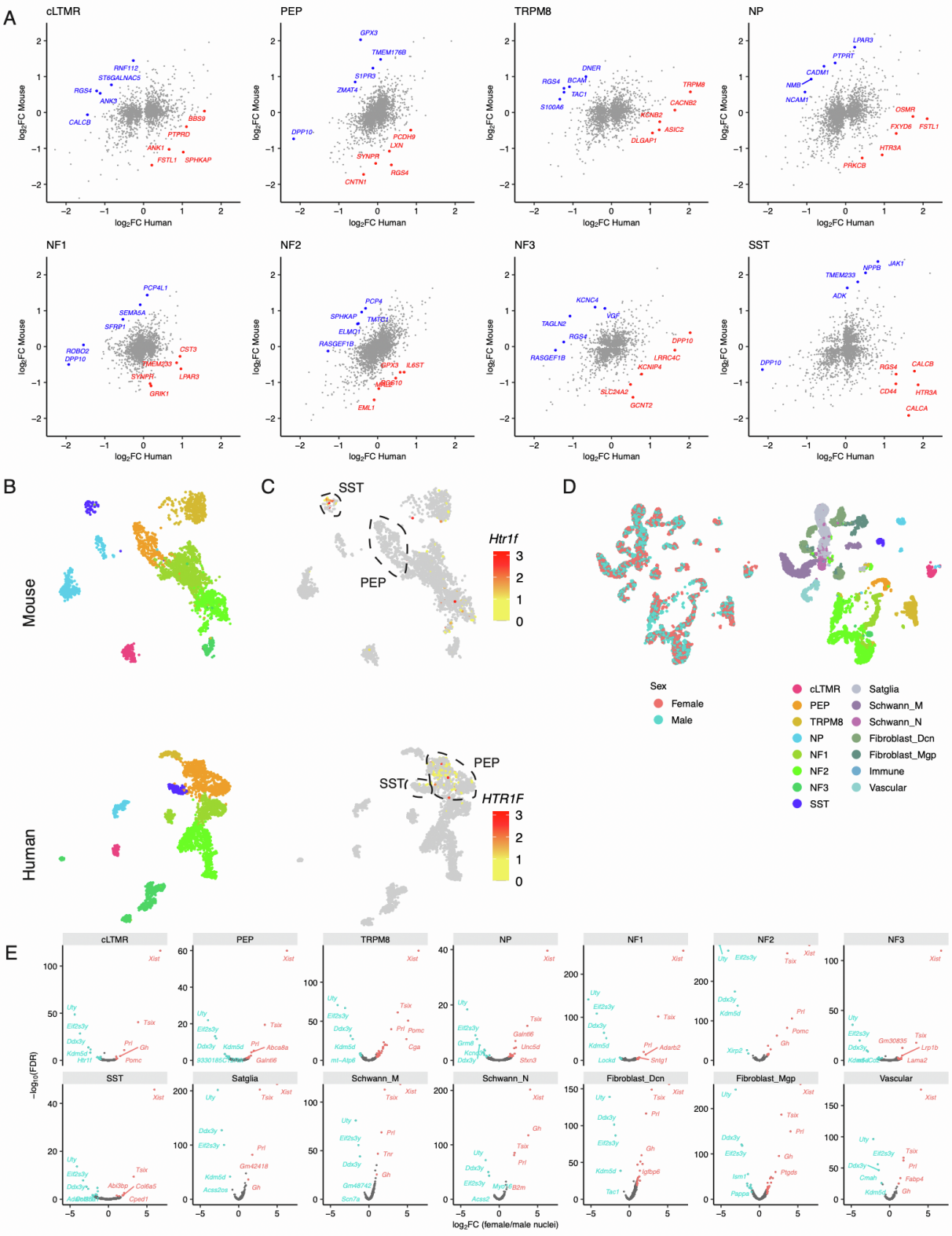
A
cLTMR
PEP
TRPM8
NP
NF1
NF2
NF3
SST
log2FC Mouse
log2FC Human
B
Mouse
Human
C
SST
PEP
Htr1f
HTR1F
D
Sex
Female
Male
cLTMR
PEP
TRPM8
NP
NF1
NF2
NF3
SST
Satglia
Schwann_M
Schwann_N
Fibroblast_Dcn
Fibroblast_Mgp
Immune
Vascular
E
−log10(FDR)
log2FC (female/male nuclei)

Figure S3. Species-specific and sex-specific features of TG cell types. Related to Figure 3.

