bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Single-Soma Deep RNA sequencing of Human DRG Neurons Reveals Novel Molecular and Cellular Mechanisms Underlying Somatosensation

Huasheng Yu[1], Dmitry Usoskin[2*], Saad S. Nagi[3*], Yizhou Hu[2], Jussi Kupari[2], Otmane Bouchatta[3], Suna Li Cranfill[1], Yijing Su[1], You Lv[1], Hongjun Song[1], Guo-Li Ming[1], Stephen Prouty[5], John Seykora[5], Hao Wu[4], Minghong Ma[1], Håkan Olausson[3#], Patrik Ernfors[2#], Wenqin Luo[1#]

1. Department of Neuroscience, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA.

2. Department of Medical Biochemistry and Biophysics, Division of Molecular Neurobiology, Karolinska Institute, Stockholm 17177, Sweden.

3. Department of Biomedical and Clinical Sciences, Linköping University, Linköping 58185, Sweden.

4. Departments of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA.

5. Department of Dermatology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA.

* These authors contributed equally to this work

Correspondence:

hakan.olausson@liu.se (Håkan Olausson)

patrik.ernfors@ki.se (Patrik Ernfors)

luow@pennmedicine.upenn.edu (Wenqin Luo)

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract:

The versatility of somatosensation arises from heterogenous human dorsal root ganglion (DRG) neurons. The critical information to decipher their functions, i.e., the soma transcriptome, is lacking due to technical difficulties. Here we developed a novel approach to isolate individual human DRG neuron somas for deep RNA sequencing (RNA-seq). On average, >9000 unique genes per neuron were detected, and 16 neuronal types were identified. Cross-species analyses revealed that touch-, cold-, and itch-sensing neuronal types were relatively conserved, while the pain-sensing neurons displayed marked divergence. Soma transcriptomes of human DRG neurons predicted novel functional features, which were confirmed using single-cell *in vivo* electrophysiological recordings. These results support a close relationship the between physiological properties of human sensory afferents and molecular profiles uncovered by the single-soma RNA-seq dataset. In summary, by conducting single-soma RNA-seq of human DRG neurons, we generated an unprecedented neural atlas for human somatosensation.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction:

The somatosensory system transduces diverse stimuli from external and internal environments to generate sensations from the body, such as pain, itch, temperature, touch, vibration, and body position[1]. Primary somatosensory neurons, which convert stimuli to electrical signals, are located in the dorsal root ganglia (DRG) and trigeminal ganglia (TG)[2]. They are greatly heterogeneous, composed of many different molecularly and functionally distinct populations[3]. Physiological recordings of human primary somatosensory afferents classify them into several broad types, based on axon conduction velocities, namely Aα, Aβ, Aδ, and C-fibers[4]. Each type is further divided based on its morphology, molecular markers, and physiological features[5]. Normal functions of the somatosensory system are fundamental to our daily lives, but abnormal somatosensation in pathological conditions, such as chronic pain and itch, are devastating human health problems. Safe and effective drug options for chronic pain and itch are still limited[6-8], and the development of novel treatment strategies is urgently needed.

Most of our current knowledge about the mammalian somatosensory system comes from model organisms, mainly rodents. However, the success rate of translating treatments from model organisms to humans is low[9,10]. There are many reasons for this, one of which is the species differences of somatosensory neurons. Studies from the somatosensory field have noticed significant species differences between rodent and human DRG neurons. For example, our group identified MRGPRX4 as a bile acid receptor for human cholestatic itch, but MRGPRX4 does not have a molecular ortholog in mice[11]. Intriguingly, TGR5, a receptor identified in mouse for bile acid-induced itch[12], is not expressed in human DRG neurons but in the surrounding satellite glial cells instead[11] Thus, it is critical to elucidate the molecular profiles and cell types of human DRG neurons for understanding human somatosensory mechanisms as well as for translational approaches.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Single-cell RNA sequencing (RNA-seq) is a powerful tool to qualitatively and quantitatively study transcripts of individual cells (soma and/or nuclei)[13]. Based on the transcriptome, including both transcript identities and expression levels, heterogenous cells can be classified into different types[14]. This approach has been successfully used to study mouse DRG and TG neurons in adults and throughout development, providing comprehensive molecular and cellular atlases for understanding mouse somatosensory neurons and their differentiation[15-17]. This approach was also successfully used to conduct single-cell RNA-seq of macaque DRG neurons[18]. However, human DRG and TG pose some unique technical challenges. Compared to model organisms, human tissue is more difficult to obtain. In addition, in human DRGs/TGs, non-neuronal cells, such as satellite glia, fibroblasts, and other cell types outnumber the neuronal cells[19,20]. Further, satellite glial cells tightly wrap around neuronal somas[19,20], making their physical separation difficult. Moreover, human DRG/TG neuronal somas are much larger than those of non-neuronal cells, so they are prone to damage by enzyme digestion and mechanical forces during single-cell isolation. Due to these difficulties, single-nucleus RNA-seq and spatial transcriptomics have been employed for human DRG/TG neurons[21-23]. Despite novel insights from these studies, the quantity of transcripts in the nucleus is much lower than those in the soma, and the nuclear transcripts may not represent the full transcriptome profile of the whole cell[24], while commercially available spatial transcriptomics lack single-neuron resolution.

To elucidate the critical information for deciphering versatile functions of human DRG neurons, i.e., their soma transcriptomes, we developed a novel approach to perform single-soma deep RNA-seq of human DRG neurons by combining laser capture microdissection (LCM) of individual neuronal soma and the Smart-seq2 protocol[18]. We successfully sequenced 1066 human DRG neurons with minimum satellite glia cell contamination from six lower thoracic and

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

lumbar DRGs of three donors, detecting an average of >9000 unique genes per neuron and identifying 16 molecularly distinctive neuron types. Cross-species analysis revealed interesting differences and similarities among human, macaque, and mouse DRG neurons. In addition, we uncovered a series of novel marker genes that can identify different types of human DRG neurons. Based on the molecular profiles, we also predicted novel physiological properties of human DRG neuron types, which were tested and confirmed by single-cell *in vivo* recordings from human sensory afferents. These results support a close relationship between functional properties of human sensory afferents and molecular profiles uncovered by the single-soma RNA-seq dataset, highlighting the precision and unique utility of this dataset in functional discoveries. In short, we have established a novel approach for conducting single-soma deep RNA-seq of adult human neurons, which revealed previously unknown human DRG neuron cell types and functional properties. Our results generated an unprecedent atlas of the neural basis for human somatosensation, providing an important reference for translating animal studies into therapeutic applications.

Results

Development of an LCM-based approach for conducting single-soma deep RNA-seq of human DRG neurons

Human DRG neurons have large soma sizes (~20-120 μm diameter), are relatively few in number, and are prone to damages caused by enzymatic digestion and mechanical dissociation[20]. As a result, sing-cell RNA-seq studies using the traditional enzymatic and mechanical dissociation method retain very few intact human DRG neurons[19,20]. In addition, enzymatic and mechanical dissociation causes transcriptome changes[25]. To overcome these hurdles, we developed a novel method that combined immediate fresh frozen human DRG tissues after postmortem extraction, cryosectioning, laser capture microdissection (LCM), and

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Smart-seq2 deep sequencing[18] (Fig. 1A). Six DRGs at the low thoracic (T11-T12) and lumbar (L2-L5) levels from three donors (one male and two females, age range: 23-61 years old, table 1) were procured through NDRI (National Disease Research Interchange, USA). Fresh frozen DRGs were cryo-sectioned, mounted onto LCM slides, and briefly stained with a HistoGene™ dye for visualization of neuronal somas (Fig. S1A-C). Individual neuronal somas were discernable under microscope and dissected by a laser. Each detached soma dropped into a tube cap for library preparation (Fig. 1A). Dissected neuronal somas exhibited a similar size distribution as the entire DRG neuron population (Fig. S1D & S1E), suggesting no obvious sampling bias. Sequencing libraries were generated following the Smart-seq2[26] protocol. A total of 1136 individual neuronal somas were dissected and passed through the final quality control for sequencing. During preliminary bioinformatic analysis, 70 cells demonstrating obvious glial cell contamination (dominant expression of APEO, FABP7, and other glia cell markers) were removed, and the remaining 1066 neuron were used for further analysis. 16 transcriptomic clusters of human DRG neurons were identified by Seurat[27] (Fig. 1B) with an average of 9486 genes detected per cell (Fig. 1C). No obvious batch effects or donor differences were observed in the clusters (Fig. S2A-E). As expected, all these cells expressed high levels of neuronal markers, *SLC17A6* (*VGLUT2*), *SYP* (Synaptophysin), and *UCHL1* (*PGP9.5*) (Fig. 1D, S2F & S2G).

Anatomical and molecular features of the clustered cell types

Based on the soma size distribution and the expression levels of neurofilament heavy chain (*NEFH*) (Fig. 1E & 1F), the 16 clusters contained neurons corresponding to the unmyelinated, small diameter C-fiber afferents (cluster 1-7) and the myelinated, large diameter A-fiber afferents (cluster 8-16). The two groups could be further divided based on the expression of

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

calcitonin related polypeptide alpha (*CALCA (CGRP)*), a peptidergic neuron marker, and PR domain zinc finger protein 12 (*PRDM12*), a transcriptional regulator critical for human pain perception[28] (Fig. 1F). Taken all information into consideration, clusters 1-4 were classified as non-peptidergic C-fiber neurons, clusters 5-7 as peptidergic C-fiber neurons, clusters 8-12 as peptidergic A-fiber neurons, clusters 13-15 as low-threshold mechanoreceptors A-fiber (A-LTMRs) neurons, and cluster 16 as an unknown A-fiber neuron type. Based on the unique markers and functional genes expressed in each cluster (Fig. 1G, S2H & S3), we named the human DRG neuron clusters (Fig. 1H) using a nomenclature system according to the following rules: 1) The "h" at the beginning of each cluster name indicated "human"; 2) mouse nomenclature for conserved subtypes (i.e. most non-peptidergic fibers and A-LTMRs); 3) human peptidergic neuron types were named as hPEP.(marker gene). Briefly, for non-peptidergic C-fiber neurons, Cluster 1 was named hTRPM8, Cluster 2 was C-fiber low-threshold mechanoreceptors (hC.LTMR), Cluster 3 was type I non-peptidergic nociceptors (hNP1), and Cluster 4 was type II non-peptidergic nociceptor (hNP2). For peptidergic C-fiber neurons, Cluster 5 was named hPEP.SST, Cluster 6 hPEP.TRPV1/A1.1, and Cluster 7 hPEP.TRPV1/A1.2. For peptidergic A-fiber neurons, Cluster 8 was named hPEP.PIEZO^{h} (the superscript 'h' means 'high'), Cluster 9 hPEP.KIT, Cluster 10 hPEP.CHRNA7, Cluster 11 hPEP.NTRK3, and Cluster 12 hPEP.0 (no distinctive molecular marker). For A-LTMRs, Cluster 13 was named Aδ low-threshold mechanoreceptors (hAδ.LTMR), Cluster 14 Aβ low-threshold mechanoreceptors (hAβ.LTMR), and Cluster 15 proprioceptors (hPropr). Cluster 16 was names as hATF3. All clusters had specific distinctive molecular markers, except hPEP.0 (Fig. S3).