A. Comparison of cell-type-specific gene expression patterns in mouse and human TG. For each cell type, scatter plots display the Log2FC of gene expression in each human TG cell type compared to all other human TG cell types (x-axis) and the Log2FC of its mouse orthologue expression in each mouse TG cell type compared to all other mouse TG cell types (y-axis). The five most differentially expressed human-specific genes (ordered by the difference between human Log2FC and mouse Log2FC) in each cell type are colored red and the five most differentially expressed mouse-specific genes (ordered by the difference between mouse Log2FC and human Log2FC) are colored blue. All other genes are grey.

B. UMAP plots of snRNA-seq data of 3,000 nuclei downsampled from 15,303 mouse TG neurons (top row) or 3,000 nuclei downsampled from 3,873 human (bottom row) TG neurons. Nuclei are colored by cell type.

C. UMAP plots of snRNA-seq data of 3,000 nuclei downsampled from 15,303 mouse TG neurons (top row) or 3,000 nuclei downsampled from 3,873 human (bottom row) TG neurons. Nuclei are colored by log2 expression of Htr1f/HTR1F. Peptidergic nociceptors (PEP) and *Sst*-expressing pruriceptors (SST) are circled.

D. UMAP plots of snRNA-seq data of 6,000 nuclei downsampled from 59,921 mouse naive TG nuclei colored by sex (Left) or cell type (Right).

E. Volcano plots of sex differences in TG cell-type-specific gene expression. Differential gene expression analysis was performed between females and males in each TG cell type. Each point is a gene, which are colored salmon if they are expressed significantly more in females than males, or cyan if they are expressed significantly more in males than females. Significance is defined as FDR < 0.05 and Log2FC >1.

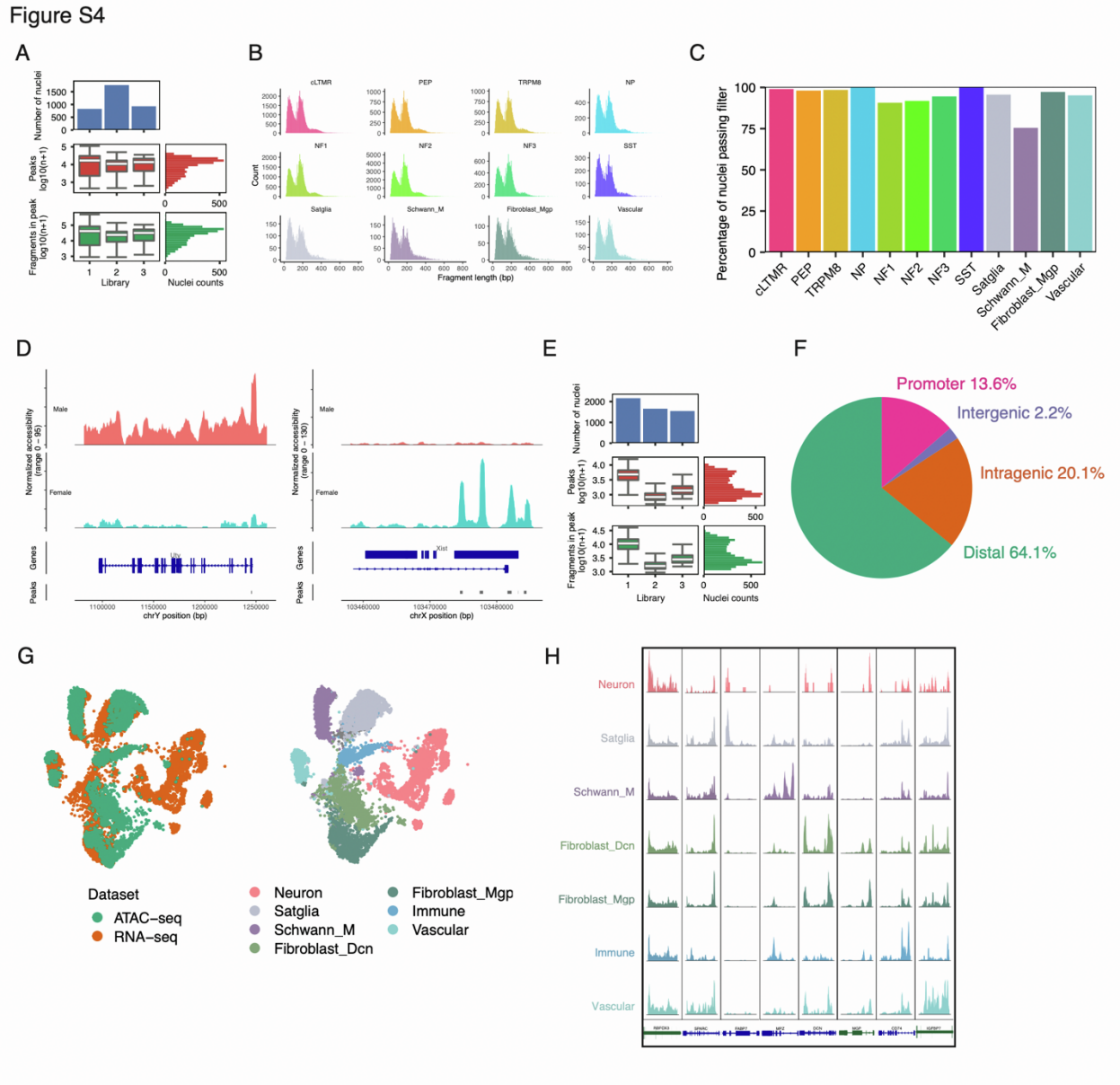

Figure S4. snATAC-seq of human and mouse TG. Related to Figure 5.

A. Mouse TG snATAC-seq metrics. Top row displays number of nuclei per biological replicate (n = 3) with > 600 fragments/nucleus. Middle row displays box plots of number of fragments per nucleus (log_{10} transformed) in each sample and the bottom row displays the number of fragments per peak (log_{10} transformed). Boxes indicate quartiles and whiskers are 1.5-times the interquartile range (Q1-Q3). The median is a white line inside each box. The distribution is aggregated across all samples and displayed on the horizontal histogram.

B. Distribution of mouse TG snATAC-seq fragment lengths in each cell type.

C. Percentage of each mouse TG cell type analyzed by snATAC-seq that passed quality control filters after doublet removal and received confident cell type assignments by anchoring (see methods).

D. Chromatin accessibility is displayed at the genomic loci of *Uty* or *Xist*, genes which are selectively expressed in males and females, respectively. At each genomic locus, chromatin accessibility in is displayed as the average fraction of transposase-sensitive fragments per nucleus at that region (moving average of 50 bins per genomic region). Accessibility at each locus (Y-axis) is scaled to the max value across all cell types (column).

E. Human TG snATAC-seq metrics. Top row displays number of nuclei with > 600 fragments. Middle row is a box plot of number of fragments per nucleus (log_{10} transformed) and the bottom row displays the number of fragments per peak (log_{10} transformed). Boxes indicate quartiles and whiskers are 1.5-times the interquartile range (Q1-Q3). The median is a white line inside each box. The distribution is displayed on the horizontal histogram.

F. Fraction of human TG snATAC-seq reads that map to promoter regions (-1,000bp to +100bp of transcription start site [TSS]), intragenic regions (within gene body excluding promoter region), distal regions (<200 kb upstream or downstream of TSS excluding promoter and intragenic regions), and intergenic regions (>200kb upstream or downstream of TSS) across 3 donor samples.

G. UMAP plots of human TG nuclei profiled by snATAC-seq data anchored to human TG nuclei profiled by snRNA-seq. Left) plot show 5,035 nuclei profiled by snATAC-seq with prediction score > 0.5 (green) and 38,028 reference nuclei from snRNA-seq data downsampled to display 7,000 nuclei (orange). Right) UMAP plot on left colored by cell type.

H. For each human TG cell type (rows), chromatin accessibility is displayed at cell-type-specific genes (columns). Human TG snATAC-seq data is displayed as the average frequency of sequenced DNA fragments per cell for each cell type, grouped by 50 bins per genomic region; Y-axis is scaled for each gene (column).