In addition, Conos[29] was used as an independent method for clustering. Cluster structure revealed by Conos analysis reproduced the Seurat results (Fig. S4A). We also employed a neural network-based probabilistic scoring module[15,18] that learned human cell type features

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

based on their molecular profiles (Fig. S4B-F). The accuracy score of our Seurat clustering assignment by the learning module was near 90% (Fig. S4B), i.e. most cells were accurately assigned to their corresponding clusters (Fig. S4C & S4D). Moreover, the cell type consistency was validated by cell type purity and probabilistic similarity (Fig. S4E & S4F). In short, different analyses methods confirmed the robustness of the clustering and strongly supported the cell type assignment. We also used Conos co-clustering on integrated data from our human dataset and a recent human single-nucleus RNA-seq dataset by Nguyen et al[23] (Fig. S5). This analysis showed that while some clusters displayed a one-to-one match between the two datasets, some did not have a mutual match. This mismatch could be caused by biological variations such as nucleus vs cytoplasm RNA species and quantity[24] and technical differences (the increased resolution obtained by deep sequencing in the present study or variability caused by the different technology platforms [26,30]).

Cross-species comparison of DRG neuron types

To investigate the evolutional conservation and divergence of DRG neurons, we performed a cross-species comparison between our human dataset, a mouse single-cell RNA-seq datasets (Sharma)[16], and a macaque dataset (Kupari, SmartSeq2 dataset)[18]. To identify corresponding clusters between human and previously published mouse and macaque datasets, we used three different strategies: Conos pairwise co-clustering, probabilistic neural network learning, and machine-learning based hierarchical clustering of an integrated dataset of human and macaque. Conos was developed to identify recurrent cell clusters across different RNA-seq platforms and species, and propagation of information between the datasets. Probabilistic neural network analysis was based on training with one dataset and then testing others for pattern recognition and probability output[18]. In hierarchical clustering analysis, each query neurons, either human or macaque, was assigned weights of the sensory-type associated

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

patterns by a neural network, which was trained with gene patterns reflecting species specific and shared cross-species features in the different sensory neuron types. The weighted gene patterns were then used for dimensional reduction and nearest neighbor analysis to infer the hierarchical relationship. In all these analyses, some human DRG neuron cell types showed high correlation to those of mouse and macaque, including hAδ.LTMR, hAβ.LTMR, hProprioceptor, hC.LTMR, NP1, NP2, hPEP.SST (NP3 neurons) and TRPM8 neurons (Fig. 2, & S6). hPEP.TRPV1/A1.1 and hPEP.TRPV1/A1.2 corresponded to PEP1 (Usoskin nomenclature), specifically to subclass PEP1.4/CGRP-γ. This conclusion was supported by both Conos label propagation analysis (Fig. 2A) and probabilistic neural network analysis (Fig. S6A), suggesting these clusters to represent C-fiber thermoreceptors and nociceptors. Notably, four types of mouse C-fiber nociceptors have been evidenced[18,31] (nomenclature from Emory & Ernfors/Sharma), but our analysis indicated that mouse clusters PEP1.1/CGRP-α, PEP1.2/CGRP-β and PEP1.3/CGRP-ε are not evolutionarily conserved, as they were absent in the human dataset. hPEP.CHRNA7 showed high correlation to mouse PEP2/CGRP-ζ and macaque PEP2 (Fig. 2A-C, Fig.S6A, B), while hPEP.KIT corresponded to mouse PEP3/CGRP-η and macaque PEP3, suggesting these clusters functionally belong to A-fiber nociceptors. Interestingly, hPEP.PIEZO^{h}, hPEP.NTRK3, and hPEP.0 did not have any direct corresponding cell types in mouse and macaque. Nevertheless, hPEP.PIEZO^{h} showed some similarity to mouse PEP3 (CGRP-η) and macaque PEP1 and PEP3 in co-clustering (Fig. 2A-B & S6C-F), neural network learning and hierarchical clustering, indicating this cell type to represent an A-fiber mechanosensory nociceptor (Fig. 2A-C, S6A-B). hPEP.NTRK3 showed the greatest divergence by Conos propagation analysis, as it was most similar to mouse PEP2 (CGRP-ζ) but macaque PEP3 (Fig. 2A-B). Probabilistic neural network learning revealed similar scores to both mouse and macaque PEP2 (CGRP-ζ) and PEP3 (CGRP-η) (Fig. S6A & S6B), while hierarchical clustering revealed hPEP.NTRK3 to represent a divergent macaque PEP3-like neuron type (Fig.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

2C). Thus, hPEP.NTRK3 and hPEP.PIEZO^{h} share molecular features with both mouse and macaque PEP2 and PEP3 and likely emerge as human specific sensory neuron types. hPEP.0, a type of human PEP A-afferents, showed no similarity to any mouse DRG neurons but some relationship to macaque PEP3 and PEP2 (Fig. 2A-C).

We next performed a transcription factor associated gene regulatory network analysis (TF-GRNs) using machine learning to identify shared and species-specific TF-GRNs contributing to the similarities and differences between different sensory neuron subtypes and species. Evolutionarily conserved TF-GRNs defining C-fiber nociceptors, A-fiber nociceptors, A-LTMRs, TRPM8, C-fiber pruriceptors/nociceptors (hNP1, hNP2, hPEP.SST), and C-LTMRs were observed (yellow boxes in Fig. 2D) as well as species specific mouse- and or primate-specific networks, such as for C-fiber nociceptors, hTRPM8, hNP1, hNP2, hPEP.SST and hC.LTMR (green box in Fig 2D). Among cross-species conserved transcriptional regulators, some were previously known to drive sensory neuron diversification in mouse, including *ZEB2* in C-fiber nociceptors[32], *SHOX2* in A-LTMRs[33,34], *RUNX3* in proprioceptors[35], *FOXP2* in TRPM8, *RUNX1* in NP1[36], *ZNF52* and *POU4F2* in C-LTMRs[16,37].

Similarities and differences of top marker genes across species

Although cell type homologs between the three different species could be identified for many DRG neuron types, we observed interesting molecular differences between the corresponding cell types in human, macaque (Kupari)[18], and mouse (Sharma)[16]. We selected the top 10 marker genes from each human type, and the expression of these genes was mapped in different types of macaque and mouse DRG neurons (Fig. S7). In general, the expression patterns of these genes were more similar between human and macaque than between human

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and mouse, reflecting the evolutionary distances between rodent, non-human primate, and human. Some genes were expressed in the corresponding populations of DRG neurons across all three species. For example, *TRPM8* was expressed in C-fiber cold-sensing neurons, and *IL31RA* was expressed in non-peptidergic population (NP3) and hPEP.SST. Some genes, such as *CHRNA7*, a nicotinic acetylcholine receptor, were specific for primate but low in the corresponding mouse DRG neurons (Fig. S7). Moreover, some marker genes were specific for human DRG neuron types. For example, Mechanosensory Transduction Mediator Homolog (*STUM*) was specifically expressed in h.TRPM8, and calsequestrin 2 (*CASQ2*) was specifically expressed in hC.LTMR. These genes are not specific for or expressed in macaque and mouse corresponding neuron types (Fig. S7). Genes specifically enriched in human DRG neurons might play unique functions in human physiological or pathological somatosensation.

Molecular marker expressed in C fiber pruriceptors/nociceptors

Based on the sequencing results, we established specific marker genes or their combination to uniquely identify the different types of human DRG neurons. We validated expression of these specific markers using multiplex fluorescent *in situ* hybridization (FISH, RNASCOPE). hNP1, hNP2 and hPEP.SST were potential pruriceptors and nociceptors based on the functional gene expression (Fig. 3A). *MRGPRX1*, a Mas-related GPCR family member that senses various pruritogens[38], was exclusively expressed in hNP1 and hNP2 populations (Fig. 3B1). Several other itch-related receptors, such as *HRH1* and *IL31RA* (Fig. 3C1-C2 & 3F1-F2), were also expressed in hNP1 and hNP2, suggesting that these two populations function to detect various pruritogens and transmit itch signals. In mice, NP1 and NP2 afferents are well-characterized by specific expression of two different Mrgpr members: NP1 neurons (~20% of total DRG neurons) express *Mrgprd*[39] while NP2 neurons (~5% of total DRG neurons) express *MrgprA3*[15]. In human, however, though still expressed in NP1 afferents, the expression level of *MRGPRD* was

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

considerably lower (Fig. S8A), whereas *MRGPRA3* gene does not have a human orthologue in the genome. Expression of other MRGPR members helped to separate the two clusters, such as enrichment of *MRGPRX4* in hNP2 but not hNP1 (Fig. S8B). Similar to mouse, we found more hNP1 than hNP2 neurons in our dataset (Fig. 1H), and *PIEZO2* was expressed at a higher level in hNP1 than in hNP2 neurons ((Fig. S9A1 & 3B1), suggesting that hNP1 neurons were sensitive to mechanical force. Consistent with this molecular feature, recordings from human afferents have found that some human histaminergic itch-sensing fibers are mechanosensitive[40]. In short, despite the obvious differences in top molecular markers of NP1 and NP2 populations from mice to humans, these two populations may have relatively conserved physiological functions.

The hPEP.SST population displayed highly specific expressions of somatostatin (*SST*), cholecystokinin (*CCK*), and an enriched level of *GFRA3* (Fig. 3A, S8C). This cluster corresponded to mouse and macaque NP3 population (Fig. 2), which are also marked by the expression of *SST*[15,16]. Here we classified this cluster as peptidergic rather than non-peptidergic because of the co-expression of a classic peptidergic marker, *CALCA* (Fig. 1E, 3D1-D2), which is barely detected in the corresponding mouse neurons (Fig. S9A2) [16,18]. Given the previously established role of mouse NP3 neurons in itch sensation[41] and the expression of itch-sensing receptor, such as *HRH1*, *IL31RA* (Fig. 3E1-F2), and *CYSLTR2* (Fig. S12A), hPEP.SST afferents could also function in itch sensation, especially in inflammatory conditions.. Neither *PIEZO1* (Fig. S8D) nor *PIEZO2* (Fig. 3B1) was detected in hPEP.SST neurons, indicating that these afferents may not be mechanosensitive. Indeed, some human histaminergic itch-sensing fibers are insensitive to mechanical forces[40].

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Molecular markers expressed in C-fiber thermoreceptors and peptidergic nociceptors

The hTRPM8 population was distinguished from other cell types by their specific expression of *STUM* and high-level expression of *TRPM8* (Fig. 4A). Almost all *STUM*$^+$ neurons expressed *TRPM8* (Fig. 4B1 & 4B2). Since *TRPM8* is a receptor for cold temperature and cooling chemical agents (such as menthol)[42-45], the hTRPM8 population likely functions as cold- and menthol-sensing afferents. Notably, the cluster-specific marker *STUM* was not detected in mouse TRPM8 neurons from the Sharma and Zeisel datasets[16,31]. In the macaque (Kupari) dataset[18], *STUM* was more broadly expressed in several clusters (Fig. S9B1). Nevertheless, some molecular markers, such as *FOXP2* and *GPR26*, were shared among mouse[16,31], macaque[18] (Fig. S9A3, S9A4, S9B2 & S9B3), and human TRPM8 cold-sensing neurons (Fig. S8E & S8F). Intriguingly, some hTRPM8 neurons also expressed the heat-sensing receptor *TRPV1* (Fig. 4C1), suggesting that these neurons might also be activated by heat stimuli. Consistent with this molecular pattern, human physiological recordings have identified cold-sensitive C-fibers that also respond to heat[46]. Thus, the hTRPM8 population is likely to be polymodal, functioning in temperature (both cold and heat) and chemical sensation.

Two *CGRP*$^+$ peptidergic C clusters were featured by the overlapping high expression of *TRPV1* and *TRPA1* (Fig. 4C1-D2). We therefore named them hPEP.TRPV1/A1.1 and hPEP.TRPV1/A1.2. Since *TRPV1* is activated by noxious heat and capsaicin, and *TRPA1* can sense noxious cold and various noxious chemicals[47,48], they are likely C-fiber peptidergic thermoreceptors and nociceptors, sensing noxious thermal and chemical stimuli and transmitting pain signals. hPEP.TRPV1/A1.1 lacked expression of *PTGER3* and *PIEZO2*, whereas hPEP.TRPV1/A1.2 was PTGER3$^+$ and PROKR2$^+$, two molecular markers found in deep tissue innervating neurons in mice[49], and expressed a low level of *PIEZO2* (Fig. 3B1, 4D1 & S8G). In addition, prostaglandin I2 receptor (*PTGIR*) (Fig. S8H), a molecular marker found in

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

viscera-innervating mouse DRG neurons[50], was more prominently expressed in hPEP.TRPV1/A1.2 than hPEP.TRPV1/A1.1. This expression pattern indicates that hPEP.TRPV1/A1.1 and hPEP.TRPV1/A1.2 might innervate different peripheral targets, with hPEP.TRPV1/A1.1 serving as cutaneous thermal and chemical nociceptive afferents and hPEP.TRPV1/A1.2 serving as viscera- and/or deep tissue-innervating chemical nociceptive afferents.

Molecular markers expressed in A-fiber peptidergic nociceptors

Five clusters of peptidergic A-fibers nociceptors were identified. They had relatively large soma sizes and high expression levels of *NEFH*, *CALCA*, *PRDM12, SCN10A* and other maker genes (Fig. 1E, 1F, 5A, & 5D1). The hPEP.PIEZO^{h} cluster was named on the basis of a relatively high number of *PIEZO1* transcripts (the overall expression of *PIEZO1* in human DRG neurons was low, Fig. S8D) and the highest number of *PIEZO2* transcripts among all PEP clusters, an expression level comparable to that in hC.LTMR and hAδ.LTMR (Fig. 5B1). This cluster could also be identified by its expression of *PTGER3* but not *TRPA1* (Fig. 5B1 & 5B2). Interestingly, the adrenoreceptor, *ADRA2C*, a molecular marker for sensory fibers innervating arteries and blood vessels[51-54], was specifically detected in this cluster (Fig. S8I). In addition, *GPR68*, a membrane receptor reported to sense flow within the vascular system[55], was also expressed in the hPEP.PIEZO^{h} population (Fig. S8J). Given functions of *PIEZO1* and *PIEZO2* in mediating neuronal sensing of blood pressure and the baroreceptor reflex[56] and the expression and function of *ADRA2C* and *GPR68*, we proposed that some hPEP.PIEZO^{h} afferents might innervate blood vessels and sense the blood pressure or flow. This cluster also expressed high levels of *PTGIR* (Fig. S8H). Since mouse *PTGIR*$^{+}$ DRG neurons innervate the bladder[50], and because *PIEZO2* in mouse sensory neurons is known to be required for sensing the bladder

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

pressure to coordinate urination[57], some hPEP.PIEZO^{h} afferents might also sense non-noxious mechanical forces in visceral organs and participate in reflexes like urination.

The hPEP.KIT cluster had the specific expression of a receptor tyrosine kinase, *KIT*, and a medium expression level of *PIEZO2* (Fig. 5C1 & 5C2). In mouse DRG neurons, however, *Kit* is highly expressed in four peptidergic clusters[16] (Fig. S7). Cross-species analysis as well as the co-expression patterns with additional molecular markers, such as *SCG2* and *OPCML* (S9B4 & S9B5), suggest the cluster mainly corresponded to the mouse PEP3/CGRP-η and macaque PEP3 population (Fig. 2 & S7), which are A-fiber high threshold mechanoreceptors (HTMRs), forming circumferential endings around hair follicles and mediating hair pulling pain[58,59]. Thus, the hPEP.KIT cluster likely functions as fast-conducting mechano-nociceptors.

The third peptidergic A-fiber cluster, hPEP.CHRNA7, featured high expression of *CHRNA7* but almost no expression of *PIEZO2* (Fig. 5D1, 5C1, & 5D2). All our cross-species analysis revealed that this cell type corresponds to mouse PEP2/CGRP-ζ and macaque PEP2 populations, which are also marked by the unique expression of *CHRNA7*[16,18] (Fig. S7). Interestingly, this cluster also expressed *PVALB* (Fig. 1G), a molecular marker for proprioceptors in mouse. Since retrograde tracing from the mouse gastrocnemius muscle labeled *CHRNA7*$^{+}$ DRG neurons[50], hPEP.CHRNA7 may contain muscle- or deep tissue-innervating sensory afferents. Studies of *PIEZO2*-deficient patients reveal the existence of A-fibers that mediate PIEZO2-independent deep pressure sensation[60,61]. This cluster might be a candidate for such a function.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

hPEP.NTRK3 is a population of peptidergic A-fibers with high expression of *NTRK3* and *S100A4* (Fig. 5E1 & 5E2). Neurons in this cluster expressed a low level of *PIEZO2*. Finally, hPEP0 is a population of peptidergic A-fibers that expressed *CALCA* and a moderate level of PIEZO2 but lack of other specific marker genes. Potential function of hPEP.NTRK3 and hPEP0 are unclear but could be some types of A-fiber mechano-nociceptor[62] or candidates for A-fiber silent nociceptors[63].

Validation of molecular marker expressed in C-LTMRs, A-LTMRs and an ATF3 population

LTMRs were featured by their high expression level of *PIEZO2* and other marker genes (Fig. 6A & 6B1). hC.LTMR, the putative human C-tactile nerve fibers, also highly expressed *GFRA2* and *ZNF521* (Fig. S8K & S8L), similar to the mouse and macaque homologues (Fig. S9A5, S9A6, S9B6, S9B7). However, different from mice, where C.LTMR cells are characterized by exclusive expression of tyrosine hydroxylase (TH) and *SLC17A8* (*VGLUT3*) (Fig. S9A7 & S9A8)[15], these two genes were barely detectable in human DRG neurons (Fig. S8M & S8N). *CASQ2* was a specific marker gene for hC.LTMR (Fig. 6B1 & 6B2), but it was not detected in mouse and macaque C-LTMRs[15,18] (Fig. S7). Multiplex FISH confirmed that $CASQ2^{+}$ cells expressed high levels of *PIEZO2* (Fig. 6B2). hC.LTMRs likely mediate innocuous affective touch sensation[64-66].

A-LTMRs were featured by large diameter somas and high expression of *NTRK2* and *NTRK3* (Fig. 1E & 6C1-D2) but lack of *SCN10A*[67,68] and *PRDM12*[28], two genes highly associated with human nociception. We identified 4 clusters in this category. One cluster was named hAδ.LTMR based on its high expression level of *NTRK2* and *PIEZO2*, but low level of *NTRK3* (Fig. 6B1, 6C1-D2), a molecular feature similar to the mouse Aδ-LTMRs (Fig. S9A9 & S9A10). LTMR afferents with Aδ conduction velocity are seldom encountered in microneurography recordings

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

from the human lower extremity[62] though, so this name may indicate molecular rather than physiological features. Voltage-gated potassium channel *KCNV1* was enriched in this cluster and could serve as a molecular marker for identifying this population (Fig. 6C1 & 6C2). hAβ.LTMR expressed higher level of *NTRK3* but a lower level of *NTRK2* compared to hAδ.LTMRs (Fig. 6D1 & 6D2). hPropr expressed a high level of a proprioceptor marker *PVALB* and *REEP5* (Fig. 6E1 & 6E2). *ASIC1*, a proton-sensing receptor[69], was also highly expressed in some of these neurons (Fig. S8O).

We also identified a cluster named as hATF3, which contained large diameter neurons and corresponded the "unknown" cluster first identified in male mice by the Sharma RNA-seq dataset[16]. This cluster in both human and mouse datasets showed conserved molecular profiles, expressing high levels of *ATF3*, a marker of neural injury, and genes *GADD45A* and *SDC1* (Fig.S7, S8P S9A11-13). In human, these neurons also expressed high levels of neuropeptide *ADCYAP1*, (Fig. 6F1 & 6F2), as well as a precursor of the signaling peptide urocortin (*UCN*) (Fig. S7). This cluster was found in all three human donors, potentially with more neurons from the two older female donors (the current n number is too low to draw a confident conclusion). Sequencing more human DRG neurons from additional donors would be necessary to differentiate whether the hATF3 population displays any sex preference and/or reflects age-related degeneration.

The single-soma deep RNA-seq dataset provides novel insights into the molecular and cellular mechanisms underlying human itch and pain sensation

Physiological recordings have identified at least two groups of C-fibers responding to different pruritogens. One group respond to cowhage, a plant triggering intense itch in human[70], and

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

comprise mechano-sensitive polymodal units[40]. The other group responds to histamine with sustained discharges but are mechanically insensitive[40]. To provide insights into the molecular and cellular mechanisms underlying these response properties, we analyzed the expression of histamine and cowhage receptors, and the mechanoreceptor *PIEZO2* in the three potential itch populations, hNP1, hNP2 and hPEP.SST. PIEZO2 was highly expressed in hNP1, with low expression in hNP2 and almost no detectable expression in hPEP.SST (Fig. 3B1). Protease-activated receptor *F2RL1*, the receptor mediating cowhage induced itch in humans[70], was exclusively expressed in hNP1 (Fig. S11A). Thus, the hNP1 population likely contains C-fiber pruriceptive afferents sensitive to cowhage and mechanical forces. Of the four histamine receptors, *HRH1* was expressed in all three itch populations, *HRH2* was expressed in hNP2 and hPEP.SST but not in hNP1, *HRH3* had low expression in hNP1 and hPEP.SST, while *HRH4* was not detected (Fig. S11B-E). Thus, hNP2 and hPEP.SST clusters are good candidates for histamine-sensitive but mechano-insensitive itch-sensing C-fibers.

Some chemicals trigger both itch and pain sensation in humans, such as prostaglandin E2 (PGE2), serotonin, and bradykinin[71]. We examined the expression of all known receptors for these ligands in our dataset. Interestingly, we found that some of these receptors were expressed in both itch- and pain-associated populations. For example, the PGE2 receptor *PTGER2* had low expression in hPEP.SST putative itch-sensing neurons (Fig. S11F), while another PGE2 receptor, *PTGER3*, was expressed in the hPEP.TRPV1/A1 putative nociceptive population (Fig. 4D1). Serotonin receptors *HTR1B* and *HTR1F* were expressed in the hPEP.TRPV1/A1.2 nociceptive population, while a subset of cells in hNP2 and hPEP.SST itch populations expressed low levels of another serotonin receptor, *HTR5A* (Fig. S11G-I). Bradykinin receptor *BDKRB2* was expressed in hNP1 at low levels and in hPEP.TRPV1/A1.1 at high levels (Fig. S11J). The expression patterns of these receptors raise the possibility that the

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

dual somatosensory effects of a single compound might be explained by receptor expression in different functional groups of primary sensory afferents.

Mining single-soma deep RNA-seq dataset for validating novel drug targets and obtaining insights into mechanisms underlying drug effects

A subset of itch-related receptors exhibited conserved expression patterns in hNP1, hNP2 and hPEP.SST populations among human, macaque, and mouse, such as the histamine receptor *HRH1*, leukotriene receptor *CYSLTR2*, and interleukin receptors *IL31RA* and *OSMR* (Fig. S12A, S10)[16,18]. On the other hand, some itch receptors exhibited distinct expression patterns across different species. For example, *MRGPRX4*, a bile acid and bilirubin receptor for human cholestatic itch[20,72], was specifically expressed in the NP2 population in human (Fig. S12A) and macaque DRG neurons but has no ortholog in mice. *HRH4* mediates histamine-dependent itch in mice[73], but expression of *HRH4* was barely detected in human itch-sensing populations (Fig. S12B), suggesting that *HRH4* may not be directly involved in histaminergic itch in humans. Some receptors has been proposed to mediate non-histaminergic itch in mice, such as *IL7R* and *TLR7*[74], but they were not enriched in human itch-sensing populations (Fig. S12B). Moreover, some drugs display anti-itch effects with unclear mechanisms, such as *IL4R*, *IL13RA1* monoclonal antibodies[75,76], and gabapentin[77,78]. From our dataset, *IL4R* was not highly expressed in human DRG neurons, whereas *IL13RA1* was enriched in hTRPM8 and PEP cluster (Fig. S12B). These expression patterns suggest that *IL4R* might contribute to itch modulation in non-DRG-neuron cells and that *IL13RA1* could affect itch by cross-modality antagonization. Gabapentin inhibits neurotransmitter release by acting on α2δ-1 and α2δ-2 voltage-dependent calcium channels *CACNA2D1* and *CACNA2D2*[79]. Interestingly, both receptors were highly expressed in human itch and pain DRG neurons (Fig. S11K & S11L),

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

suggesting that one potential mechanism by which gabapentin could provide clinical benefit is through inhibiting synaptic transmission of itch- and pain-sensing primary afferents.