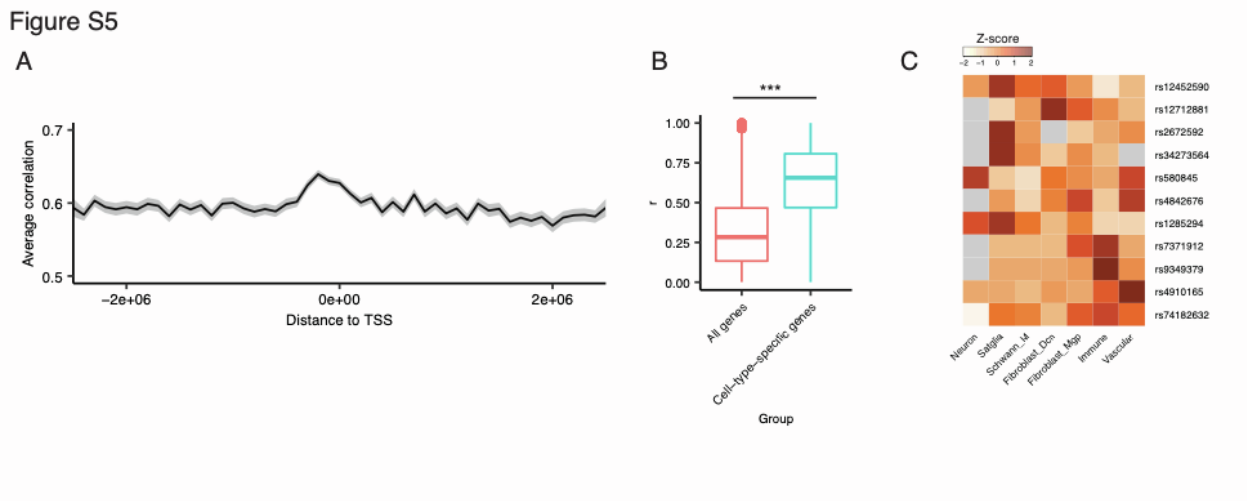

Figure S5. Cell-type-specific gene regulatory mechanisms in TG. Related to Figure 6.

A. Moving average (100 kb bins, black line) of the Pearson's correlation of 43,885 pairs of differentially accessible peaks (DA peaks, Log2FC > 0.5, FDR < 0.05, comparing accessibility in one TG cell type to all others) and cell-type-specific genes (Log2FC > 0.5 FDR < 0.05, comparing gene expression in one TG cell type to all others) on the same chromosome in the respective cell type as a function of distance to each gene's TSS. Standard error of the mean is shown.

B. Boxplot displays the Pearson's correlation for each cell-type-specific peak (Log2FC > 0.5, FDR < 0.05, comparing each cell type to all other cell types) and either cell-type-specific gene expression (Log2FC >0.5, FDR <0.05, comparing each cell type to all other cell types) on the same chromosome for the respective cell type or all genes on that chromosome. Boxes indicate quartiles and whiskers are 1.5-times the interquartile range (Q1-Q3). The median is a line inside each box. ****P* < 0.001, Student's t-test.

C. Heatmap displays Z-score (row scaled) of mean fraction of human TG snATAC-seq fragments in each cell type that map to a 1,000 bp window around each migraine-associated genomic locus.