Endogenous and exogenous opioids acting through opioid receptors modulate pain and itch. Agonists of the μ-opioid receptor (*OPRM1*) alleviate pain but elicit itch in humans and model organisms, whereas agonists of the κ-opioid receptor (*OPRK1*) inhibit itch in mice[80,81]. Unlike the lack of any clear expression patterns of opioid receptors in mouse DRG neurons, we found a relatively modality-specific expression patterns of opioid receptors in human DRG neurons. Transcripts of δ-opioid receptor (*OPRD1*) were preferentially expressed in itch populations hNP1 and hNP2, while *OPRM1* was enriched in all hPEP clusters (Fig. S11M & S11N). Since opioid receptors are inhibitory GPCRs, our results suggest that activation of *OPRM1* could directly inhibit human nociceptive afferents while *OPRD1* could be a molecular target for inhibiting itch transmission. In addition, activation of *OPRM1* may elicit itch sensation by central mechanisms or by dis-inhibiting antagonization effects from the nociceptors. On the other hand, *OPRK1* was barely detected in human DRG neurons (Fig. S11O), indicating *OPRK1* agonists may relieve itch through indirect or central mechanisms. Finally, our dataset revealed a series of additional genes, including GPCRs, ion channels, chemokine and cytokine receptors, that were specifically expressed in human itch populations and could be potential novel anti-itch targets (Fig. S12C & S13-S16). In short, our single-soma deep RNA-seq dataset of human DRG neurons provides important insights to understand normal somatosensation and for discovering novel molecular targets for treating pain and itch disorders. Our analyses also highlight that it is critical to validate molecular targets when translating anti-nociception and anti-itch treatment results from animal models to humans.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Microneurography evaluation of novel physiological properties of human sensory afferents predicted by single-soma deep RNA-seq dataset

Our single-soma deep RNA-seq dataset of human DRG neurons also provided novel information to predict the physiological properties and functions of somatosensory afferents. The PEP3/CGRP-η peptidergic A-fiber population of fast-conducting hair pull-sensitive mechano-nociceptors has been identified in mice and found to be CGRP^{+}, PIEZO2^{+}, but TRPV1^{-}[82] (Fig. S17A). In our dataset, hPEP.KIT neurons showed similar expression patterns of these marker genes (Fig. S17B). Indeed, cell type correlation analysis between human and mouse revealed that hPEP.KIT correspond to the mouse PEP3/CGRP-η cluster[15,18] (Fig. 2A-C and Fig. S6). Interestingly, our dataset revealed that hPEP.KIT neurons expressed a low level of *TRPM8* (Fig. S17B), a feature that was not found in the mouse DRG neuron dataset (Fig. S17A). Multiplex FISH validated the co-expression of *KIT*, *PIEZO2* and *TRPM8* in human DRG neurons (Fig. S17C). Thus, we hypothesized that some fast-conducting A-fiber HTMRs in human skin could be activated by cooling but not heating. To test this idea, using the *in vivo* electrophysiological technique of microneurography, single-unit axonal recordings were performed in the radial, antebrachial, and peroneal nerves of healthy participants (Fig. 7A-C).

The A-HTMRs (n=10) were identified by their insensitivity to soft-brush stroking (they responded to a rough brush) and high indentation thresholds (≥4 mN); further, they had Aβ-range conduction velocities (>30 m/s, Fig. 7A-F). Remarkably, a subtype of these A-HTMRs (n=5) responded to cooling but not heating stimuli (Fig. 7G-H). Compared to mechanically evoked responses, those evoked by cooling (30 to 0°C at 4°C per second) were relatively weak (a few spikes per second) but reproducible (tested in triplicates per recording) and followed the temperature decrements for several seconds before plateauing out (Fig. S18A). Further, the responses to cooling/cold persisted during the sustained phase (0°C, Fig. S18A). These

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

observations confirm our prediction that some A-HTMRs are cold-sensitive. We also recorded from the other type of Aβ-HTMR (n=5, referred to as A-HTMR cooling- in Fig. 7A-C), which, as reported previously[62], did not respond to heating or to cooling (data not shown).

Human C-LTMRs are readily found during microneurography recording from the upper limb, but they seem to sparsely innervate the distal lower limb[83]. This is consistent with our sequencing results wherein the hC.LTMRs constituted a small population of neurons in lower thoracic and lumbar level DRGs. Unexpectedly, our sequencing data revealed that human hC.LTMRs had almost no expression of the cold and menthol receptor, *TRPM8*, though C.LTMRs display sensitivity to cooling in both humans and mice [84,85]. Instead, hC.LTMRs expressed *TRPV1* (Fig. S19A). Multiplex FISH confirmed that *CASQ2*$^{+}$ hC.LTMRs were TRPM8^{-} but TRPV1^{+} (Fig. S19B). This expression pattern suggested that hC.LTMRs might respond to heating and capsaicin, a novel physiological property that has not been discovered in rodent or non-human primate models, but no response to menthol. In human microneurography, the C-LTMRs (n=11) were identified by their soft-brush sensitivity, low indentation threshold (in this case, ≤0.7 mN), and slow conduction velocity (~1 m/s, n=11, Fig. 7A-C & 7I-J). Consistent with mouse C-LTMRs, these human C-LTMRs responded to hair movement (Fig. 7K) and dynamic cooling (Fig. 7L). In two of them, after having confirmed the cooling response, we applied menthol to their individual receptive fields leading to a cold sensation, but those recorded C-LTMRs, while still responsive to mechanical and thermal stimuli, were not activated by menthol[83] (Fig. S19C & S19J). Remarkably, microneurography recordings showed that a subset of human C-LTMRs responded to dynamic heating (5 out of 11 units, Fig. 7M). In three of them, after having confirmed the heating response, we applied capsaicin to their receptive fields, and, consistent with the sequencing results, all three were activated by capsaicin (Fig. S19D). Compared to mechanical stimulation, the C-LTMR thermal responses were relatively weak (just a few spikes

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

per second) but reproducible (tested in triplicates per modality per recording). Further, a comparison of C-LTMR responses to cooling (from 30 to 0°C at 4°C per s) and heating (up to 50°C) ramps showed a preference for intermediate cold and warm temperatures (Fig. S18B). The existence of a polymodal (mechano-heat-cold) C-LTMR type is novel and confirms the sequencing predictions. Further, the lack of menthol response in C-LTMRs suggests that they express a non-TRPM8 cold receptor.

For comparison, our recordings also identified C-HTMRs (n=11) that did not respond to soft-brush stroking and hair movement (they responded to a rough brush), had high indentation thresholds (in this case, ≥10 mN), and slow conduction velocities (~1 m/s, Fig. 7A-C & S19E-H). Based on their temperature responses (tested in 9 units), a mix of C mechano-heat (n=6), C mechano-cold (n=2), and C polymodal (mechano-heat-cold, n=1) subtypes were identified. An example of heating and capsaicin responses of a C mechano-heat nociceptor is shown in Fig. S19I. Collectively, these results highlight the accuracy and utility of our single-soma deep RNA-seq dataset of human DRG neurons for functional characterizations of the human somatosensory system.

Discussion

While somatosensation is critical for the daily lives and survival of humans, a malfunction can have devastating consequences, such as during chronic pain, chronic itch, and peripheral neuropathy. At present, the molecular and cellular mechanisms underlying normal and pathological human somatosensation remain largely elusive. In this study, we developed a novel LCM-based approach for single-soma deep RNA-seq and sequenced over 1000 human DRG neurons. This approach allowed us to generate the first single-cell RNA-seq dataset from

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

adult human neuronal somas, as far as we are aware. The sequenced human DRG neurons were clustered into 16 molecular groups displaying similarities to and differences from macaque and mouse DRG neurons. We believe that our dataset provides important insights into human pain and itch sensory phenomena, explains mechanisms of drug effects, and as also exemplified in the study, represents a rich resource to identify new molecular targets for modulating the activity of itch- and pain-sensing primary afferents. Moreover, we discover novel physiological properties of human DRG neurons, which are predicted by molecular profiles from this dataset.

Single-cell RNA-seq of human DRG/TG neurons has been technically challenging. One main hurdle is to isolate sensory neurons from a large population of non-neuronal cells. The traditional enzymatic and mechanical dissociation method is incompatible with human DRG/TG neurons, as they are large, fragile, and few in number. Strategies, including spatial transcriptomics and single-nucleus RNA-seq [22,23,86], have generated some pioneering datasets characterizing the molecular profiles of human DRG and TG neurons. Nevertheless, the number of transcripts, sequencing depth, or single-cell resolution of the previous studies needs to be improved. Moreover, though nuclear and soma transcripts are intrinsically linked and overlapping to some extent, the soma contains much more transcripts both quantitatively and qualitatively. Soma transcripts are also one step closer than nuclear transcripts for cellular functions. Thus, soma transcripts are better, if available, for cell type clustering and functional interpretation. In this study, we developed a new strategy by combining fresh frozen human DRG tissues, cryo-section, laser capture microdissection of individual neuronal soma, and Smart-seq2 deep sequencing. Fresh frozen tissue and cryosection techniques minimized transcriptomic changes during the tissue transportation and single-cell isolation process. LCM allowed isolation of DRG neuronal soma with minimal contamination from surrounding non-

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

neuronal cells, while preserving information about cellular morphology and localization, and the Smart-seq2 protocol enabled a high recovery rate of mRNA molecules. Although LCM has been used for isolating a group of neurons for RNA-seq[6,7], the successful application of LCM for single-cell transcriptomic analysis has not been achieved before. Thus, we have established a new method for single-soma deep RNA-seq of human DRG neurons. Our approach should be readily applicable to other human neurons with large soma sizes, such as trigeminal ganglion neurons, nodose ganglion neurons, motor neurons, etc.

Different single-cell RNA-seq approaches, including single-nucleus RNA-seq[21,23,86], spatial transcriptomics[22], and our LCM-based single soma RNA-seq, have generated four datasets of transcriptome profiles and cell type clusters of human DRG neurons. These datasets and results overlap to some extent but also exhibit some obvious differences (an example shown in Fig. S5). The observed differences are likely caused by both biological and technical features associated with each method. Given the high sequencing depth of transcripts from the neuronal soma, our approach is powerful for molecular discovery, especially for functional molecules expressed at a low level. For example, our approach detected the specific expression of *MRGPRX1*, a pruritogen-sensing GPCR, in hNP1 and hNP2 neurons, while the previous datasets barely detected this transcript[22,23].

Our analysis of the expression of receptors, ion channels, and neuropeptides in human itch-sensing DRG neurons have identified a set of potential new targets for modulating the activities of these sensory afferents. Furthermore, the sequencing and microneurography results also suggested the existence of a non-TRPM8 cold receptor in human hC.LTMRs. We noticed a low expression level of *TRPA1*, which also has cold sensitivity[87,88], in this population (Fig. 4C1).