Figure S6

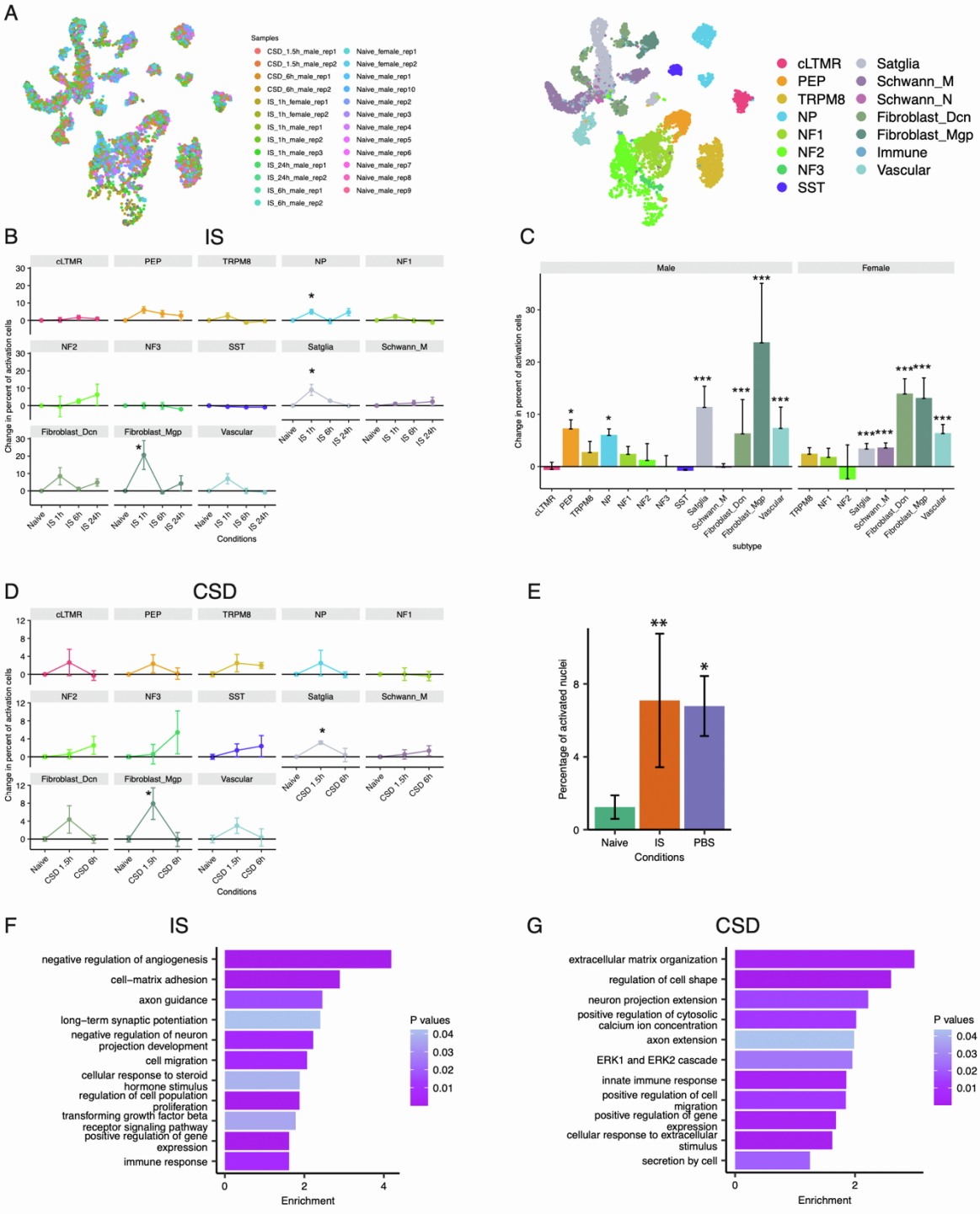
A
Samples
CSD_1.5h_male_rep1
CSD_1.5h_male_rep2
CSD_6h_male_rep1
CSD_6h_male_rep2
IS_1h_female_rep1
IS_1h_female_rep2
IS_1h_male_rep1
IS_1h_male_rep2
IS_1h_male_rep3
IS_24h_male_rep1
IS_24h_male_rep2
IS_6h_male_rep1
IS_6h_male_rep2
Naive_female_rep1
Naive_female_rep2
Naive_male_rep1
Naive_male_rep10
Naive_male_rep2
Naive_male_rep3
Naive_male_rep4
Naive_male_rep5
Naive_male_rep6
Naive_male_rep7
Naive_male_rep8
Naive_male_rep9
cLTMR
PEP
TRPM8
NP
NF1
NF2
NF3
SST
Satglia
Schwann_M
Schwann_N
Fibroblast_Dcn
Fibroblast_Mgp
Immune
Vascular
B
IS
Change in percent of activation cells
Conditions
Naive
IS 1h
IS 6h
IS 24h
C
Male
Female
Change in percent of activation cells
subtype
D
CSD
Change in percent of activation cells
Conditions
Naive
CSD 1.5h
CSD 6h
E
Percentage of activated nuclei
Naive
IS
PBS
Conditions
F
IS
negative regulation of angiogenesis
cell−matrix adhesion
axon guidance
long−term synaptic potentiation
negative regulation of neuron projection development
cell migration
cellular response to steroid hormone stimulus
regulation of cell population proliferation
transforming growth factor beta receptor signaling pathway
positive regulation of gene expression
immune response
P values
0.04
0.03
0.02
0.01
Enrichment
G
CSD
extracellular matrix organization
regulation of cell shape
neuron projection extension
positive regulation of cytosolic calcium ion concentration
axon extension
ERK1 and ERK2 cascade
innate immune response
positive regulation of cell migration
positive regulation of gene expression
cellular response to extracellular stimulus
secretion by cell
P values
0.04
0.03
0.02
0.01
Enrichment

Figure S6. snRNA-seq of TG in mouse models of headache. Related to Figure 7.

A. UMAP plots display snRNA-seq data of 96,933 mouse TG nuclei (59,921 naive, 37,012 IS/CSD) from each library used in headache models analyses. Each library is downsampled to display 200 nuclei. Nuclei are colored by library (left) or cell type (right).