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Whether *TRPA1* or some currently unknown cold receptors mediate cooling sensitivity in hC.LTMRs will be of interest for future research. Using microneurography, we also demonstrated heat/capsaicin responses in C-LTMRs, a functional validation of TRPV1 expression in these afferents (Fig. S19D). To our knowledge, this is the first evidence of such properties of human non-nociceptive neurons. The discovery of the hPEP.KIT population indicated a potential role for PIEZO2 in human mechano-nociception. This population also responded to innocuous and noxious cooling, which is a curious property of a large-diameter myelinated nociceptor. Our discovery of TRPM8 expression in these neurons provides a molecular explanation for this property. Moreover, a likely molecular type for C-HTMRs, which display a mixture of responsiveness to mechanical forces and temperature, is the hPEP.TRPV1/A1.1 population. Since hPEP.TRPV1/A1.1 neurons have no expression of *PIEZO2* (Fig. 3B1) or *PIEZO1* (Fig. S8D), a non-*PIEZO* mechanoreceptor may exist in these neurons for mediating mechanical pain sensation. This is consistent with reports that patients with *PIEZO2* loss-of-function mutations have normal mechanical pain threshold and sensitivity[60,62]. The other suggested mechanical pain channel, TMEM120A (TACAN), is broadly expressed in all types of human DRG neurons (Fig. S8Q). This expression pattern does not support its role as the mechanical pain receptor in humans. C-LTMRs and A-HTMRs responded to temperature changes in human microneurography, even if those were weak responses, suggests that the requisite circuity is already in place; this may have important implications for thermal hypersensitivities in pathological states. Indeed, there is indirect evidence from human psychophysics and targeted pharmacology, pointing to the role of C-LTMRs in mediating acute cold allodynia[89] and the role of A fibers in signaling chronic cold allodynia[90]. TAC1 (substance P) is a specific marker for mouse nociceptors, and its receptor, NK1R, is a molecular target for chronic pain and itch treatment[10]. Though working effectively in rodent models, NK1R antagonists have failed several chronic-pain-related clinical trials[91]. Our dataset showed specific

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

expression of TAC1 in four types of peptidergic afferents, hPEP.SST, hPEP.TRPV1/A1.1, hPEP.TRPV1/A1.2, and hPEP.CHRNA7 (Fig. S8R). One potential reason for the failure of NK1R antagonist in treating chronic pain could be a mismatch between the tested chronic pain conditions and the types of nociceptive afferents involved. In short, our dataset generates an important atlas for understanding human somatosensation and a valuable reference for designing translational studies.

Cross-species comparisons among human, macaque, and mouse DRG neurons suggested that the broader functional groups of DRG sensory afferents are conserved across species. Nevertheless, there are also critical differences at molecular and cellular levels. For example, the molecular profiles of NP1 and NP2 afferents in human and macaque are more similar to each other than to their corresponding subtypes in the mouse (Fig. S7). The MRGPRX family, which constitutes their main markers and functional molecules, has only 4 members in human and macaque[92], but expands into three subfamilies with over 50 genes in the mouse[93]. Since NP1 and NP2 afferents mainly function to detect irritants in the external environment, this molecular change may reflect the adaptation of different animal species to their living environment. PEP afferents showed the most evolutionary divergence between mouse and human and exhibited significant molecular diversity in our dataset. The hPEP.PIEZO^{h} cluster seems to be a primate specific population or greatly expanded in the human and macaque (Fig. 5A & 5B1-B2), as very few mouse *CGRP*+ DRG neurons show such a high expression level of *PIEZO2* and *PIEZO1* (Fig. S9A1, S9A2 & S9A14). Since both human and macaque have much larger body sizes than mice, the task to regulate blood pressure and local blood flow in different body parts would be significantly more demanding for them. Validating and characterizing different molecularly defined human DRG neuron subtypes, especially the diverse PEP clusters that we identified, will be another important future research direction.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

In this study, we sequenced over 1000 DRG neurons from 6 DRGs at the thoracic (T11-T12) and lumbar (L2-L5) levels of three healthy Caucasian human donors, with ~150 to 200 neurons per DRG, which represent 1-2% of the total neurons within a human DRG. Thus, it is necessary to sequence more neurons for a reliable representation. We know the existence of different types of RA and SA Aβ-LTMRs and different types of proprioceptors in humans, but our current dataset did not have the resolution to separate them. Thus, with more human DRG neuron sequenced, we anticipate an even greater heterogeneity to be discovered. In addition, DRGs at different spinal levels innervate different peripheral target tissues[5]. Sequencing DRG neurons from different spinal levels will help to uncover molecular and cellular mechanisms underlying physiological functions. For instance, single-cell transcriptomic studies of sacral DRG neurons might elucidate the sensory mechanisms of genital sensation, whereas the characterization of cervical DRG neurons might help understand how the human hand mediates fine tactile sensation. Moreover, increased sampling from donors representing different demographics would also be critical for investigating sex, race, and age-related differences. Lastly, pathological conditions greatly alter the transcriptomic landscape[94,95]. Systematic comparison of molecular and cellular changes between healthy donors and those with chronic itch or chronic pain would be of great value, if samples are available, for understanding pathological mechanisms and identifying molecular targets for effective treatments.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Reference:

1 Nascimento, A. I., Mar, F. M. & Sousa, M. M. The intriguing nature of dorsal root ganglion neurons: Linking structure with polarity and function. *Prog Neurobiol* **168**, 86-103, doi:10.1016/j.pneurobio.2018.05.002 (2018).

2 Purves, D. A., George J.; Fitzpatrick, David; Katz, Lawrence C.; LaMantia, Anthony-Samuel; McNamara, James O.; Williams, S. Mark. The Major Afferent Pathway for Mechanosensory Information: The Dorsal Column-Medial Lemniscus System. *Neuroscience. 2nd edition* (2001).

3 Azimi, E., Xia, J. & Lerner, E. A. Peripheral mechanisms of itch. *Itch-Management in Clinical Practice* **50**, 18-23 (2016).

4 Abraira, V. E. & Ginty, D. D. The sensory neurons of touch. *Neuron* **79**, 618-639, doi:10.1016/j.neuron.2013.07.051 (2013).

5 Luo, L. *PRINCIPLES OF NEUROBIOLOGY*. 265-267 (Garland Science, 2016).

6 Mills, S. E. E., Nicolson, K. P. & Smith, B. H. Chronic pain: a review of its epidemiology and associated factors in population-based studies. *Br J Anaesth* **123**, e273-e283, doi:10.1016/j.bja.2019.03.023 (2019).

7 Manglik, A. *et al.* Structure-based discovery of opioid analgesics with reduced side effects. *Nature* **537**, 185-190, doi:10.1038/nature19112 (2016).

8 Yosipovitch, G., Rosen, J. D. & Hashimoto, T. Itch: From mechanism to (novel) therapeutic approaches. *J Allergy Clin Immunol* **142**, 1375-1390, doi:10.1016/j.jaci.2018.09.005 (2018).

9 Burma, N. E., Leduc-Pessah, H., Fan, C. Y. & Trang, T. Animal models of chronic pain: advances and challenges for clinical translation. *Journal of neuroscience research* **95**, 1242-1256 (2017).

10 Hill, R. NK1 (substance P) receptor antagonists–why are they not analgesic in humans? *Trends in pharmacological sciences* **21**, 244-246 (2000).

11 Yu, H. *et al.* MRGPRX4 is a bile acid receptor for human cholestatic itch. *Elife* **8**, e48431 (2019).

12 Alemi, F. *et al.* The TGR5 receptor mediates bile acid–induced itch and analgesia. *The Journal of clinical investigation* **123**, 1513-1530 (2013).

13 Tang, F. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* **6**, 377-382, doi:10.1038/nmeth.1315 (2009).

14 Hwang, B., Lee, J. H. & Bang, D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp Mol Med* **50**, 96, doi:10.1038/s12276-018-0071-8 (2018).

15 Usoskin, D. *et al.* Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. *Nature neuroscience* **18**, 145-153 (2015).

16 Sharma, N. *et al.* The emergence of transcriptional identity in somatosensory neurons. *Nature* **577**, 392-398 (2020).

17 Wang, K. *et al.* Single-cell transcriptomic analysis of somatosensory neurons uncovers temporal development of neuropathic pain. *Cell research* **31**, 904-918 (2021).

18 Kupari, J. *et al.* Single cell transcriptomics of primate sensory neurons identifies cell types associated with chronic pain. *Nature communications* **12**, 1-15 (2021).

19 Haberberger, R. V., Barry, C., Dominguez, N. & Matusica, D. Human Dorsal Root Ganglia. *Front Cell Neurosci* **13**, 271, doi:10.3389/fncel.2019.00271 (2019).

20 Stuart, T. *et al.* Comprehensive integration of single-cell data. *Cell* **177**, 1888-1902. e1821 (2019).

21 Jung, M. *et al.* Cross-species transcriptomic atlas of dorsal root ganglia reveals species-specific programs for sensory function. *Nature Communications* **14**, 366 (2023).

22 Tavares-Ferreira, D. *et al.* Spatial transcriptomics of dorsal root ganglia identifies molecular signatures of human nociceptors. *Science translational medicine* **14**, eabj8186 (2022).

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

23 Nguyen, M. Q., von Buchholtz, L. J., Reker, A. N., Ryba, N. J. & Davidson, S. Single-nucleus transcriptomic analysis of human dorsal root ganglion neurons. *Elife* **10**, e71752 (2021).

24 Ding, J. *et al.* Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. *Nature biotechnology* **38**, 737-746 (2020).

25 Mattei, D. *et al.* Enzymatic dissociation induces transcriptional and proteotype bias in brain cell populations. *International journal of molecular sciences* **21**, 7944 (2020).

26 Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nature protocols* **9**, 171-181 (2014).

27 Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587. e3529 (2021).

28 Chen, Y.-C. *et al.* Transcriptional regulator PRDM12 is essential for human pain perception. *Nature genetics* **47**, 803-808 (2015).

29 Barkas, N. *et al.* Joint analysis of heterogeneous single-cell RNA-seq dataset collections. *Nature methods* **16**, 695-698 (2019).

30 Jovic, D. *et al.* Single-cell RNA sequencing technologies and applications: A brief overview. *Clin Transl Med* **12**, e694, doi:10.1002/ctm2.694 (2022).

31 Zeisel, A. *et al.* Molecular architecture of the mouse nervous system. *Cell* **174**, 999-1014. e1022 (2018).

32 Jeub, M. *et al.* The transcription factor Smad-interacting protein 1 controls pain sensitivity via modulation of DRG neuron excitability. *PAIN®* **152**, 2384-2398 (2011).

33 Scott, A. *et al.* Transcription factor short stature homeobox 2 is required for proper development of tropomyosin-related kinase B-expressing mechanosensory neurons. *Journal of Neuroscience* **31**, 6741-6749 (2011).

34 Abdo, H. *et al.* Dependence on the transcription factor Shox2 for specification of sensory neurons conveying discriminative touch. *European Journal of Neuroscience* **34**, 1529-1541 (2011).

35 Levanon, D. *et al.* The Runx3 transcription factor regulates development and survival of TrkC dorsal root ganglia neurons. *The EMBO journal* **21**, 3454-3463 (2002).

36 Chen, C.-L. *et al.* Runx1 determines nociceptive sensory neuron phenotype and is required for thermal and neuropathic pain. *Neuron* **49**, 365-377 (2006).

37 Vermeiren, S., Bellefroid, E. J. & Desiderio, S. Vertebrate sensory ganglia: common and divergent features of the transcriptional programs generating their functional specialization. *Frontiers in Cell and Developmental Biology* **8**, 587699 (2020).

38 Liu, Q. & Dong, X. The role of the Mrgpr receptor family in itch. *Pharmacology of itch*, 71-88 (2015).

39 Dong, P. *et al.* TRPC3 is dispensable for β-alanine triggered acute itch. *Scientific reports* **7**, 1-12 (2017).

40 Namer, B. *et al.* Separate peripheral pathways for pruritus in man. *Journal of neurophysiology* **100**, 2062-2069 (2008).