B. Change in percentage of transcriptionally activated nuclei in each TG cell type after IS. Significant effects of IS are observed in NP, satellite glia, and fibroblasts (*P<0.05) and strong trends of IS activation are observed in PEP and vascular cells (P<0.1, one-way ANOVA: cLTMR: $F(3, 15) = 0.52$, $p = 0.67$; PEP: $F(3, 16) = 2.65$, $p = 0.08$, Tukey HSD post hoc testing: $p = 0.06$ between 1h after IS and naive; TRPM8: $F(3, 17) = 2.09$, $p = 0.14$; NP: $F(3, 16) = 4.92$, $p = 0.01$, Tukey HSD post hoc testing: $p = 0.045$ between 1h after IS and naive; NF1: $F(3, 13) = 1.76$, $p = 0.20$; NF2: $F(3, 15) = 0.56$, $p = 0.65$; NF3: $F(3, 13) = 0.56$, $p = 0.6529$; SST: $F(3, 16) = 1.39$, $p = 0.28$; Satglia: $F(3, 18) = 4.83$, $p = 0.01$, Tukey HSD post hoc testing: $p = 0.01$ between 1h after IS and naive; Schwann_M: $F(3, 18) = 0.61$, $p = 0.61$; Fibroblast_Dcn: $F(3, 18) = 1.57$, $p = 0.23$; Fibroblast_Mgp: $F(3, 18) = 3.62$, $p = 0.03$, Tukey HSD post hoc testing: $p = 0.043$ between 1h after IS and naive; Vascular: $F(3, 18) = 3.54$, $p = 0.04$, Tukey HSD post hoc testing: $p = 0.06$ between 1h after IS and naive).

C. Comparison of the percentage of transcriptionally activated nuclei in each TG cell type after IS between males and females. In males, significant increases in the fraction of activated nuclei ($P < 0.05$, permutation test) after IS compared to naïve are observed in PEP, NP, satellite glia, fibroblasts, and vascular cells. In females, significant increases in the fraction of activated nuclei ($P < 0.05$, permutation test) after IS compared to naive are observed in satellite glia, Schwann cells, fibroblasts, and vascular cells. Neuronal subtypes with < 50 cells are not shown. Note that the slight differences in P-values between this figure and Figures S6B are secondary to splitting comparisons by sex and by the use of permutation test (this figure) vs. ANOVA (Figure S6B).

D. Change in percentage of transcriptionally activated nuclei in each TG cell type after CSD. Significant effects of CSD were observed in satellite glia and fibroblasts (*P<0.05, one-way ANOVA: cLTMR: $F(2, 11) = 0.99$, $p = 0.40$; PEP: $F(2, 11) = 1.18$, $p = 0.34$; TRPM8: $F(2, 12) = 1.81$, $p = 0.21$; NP: $F(2, 11) = 0.98$, $p = 0.41$; NF1: $F(2, 9) = 0.06$, $p = 0.95$; NF2: $F(2, 11) = 1.36$, $p = 0.30$; NF3: $F(2, 9) = 0.96$, $p = 0.42$; SST: $F(2, 11) = 0.77$, $p = 0.49$; Satglia: $F(2, 12) = 5.55$, $p = 0.02$, Tukey HSD post hoc testing: $p = 0.02$ between 1.5h after CSD and naive; Schwann_M: $F(2, 12) = 0.93$, $p = 0.42$; Fibroblast_Dcn: $F(2, 12) = 2.56$, $p = 0.12$; Fibroblast_Mgp: $F(2, 12) = 5.57$, $p = 0.02$, Tukey HSD post hoc testing: $p = 0.02$ between 1.5h after CSD and naive; Vascular: $F(2, 12) = 1.32$, $p = 0.30$).

E. Fraction of nuclei in each headache model that display a transcriptionally activated naïve in naive mice, 1h after PBS, or 1h after IS as defined by the expression of intermediate early genes (IEG) (expression defined as 2 standard deviations above the mean IEG score across all nuclei of the same cell type, see methods). There are significantly more transcriptionally activated cells 1h after PBS and IS than in control samples (one-way ANOVA: $F (2, 13) = 10.02$, $p = 0.002$; Tukey HSD post hoc testing: $p < 0.01$ between IS and naive, $p < 0.05$ between PBS and naive).

F. Gene ontology analysis of genes that are differentially regulated in IS-activated cells compared to control cells (Log2FC > 1, FDR < 0.05, from Figure 7D). Significant ontology terms are defined as (Enrichment > 2, FDR < 0.05, IS-activated genes compared to all expressed genes in cells included in the analysis).

G. Gene ontology analysis of genes that are differentially regulated in CSD-activated cells compared to control cells (Log2FC > 1, FDR < 0.05, from Figure 7D). Significant ontology terms are defined as (Enrichment > 1.5, FDR < 0.05, CSD-activated genes compared to all expressed genes in cells included in the analysis).

Figure S7

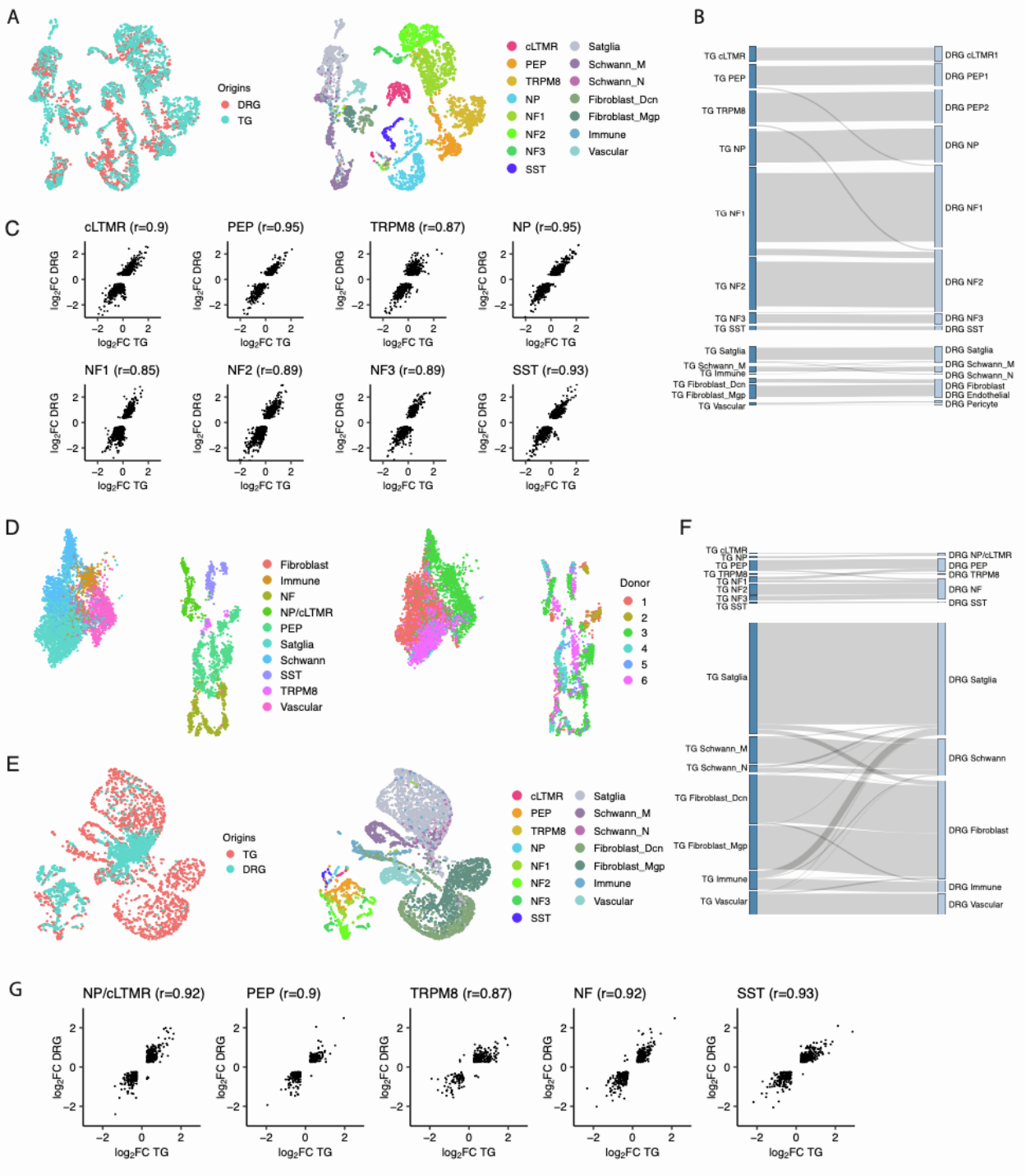
A
Origins
DRG
TG
cLTMR
PEP
TRPM8
NP
NF1
NF2
NF3
SST
Satglia
Schwann_M
Schwann_N
Fibroblast_Dcn
Fibroblast_Mgp
Immune
Vascular
B
TG cLTMR
TG PEP
TG TRPM8
TG NP
TG NF1
TG NF2
TG NF3
TG SST
TG Satglia
TG Schwann_M
TG Immune
TG Fibroblast_Dcn
TG Fibroblast_Mgp
TG Vascular
DRG cLTMR1
DRG PEP1
DRG PEP2
DRG NP
DRG NF1
DRG NF2
DRG NF3
DRG SST
DRG Satglia
DRG Schwann_M
DRG Schwann_N
DRG Fibroblast
DRG Endothelial
DRG Pericyte
C
cLTMR (r=0.9)
PEP (r=0.95)
TRPM8 (r=0.87)
NP (r=0.95)
NF1 (r=0.85)
NF2 (r=0.89)
NF3 (r=0.89)
SST (r=0.93)
log2FC DRG
log2FC TG
D
Fibroblast
Immune
NF
NP/cLTMR
PEP
Satglia
Schwann
SST
TRPM8
Vascular
Donor
1
2
3
4
5
6
E
Origins
TG
DRG
cLTMR
PEP
TRPM8
NP
NF1
NF2
NF3
SST
Satglia
Schwann_M
Schwann_N
Fibroblast_Dcn
Fibroblast_Mgp
Immune
Vascular
F
TG cLTMR
TG NP
TG PEP
TG TRPM8
TG NF1
TG NF2
TG NF3
TG SST
TG Satglia
TG Schwann_M
TG Schwann_N
TG Fibroblast_Dcn
TG Fibroblast_Mgp
TG Immune
TG Vascular
DRG NP/cLTMR
DRG PEP
DRG TRPM8
DRG NF
DRG SST
DRG Satglia
DRG Schwann
DRG Fibroblast
DRG Immune
DRG Vascular
G
NP/cLTMR (r=0.92)
PEP (r=0.9)
TRPM8 (r=0.87)
NF (r=0.92)
SST (r=0.93)
log2FC DRG
log2FC TG