41 McNeil, B. & Dong, X. Peripheral mechanisms of itch. *Neuroscience bulletin* **28**, 100-110 (2012).

42 Colburn, R. W. *et al.* Attenuated cold sensitivity in TRPM8 null mice. *Neuron* **54**, 379-386 (2007).

43 Bautista, D. M. *et al.* The menthol receptor TRPM8 is the principal detector of environmental cold. *Nature* **448**, 204-208 (2007).

44 Dhaka, A. *et al.* TRPM8 is required for cold sensation in mice. *Neuron* **54**, 371-378 (2007).

45 Samour, M. S., Nagi, S. S. & Mahns, D. A. Cav3. 2-expressing low-threshold C fibres in human hairy skin contribute to cold allodynia—a non-TRPV1-and non-TRPM8-dependent phenomenon. *Pain* **156**, 1566-1575 (2015).

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

46 Campero, M., Baumann, T., Bostock, H. & Ochoa, J. Human cutaneous C fibres activated by cooling, heating and menthol. *The Journal of physiology* **587**, 5633-5652 (2009).
47 Caterina, M. J. *et al.* The capsaicin receptor: a heat-activated ion channel in the pain pathway. *Nature* **389**, 816-824 (1997).
48 Bandell, M. *et al.* Noxious cold ion channel TRPA1 is activated by pungent compounds and bradykinin. *Neuron* **41**, 849-857 (2004).
49 Liu, S. *et al.* A neuroanatomical basis for electroacupuncture to drive the vagal–adrenal axis. *Nature* **598**, 641-645 (2021).
50 Yang, F.-C. *et al.* Genetic control of the segregation of pain-related sensory neurons innervating the cutaneous versus deep tissues. *Cell reports* **5**, 1353-1364 (2013).
51 Albrecht, P. J. *et al.* Pathologic alterations of cutaneous innervation and vasculature in affected limbs from patients with complex regional pain syndrome. *Pain* **120**, 244-266 (2006).
52 Morelli, C. *et al.* Identification of a population of peripheral sensory neurons that regulates blood pressure. *Cell reports* **35**, 109191 (2021).
53 Albrecht, P. J. *et al.* Excessive peptidergic sensory innervation of cutaneous arteriole–venule shunts (AVS) in the palmar glabrous skin of fibromyalgia patients: Implications for widespread deep tissue pain and fatigue. *Pain medicine* **14**, 895-915 (2013).
54 Bowsher, D. *et al.* Absence of pain with hyperhidrosis: a new syndrome where vascular afferents may mediate cutaneous sensation. *PAIN®* **147**, 287-298 (2009).
55 Xu, J. *et al.* GPR68 senses flow and is essential for vascular physiology. *Cell* **173**, 762-775. e716 (2018).
56 Zeng, W.-Z. *et al.* PIEZOs mediate neuronal sensing of blood pressure and the baroreceptor reflex. *Science* **362**, 464-467 (2018).
57 Marshall, K. L. *et al.* PIEZO2 in sensory neurons and urothelial cells coordinates urination. *Nature* **588**, 290-295 (2020).
58 Willimann, B. & Trüeb, R. M. Hair pain (trichodynia): frequency and relationship to hair loss and patient gender. *Dermatology* **205**, 374-377 (2002).
59 von Buchholtz, L. J. *et al.* Decoding cellular mechanisms for mechanosensory discrimination. *Neuron* **109**, 285-298. e285 (2021).
60 Chesler, A. T. *et al.* The role of PIEZO2 in human mechanosensation. *New England Journal of Medicine* **375**, 1355-1364 (2016).
61 Case, L. K. *et al.* Innocuous pressure sensation requires A-type afferents but not functional PIEZO2 channels in humans. *Nature communications* **12**, 657 (2021).
62 Nagi, S. S. *et al.* An ultrafast system for signaling mechanical pain in human skin. *Science advances* **5**, eaaw1297 (2019).
63 Schmidt, R. F. *Encyclopedia of Pain*. pp 2171–2173 (Spriner 2007).
64 Löken, L. S., Wessberg, J., McGlone, F. & Olausson, H. Coding of pleasant touch by unmyelinated afferents in humans. *Nature neuroscience* **12**, 547-548 (2009).
65 Olausson, H., Wessberg, J., McGlone, F. & Vallbo, Å. The neurophysiology of unmyelinated tactile afferents. *Neuroscience & Biobehavioral Reviews* **34**, 185-191 (2010).
66 Huzard, D. *et al.* The impact of C-tactile low-threshold mechanoreceptors on affective touch and social interactions in mice. *Science Advances* **8**, eabo7566 (2022).
67 Xue, Y., Chidiac, C., Herault, Y. & Gaveriaux-Ruff, C. Pain behavior in SCN9A (Nav1. 7) and SCN10A (Nav1. 8) mutant rodent models. *Neuroscience Letters* **753**, 135844 (2021).
68 Duan, G. *et al.* A SCN10A SNP biases human pain sensitivity. *Molecular pain* **12**, 1744806916666083 (2016).
69 Walder, R. Y. *et al.* ASIC1 and ASIC3 play different roles in the development of hyperalgesia after

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

inflammatory muscle injury. *The Journal of pain* **11**, 210-218 (2010).

70 Reddy, V. B., Iuga, A. O., Shimada, S. G., LaMotte, R. H. & Lerner, E. A. Cowhage-evoked itch is mediated by a novel cysteine protease: a ligand of protease-activated receptors. *Journal of Neuroscience* **28**, 4331-4335 (2008).

71 Schmelz, M. *et al.* Chemical response pattern of different classes of C-nociceptors to pruritogens and algogens. *Journal of neurophysiology* **89**, 2441-2448 (2003).

72 Meixiong, J., Vasavda, C., Snyder, S. H. & Dong, X. MRGPRX4 is a G protein-coupled receptor activated by bile acids that may contribute to cholestatic pruritus. *Proc Natl Acad Sci U S A* **116**, 10525-10530, doi:10.1073/pnas.1903316116 (2019).

73 Bell, J., McQueen, D. & Rees, J. Involvement of histamine H4 and H1 receptors in scratching induced by histamine receptor agonists in BalbC mice. *British journal of pharmacology* **142**, 374-380 (2004).

74 Dong, X. & Dong, X. Peripheral and central mechanisms of itch. *Neuron* **98**, 482-494 (2018).

75 Harb, H. & Chatila, T. A. Mechanisms of dupilumab. *Clinical & Experimental Allergy* **50**, 5-14 (2020).

76 Veverka, K. *et al.* Interim analysis results from a proof-ofconcept study for ASLAN004 in adult moderate to severe atopic dermatitis: A double blind, randomized, placebo-controlled study. *Journal of Clinical and Aesthetic Dermatology*, S13-S13 (2022).

77 Rayner, H., Baharani, J., Smith, S., Suresh, V. & Dasgupta, I. Uraemic pruritus: relief of itching by gabapentin and pregabalin. *Nephron Clinical Practice* **122**, 75-79 (2012).

78 Matsuda, K. M., Sharma, D., Schonfeld, A. R. & Kwatra, S. G. Gabapentin and pregabalin for the treatment of chronic pruritus. *Journal of the American Academy of Dermatology* **75**, 619-625. e616 (2016).

79 Calandre, E. P., Rico-Villademoros, F. & Slim, M. Alpha2delta ligands, gabapentin, pregabalin and mirogabalin: a review of their clinical pharmacology and therapeutic use. *Expert review of neurotherapeutics* **16**, 1263-1277 (2016).

80 Cowan, A., Kehner, G. B. & Inan, S. Targeting itch with ligands selective for κ opioid receptors. *Pharmacology of itch*, 291-314 (2015).

81 Ko, M. H., Song, M., Edwards, T., Lee, H. & Naughton, N. The role of central μ opioid receptors in opioid-induced itch in primates. *Journal of Pharmacology and Experimental Therapeutics* **310**, 169-176 (2004).

82 Ghitani, N. *et al.* Specialized mechanosensory nociceptors mediating rapid responses to hair pull. *Neuron* **95**, 944-954. e944 (2017).

83 Löken, L. S., Backlund Wasling, H., Olausson, H., McGlone, F. & Wessberg, J. A topographical and physiological exploration of C-tactile afferents and their response to menthol and histamine. *Journal of Neurophysiology* **127**, 463-473 (2022).

84 Nordin, M. Low-threshold mechanoreceptive and nociceptive units with unmyelinated (C) fibres in the human supraorbital nerve. *The Journal of physiology* **426**, 229-240 (1990).

85 Middleton, S. J. *et al.* Nav1. 7 is required for normal C-low threshold mechanoreceptor function in humans and mice. *Brain* **145**, 3637-3653 (2022).

86 Yang, L. *et al.* Human and mouse trigeminal ganglia cell atlas implicates multiple cell types in migraine. *Neuron* **110**, 1806-1821. e1808 (2022).

87 Bautista, D. M. *et al.* TRPA1 mediates the inflammatory actions of environmental irritants and proalgesic agents. *Cell* **124**, 1269-1282 (2006).

88 Macpherson, L. J. *et al.* Noxious compounds activate TRPA1 ion channels through covalent modification of cysteines. *Nature* **445**, 541-545 (2007).

89 Shaikh, S., Nagi, S. S., McGlone, F. & Mahns, D. A. Psychophysical investigations into the role of

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

low-threshold C fibres in non-painful affective processing and pain modulation. *PLoS One* **10**, e0138299 (2015).

90 Jørum, E. & Opstad, P.-K. A 4-year follow-up of non-freezing cold injury with cold allodynia and neuropathy in 26 naval soldiers. *Scandinavian journal of pain* **19**, 441-451 (2019).

91 Boyce, S. & Hill, R. Discrepant results from preclinical and clinical studies on the potential of substance P-receptor antagonist compounds as analgesics. *Progress in pain research and management* **16**, 313-324 (2000).

92 Lembo, P. *et al.* Proenkephalin A gene products activate a new family of sensory neuron–specific GPCRs. *Nature neuroscience* **5**, 201-209 (2002).

93 Dong, X., Han, S.-k., Zylka, M. J., Simon, M. I. & Anderson, D. J. A diverse family of GPCRs expressed in specific subsets of nociceptive sensory neurons. *Cell* **106**, 619-632 (2001).

94 Xiao, H.-S. *et al.* Identification of gene expression profile of dorsal root ganglion in the rat peripheral axotomy model of neuropathic pain. *Proceedings of the National Academy of Sciences* **99**, 8360-8365 (2002).

95 North, R. Y. *et al.* Electrophysiological and transcriptomic correlates of neuropathic pain in human dorsal root ganglion neurons. *Brain* **142**, 1215-1226 (2019).

96 Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal* **17**, 10-12 (2011).

97 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).

98 Hu, Y. *et al.* Neural network learning defines glioblastoma features to be of neural crest perivascular or radial glia lineages. *Science Advances* **8**, eabm6340 (2022).

99 Shrikumar, A., Greenside, P. & Kundaje, A. in *International conference on machine learning.* 3145-3153 (PMLR).

100 Huynh-Thu, V. A., Irrthum, A., Wehenkel, L. & Geurts, P. Inferring regulatory networks from expression data using tree-based methods. *PloS one* **5**, e12776 (2010).

101 Aibar, S. *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nature methods* **14**, 1083-1086 (2017).

102 Vallbo, A., Olausson, H., Wessberg, J. & Kakuda, N. Receptive field characteristics of tactile units with myelinated afferents in hairy skin of human subjects. *The Journal of physiology* **483**, 783-795 (1995).

103 Vallbo, Å., Olausson, H. & Wessberg, J. Unmyelinated afferents constitute a second system coding tactile stimuli of the human hairy skin. *Journal of neurophysiology* **81**, 2753-2763 (1999).

104 Wasner, G., Schattschneider, J., Binder, A. & Baron, R. Topical menthol—a human model for cold pain by activation and sensitization of C nociceptors. *Brain* **127**, 1159-1171 (2004).

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Acknowledgements:

We thank all tissue donors and their families for the generous donations. The new knowledge pieces generated from this study would not be possible without their precious gifts. We thank NDRI for helping with human DRG procurement, Dr. Isaac Chen, Ms. Arianna Unger, and other members at the UPenn (University of Pennsylvania) department of neurosurgery for helping with patient DRG procurement, the UPenn Skin Biology and Disease Resource-based Center (SBDRC) for helping with laser capture microdissection, and the Children's Hospital of Philadelphia (CHOP) Center for Applied Genomics (CAG) for helping with sequencing. We thank Dr. Yulong Li for his encouragement in starting this new research direction and all lab members in the Luo and Ma labs for their helps in conducting this project. This project is supported by an internal funding from UPenn to Dr. Luo. Dr. Luo is also supported by NIH fundings (R01 (R01NS083702), R34 (NS118411), U01 (EY034681), and P30 (AR069589)). Dr. Ernfors is supported by the Swedish Research Council (2019-00761), Knut and Alice Wallenbergs Foundation, ERC advanced grant (PainCells 740491). Dr. Olausson is supported by Knut and Alice Wallenberg Foundation (grants KAW 2019.0047 and KAW 2019.0487). Dr. Nagi is supported by the Swedish Research Council (2021-03054), Swedish Medical Society (SLS), and ALF Grants, Region Östergötland.

Materials and Methods:

Human DRGs and subjects

Human DRG tissues were procured from National Disease Research Interchange (NDRI). The research application was approved by the NDRI Feasibility Committee (RLUW1 01). Six DRGs between T11 to L5 of three human donors aged from 23- 61 years were used in this study. The dissected DRG tissues from human donors were immediately imbedded in OCT, shipped to the

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Luo lab on dry ice and stored in -80 ℃ until use. The information of DRGs and dis-identified donor is summarized in Supplementary Table 1. As determined by the University of Pennsylvania IRB, this study was exempted from the human subject requirements.

In vivo recordings of peripheral sensory afferents of healthy human subjects were performed at Linköping University, Sweden. These subjects were recruited through social media. All participants provided written informed consent before the start of the experiment. The study was approved by the ethics committee of Linköping University (dnr 2020-04426) and complied with the revised Declaration of Helsinki.

Laser capture microdissection of human DRG neurons

The human DRGs imbedded in OCT were cryosectioned (Cryostat Leica Cm1950) into 20 μm sections and mounted onto Arcturus PEN Membrane Frame Slides (Applied biosystems, LCM0521). One of every five consecutive sections was collected for laser capture microdissection to avoid repeated dissection of the soma from the same neuron in different sections. The slides were stored in -80 ℃ until further use.

On the day of laser capture microdissection, the slides were transferred to the SBDRC laser capture microdissection (LCM) core on dry ice. Before dissection, the section was briefly stained with RNase free Arcturus™ Histogene™ staining solution (Applied biosystems, 12241-05) for better visualization of neuronal soma: 70% cold EtOH for 30s; Histogene™ staining for 10s; 70% cold EtOH for 30s; 95% cold EtOH for 30s; 100% cold EtOH for 30s and air dry for 2min. Then, the slide was put onto laser capture microdissection microscope system (Leica LMD6) for the

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

neuronal soma dissection. The laser was calibrated, and the laser intensity was adjusted to achieve best dissection efficiency. The dissected individual neuronal soma was collected in the cap of a 200 µl PCR tube containing 4 µl lysis buffer[26]. The sequencing library was generated following Smart-seq2 workflow[26]. The libraries passing through all quality controls were selected for the final sequencing.

Sequencing and sequence alignment

The libraries were pooled together (384 samples for one batch) and sequenced on NovaSeq 6000 platform by the Children's Hospital of Philadelphia (CHOP) Center for Applied Genomics (CAG). Raw sequencing data was demultiplexed with bcl2fastq2 v.2.20 (Illumina) followed by Tn5 transposon adapter sequences trimming with Cutadapt[96]. The processed reads were then aligned to human genome (GRCh38 GENCODE as the reference genome, and GRCh38.104. GTF as the annotation) using STAR v.2.7.9a49[97]. STAR quantMode GeneCounts was used to quantify unique mapped reads to gene counts.

Analysis of single soma RNA-seq data of human DRG neurons using Seurat and R

R (version 4.1.2) and Seurat (version 4.0.5) were used for the single-cell RNA-seq analysis. Six objects were created from the individual biological replicates. The data were normalized (NormalizeData) after which 4500 most variable features were selected (FindVariableFeatures). To mitigate batch effects between replicates, we used Seurat's integrated analysis approach that transforms datasets into a shared space using pairwise correspondences (or "anchors")[20]. Anchors were first defined using FindIntegrationAnchors (dims = 1:30) and the data were then integrated (IntegrateData) and scaled (ScaleData), followed by principal component analysis (PCA) (RunPCA, npcs = 50). For clustering, the final parameters were: RunUMAP,

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

reduction = pca, dims = 1:25; FindNeighbors, reduction = pca, dims = 1:25; FindClusters, resolution = 3.4. Highly similar clusters without clearly distinguishable markers were merged to produce the final 16 clusters.

Conos analysis of single soma RNA-seq data of human DRG neurons

For Conos[29] analysis, human DRG neurons were integrated using CCA space $buildGraph(k = 8, k.self=3, space = "CCA", ncomps=30, n.odgenes=2000, verbose=TRUE, snn=T, snn.k = 10). For human single soma and human single-nucleus dataset, co-integration was performed as $buildGraph(k = 8, k.self=3, space = "CCA", n.odgenes=2000, verbose=TRUE, snn=T, snn.k = 10). For Conos co-clustering mouse (Sharms) dataset was downsampled to max 300 cells per cluster, co-integration was performed as $buildGraph(k = 8, k.self=3, space = "CCA", n.odgenes=2000, verbose=TRUE, snn=T,snn.k = 10). Macaque (Kupari, SmartSeq2 dataset) was used for interspecies analysis. For Conos co-clustering macaque/human graph was built as $buildGraph(k = 4, k.self=3, space = "CCA", ncomps = 30, n.odgenes=2000, snn=F, snn.k = 10), For all UMAP plots in Conos graphs were embeded as: $embedGraph(method = "UMAP", spread=15, seed = 3). Label propagation ($propagateLabels) was run using method "diffusion".

RNAScope, microscopy, imaging, and quantification

OCT embedded freshly dissected human lumbar or thoracic DRG tissues were cryosectioned at 20 µm thickness and mounted on glass slides. The slides were stored in −80 °C to preserve RNA integrity until use. RNAscope Fluorescent Multiplex Reagent Kit and RNASCOPE probes for the targeted genes (Advanced Cell Diagnostics Inc.) were used for *in situ* hybridization. RNAscope *in situ* hybridization was performed in accordance with the manufacturer's

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

instructions. In brief, fresh frozen human DRG sections were fixed, dehydrated, and treated with protease. The sections were then hybridized with the respective target probe for 2 hr at 40°C, followed by two to three rounds of signal amplification. The sections were then mounted under coverslips, sealed with nail polish, and stored in the dark at 4°C until imaged. A Leica SP5 confocal microscope was used to capture images and ImageJ was used for image analysis. In some DRG neurons, accumulation of lipofuscin in part of cells caused strong autofluorescence in all channels. These signals were considered as non-specific background (labeled by asterisk) were excluded for analysis. (See Fig. S10 for examples).

Cell type purity calculation, probabilistic similarity estimation across cell types and the data integration across species

The assessment of cell type purity, the probabilistic similarity, and cell-type integration across species are performed using packages in a machine learning based single-cell analysis toolkit-scCAMEL, released separately at https://sccamel.readthedocs.io/.

Cell type purity calculation

Shortly, we use Shannon's entropy to evaluate the purity of a cell cluster. Top 10 principal components were calculated. For an individual cluster, the loading values of every cell in each principal component were normalized to the total value after min-max scaling. The normalized loading value of each cell in a principal component presents the expected value operator for that principal component. Besides, the information content was generated from the minus logarithm of the loading value. We hold the assumption that the distribution is discrete. Thus, we summarized a cell cluster's entropy in each principal component from all cells' entropies. Next, we calculated the product of the entropies from all calculated principal components. The product

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

raised to the power of 0.1 (multiplicative inverse of the number of principal components), and the result represents the purity of this cluster.

Probabilistic similarity estimation across cell types

The calculation of cell-type probabilistic score has been described in SWAPLINE package[98]. Briefly, a vanilla neural network model was built for cell-type classification. To train the model, we removed the cell cycle–related genes and then computed the most variable features. In addition, we ranked the marker genes for each cell type by two heuristics for the cell-type specificity of both fold change and enrichment score change. Subsequently, the ranked marker genes and the most variable genes were merged, log-transformed, and scaled by min-max normalization for learning models. The frame of the neural network model and the parameters have been described in the SWAPLINE package. The learning accuracy of the neural-network classifier was inspected against epoch numbers and was estimated by k-fold cross-validation (k = 3). The learning rate and learning epochs were selected according to the maximum point of the learning curve reaching the accuracy plateaus. Data were visualized in the radar plot, which consists of a sequence of equiangular polygon spokes with the distal vertex indicating each reference cell type. The position of each predicting cell is a linear combination of the probabilistic score against all reference cell types and then visualized as the relative position to all polygon vertices. The probabilistic scores from mouse and macaque species against human reference were visualized in violin plot.

Data integration across species

For the integration task, we applied interpretable neural-network learning. First, we took one dataset from the dataset pool. We trained a neural-network classifier by learning the

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

transcriptional features of each cell-types in this dataset and then calculated the trained cells' probabilistic scores against all cell-types. Subsequently, we used all other datasets as query datasets and calculated the probabilistic score of every cell in each query dataset via the trained classifier. Then, we took another dataset from the dataset pool and repeated the training and prediction. We repeat the training and prediction till every dataset has been used as a training reference for the predictions. Here, we consider that the probabilistic score of each cell reflects the weighted gene patterns representing each trained cell-type. Thus, we merged the probabilistic scores of all cells from all trained and predicted datasets for the principal component analysis. The most significant principal components were determined by the elbow method and subsequently used as the latent space for further downstream analysis. The tree plot was constructed with the parameter of 11 principal components, 90 nearest neighbors, and correlation metric. The trained cell-type similarity was calculated with the correlation distance and the average/UPGMA linkage and visualized in the hierarchical heatmap.

In parallel, we normalized the gene expression by interpretable learning. We transformed the gene symbols of each species into the nomenclature in Homo sapiens. We estimated the features' weights in each reference cell-type by using the DeepLift algorithm[99]. The gene expression of each cell that has been learned or predicted in one trained reference dataset, was inferred by the matrix multiplication between the features' weights and the cell-type probabilistic scores. And the final gene*cell expression matrix was calculated by the average of non-empty values across all datasets. Using this normalized expression matrix, we enriched the mostly co-expressed genes via spearman correlation. These co-expressed genes were used for inferring the TF associated gene patterns via a modified GENIE3, as described in[100,101]. The result was visualized as a hierarchical heatmap.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

In vivo electrophysiological recording of human peripheral sensory fibers

Single-unit axonal recordings (microneurography) were performed from the right posterior antebrachial cutaneous, radial, or superficial peroneal nerve of 41 healthy participants (19 males and 22 females; 19 to 41 years). All participants were seated comfortably in an adjustable chair with legs and arms stretched out (and hand pronated), supported with vacuum pillows, and covered in a blanket if they reported as being cold.