Figure S7. Comparison between snRNA-seq of TG and DRG in human and mouse. Related to Figure 1.

A. UMAP plots of 59,921 naive mouse TG nuclei profiled by snRNA-seq data in this study anchored to 16,828 male naïve mouse DRG nuclei profiled by snRNA-seq (Renthal et al. 2020). UMAP plots are colored by Left) ganglion type; Right) TG cell type.

B. Similarities between cell types of mouse TG and mouse DRG. Gray lines connect the TG classifications generated in this study and the DRG cell type to which they are most closely related when anchored to mouse DRG (Renthal et al. 2020). Plots show nuclei with anchoring prediction score > 0.5.

C. Comparison of cell-type-specific gene expression patterns in mouse DRG and TG. For each cell type, scatter plots display the Log2FC of gene expression in each mouse TG cell type compared to all other TG cell types (Log2FC > 1, FDR < 0.05, x-axis) and the Log2FC of the same gene in each mouse DRG cell type compared to all other DRG cell types (Log2FC > 1, FDR < 0.05, y-axis). Pearson's correlation between cell-type-specific gene expression in TG and DRG is displayed for each cell type.

D. UMAP plots of 6,764 human DRG nuclei profiled by snRNA-seq (Nguyen et al. 2021). Left) UMAP plots are colored by cell type. Right) UMAP plots are colored by donor.

E. UMAP plots of 38,028 human TG nuclei profiled by snRNA-seq data in this study anchored to 6,764 human DRG nuclei profiled by snRNA-seq (Nguyen et al. 2021). Left) UMAP plots are colored by cell type. Right) UMAP plots are colored by TG ganglion type.

F. Similarities between cell types of human TG and human DRG. Gray lines connect the TG classifications generated in this study and the DRG cell type to which they are most closely related when anchored to human DRG (Nguyen et al. 2021). Plots show nuclei with anchoring prediction score > 0.5.

G. Comparison of cell-type-specific gene expression patterns in human DRG and TG. For each cell type, scatter plots display the Log2FC of gene expression in each human TG cell type compared to all other TG cell types (Log2FC > 1, FDR < 0.05, x-axis) and the Log2FC of the same gene in each human DRG cell type compared to all other DRG cell types (Log2FC > 1, FDR < 0.05, y-axis). Pearson's correlation between cell-type-specific gene expression in TG and DRG is displayed for each cell type.