Under real-time ultrasound guidance (LOGIQ P9, GE Healthcare, Chicago, IL, USA), the target nerve was impaled with an insulated tungsten recording electrode (FHC Inc., Bowdoin, ME, USA). Adjacent to that, an uninsulated reference electrode was inserted just under the skin. A high-impedance preamplifier (MLT185 headstage) was attached to the skin near the recording electrode and used together with a low-noise high-gain amplifier (FE185 Neuro Amp EX, ADInstruments, Oxford, UK). Once the electrode tip was intrafascicular, single LTMRs were searched for by soft-brush stroking, and single HTMRs were searched for by coarse-brush stroking, pinching, and hair tugging in the fascicular innervation zone while making minute adjustments to the electrode position.

All recorded afferents were mechanically responsive and divided into subtypes based on the criteria used in previous studies[62,102,103]. Mechanical threshold and receptive field size were determined using Semmes-Weinstein monofilaments (nylon fiber; Aesthesio, Bioseb, Pinellas Park, FL, USA). Mechanical threshold was defined as the weakest monofilament to which the unit responded to in at least 50% of trials. Hair deflection was tested with a small pair of forceps. Care was taken to avoid skin contact during hair movement and force measurements were performed to ensure no skin/hair pulling was involved. The conduction velocity of the recorded

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

afferent was estimated from the latency responses to surface electrical stimulation of the receptive field (FE180 Stimulus Isolator, ADInstruments, Oxford, UK). Electrically and mechanically evoked spikes were compared on an expanded time scale to confirm that both responses were from the same unit. Thermal responsiveness was tested by placing a Peltier probe (7.4 x 12.2 mm, T09, QST.Lab, Strasbourg, France) onto the receptive field; after recording at least 30 s of baseline activity with the thermode in contact with the receptive field at a neutral temperature of 30°C, a series of cooling (down to 0°C) and warming (up to 50°C) stimuli were delivered at intervals of 30 s. Where needed, the thermode was mounted on a stand for better stability.

TRPV1 presence was tested by topical application of capsaicin to the receptive field (Capsina 0.075%, Bioglan AB, Malmö, Sweden). After 1 min, the skin was wiped clean, and the emergence of any spontaneous spiking activity from the recorded afferent was monitored. *TRPM8* expression was tested by placing a gauze pad onto the receptive field soaked in ethanol only (90% as control) followed by an aliquot of a solution containing menthol (400 mg of 40% L-menthol dissolved in 90% ethanol, Sigma-Aldrich, Inc., Schnelldorf, Germany[104]). The gauze pad was then covered with an adhesive film to prevent the evaporation of ethanol. In either case, after 5 minutes of application, the skin was wiped clean and the emergence of any spontaneous spiking activity from the recorded afferent was monitored. During these procedures, we documented the participants' verbal descriptions of what they felt, and if there was no obvious sensation, the procedure was repeated.

The neural activity was sampled at 20 kHz and recorded using the ADInstruments data acquisition system (LabChart software v8.1.24 and PowerLab 16/35 hardware PL3516/P,

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Oxford, UK) and subsequently exported to Spike2 (v10.13, Cambridge Electronic Design Ltd., Cambridge, UK). The recorded action potentials were carefully examined offline on an expanded time scale. Threshold crossing was used to distinguish action potentials from the noise with a signal-to-noise ratio of at least 2:1, and spike morphology was confirmed by template matching. Recordings were discarded if the analyses indicated the presence of multiple units or if the spike amplitudes were not distinct from the noise, preventing secure action potential identification. Figures were generated in GraphPad Prism (v9, GraphPad Software Inc. La Jolla, CA, USA).

Data availability

The raw and processed datasets for the single soma sequencing of human DRG neurons reported in this study will be deposited to a NIH repository once the manuscript is accepted for publication. A website containing the processed data for browsing and searching gene expression in the different human neuron types will be available once the manuscript is accepted for publication. Macaque (Kupari) data is available at

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE165569

Mouse (Zeisel) DRG data is available at

http://loom.linnarssonlab.org/clone/Mousebrain.org.level6/L6_Peripheral_sensory_neurons.loom.

Mouse (Sharma) DRG data is available at

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE139088

Code availability

Any custom code will be deposited to Github once the manuscript is accepted for publication. All analyses are based on previously published code and software.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 1. Developing a novel laser capture microdissection (LCM) based approach for single-soma deep RNA-seq of human DRG neurons

(**A**) A cartoon showing the workflow of LCM-based single-soma deep RNA-seq of human DRG neurons (left). An example of laser dissection of the soma of a human DRG neuron (right). Scale bar, 50 µm (cell) and 500 µm (tube) (**B**) UMAP plot showing the clusters of 1066 neurons. (**C-F**) Violin plots showing total number of detected genes (C), the expression of neuronal marker *SLC17A6* (D), the soma sizes (E), and the expression of *NEFH*, *PRDM12*, and *CALCA* in each cluster (F). (**G**) UMAPs showing some canonical marker gene expression in each cluster. (**H**) UMAP plot clusters with names for each cluster.

Fig. 2 Cross-species analysis of DRG neurons in human, macaque, and mouse

(**A, B**) Conos label propagation from mouse (Sharma) (**A,** combined Sharma & Usoskin nomenclature) and macaque (**B**) to human DRG neuron clusters showing the cell type correlation. For UMAPs for correspondent co-integration from which these results were inferred, see **Fig. S6. (C**) Hierarchical clustering of cell types in human and macaque. (**D**) Heatmap visualization of cross-species-conserved and species-specific transcription factor associated gene patterns across mouse, macaque, and human. Species are color coded in the right column. Names of human specific clusters in blue. Yellow boxes, conserved, green boxes, species-specific gene regulatory networks.

Figure 3. Markers and functional genes expressed in C-fiber pruriceptors/nociceptors

(A) Dot plot showing the expression of marker or functional genes in C-fiber pruriceptors/nociceptors. (**B1- F2**) Marker genes for specific labeling of each cluster and validation by multicolor *in situ* hybridization for hNP1 and hNP2 (B1-C2), hPEP.SST(D1-F2). The fluorescent images at each row show one example human DRG neuron (cell body outlined

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

by the white dashed line), and the circle chart at the right-side show quantification. The arcs indicate the percentage of marker gene positive cells over all sampled DRG neurons. The sector shaded areas indicate the approximate percentage of each cell type over all neuron population. B2 (188 cells, N=2), C2 (199 cells, N=2), D2 (100 cells, N=1), E2 (220 cells, N=2), F2 (220 cells, N=2). Scale bar, 50 µm.

Figure 4. Markers and functional genes expressed in C-fiber thermoreceptors and nociceptors

(A) Dot plot showing the expression of marker or functional genes in C-fiber thermoreceptors and nociceptors. (**B1- D2**) Marker genes for specific labeling of each cluster and validation by multicolor *in situ* hybridization for hTRPM8 (B1, B2), and hTRPV1/A1.1-hTRPV1/A1.2 (C1-D2). The arcs indicate the percentage of marker gene positive cells over all sampled DRG neurons. The sector shaded areas indicate the approximate percentage of each cell type over all neuron population. B2 (156 cells, N=2), C2 (202 cells, N=2), D2 (192 cells, N=2). Scale bar, 50 µm. Scale bar, 50 µm.

Figure 5. Markers and functional genes expressed in A-fiber nociceptors

(A) Dot plot showing the expression of marker or functional genes in A-nociceptors. (**B1- E2**) Marker genes for specific labeling of each cluster and validation by multicolor *in situ* hybridization for hPEP.PIEZO^{h} (B1, B2), hPEP.KIT (C1, C2), hPEP.CHRNA7 (D1, D2) and hPEP.NTRK3 (E1, E2). The arcs indicate the percentage of marker gene positive cells over all sampled DRG neurons. The sector shaded areas indicate the approximate percentage of each cell type over all neuron population. B2 (165 cells, N=2), C2 (173 cells, N=2), D2 (196 cells, N=2), E2 (191 cells, N=2). Scale bar, 50 µm.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 6. Markers and functional genes expressed in C- and A-LTMRs

(A) Dot plot showing the expression of marker or functional genes in C- and A-LTMRs. (**B1- F2**) Marker genes for specific labeling of each cell type and validation by multicolor *in situ* hybridization for hC.LTMR (B1, B2), hAδ.LTMR (C1, C2), hAβ.LTMR (D1, D2), hPropr (E1, E2), and hATF3 (F1, F2). The arcs indicate the percentage of marker gene positive cells over all sampled DRG neurons. The sector shaded areas indicate the approximate percentage of each cell type over all neuron population. B2 (205 cells, N=2), C2 (183 cells, N=2), D2 (188 cells, N=2), E2 (198 cells, N=2), F2 (202 cells, N-4=2). Asterisk (*) in E2 indicates non-specific lipofuscin autofluorescence. Scale bar, 50 µm.

Figure 7. Single-unit microneurography recordings of human peripheral sensory fibers reveal the novel physiological properties predicted by molecule profiles

(**A**) Location of receptive fields of single afferents from superficial peroneal (S. peroneal), posterior antebrachial cutaneous (PABCN), and radial nerve recordings (n = 47). (**B**) Mechanical (monofilament) threshold distribution of HTMRs and LTMRs in the recorded samples. (**C**) Individual and mean (±SEM) conduction velocities of different HTMR and LTMR types to surface electrical stimulation from upper and lower limbs (Field-LTMR: 40.3 ± 4.2 m/s, n=2; SA1-LTMR: 44.9 ± 2.6 m/s, n=3; SA2-LTMR: 44.9 ± 1.2 m/s, n=3; A-HTMR cooling-: 50.6 ± 4.8 m/s, n=5; A-HTMR cooling+: 48.9 ± 5.0 m/s, n=3; C-LTMR: 1.0 ± 0.05 m/s, n=8; C-HTMR: 0.7 ± 0.08 m/s, n=5). (**D**) Spike activity of a likely singular hPEP.KIT unit (A-HTMR cooling+) to repeated stimulations of the receptive field, superimposed on an expanded time scale. (**E-H**) Responses of an A-HTMR cooling+ unit to soft brushing (E), coarse brushing (F), cooling (G),

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and heating (H). (**I**) Spike activity of a hC.LTMR unit to repeated stimulations of the receptive field, superimposed on an expanded time scale. (**J-M**) Responses of a hC.LTMR unit to soft brushing (J), hair movement (K), cooling (L), and heating (M).

Supplementary figure 1. Isolation of human DRG neuronal soma by LCM

(**A**) A workflow showing steps for brief staining of human DRG section before LCM. **(B)** An image showing the human DRG cryosections mounted on the LCM slide after brief staining. (**C**) A reprehensive image showing the stained human DRG section. Scale bar, 400 μm (**D, E**) Size distribution of all DRG neurons (D) and dissected neuronal soma (E).

Supplementary figure 2. LCM RNA-seq statistics, clustering, and additional marker gene expression

(**A-C**) Violin plots showing total number of detected genes in different batches (A), donors (B), and DRG levels (C). (**D-E**) UMAP plots showing the contribution of individual batches (D) and donors (E) to each cluster. (**F-G**) Violin plots showing the expression of pan-neuronal markers *SYP* and *UCHL1*. (**H**) UMAPs showing expression pattern of additional canonical marker genes in each cluster.

Supplementary figure 3. Heatmap of marker genes in human DRG clusters

A heatmap showing the ten most specific genes for each cluster.

Supplementary figure 4. Validation of human DRG neuron clustering

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(**A**) Clustering of human DRG neurons by Conos. (**B**) The accuracy of neural-network classifier in learning human DRG neurons was visualized as learning curve. The Y-axis represents the learning accuracy, and the X-axis represents the training epoch numbers. Following the training epochs, the maximum accuracy plateaus of this learning curve reaching ~88%. (**C**) A radar-plot showing cell-type fractions of the human DRG neurons from the neural-network scoring module trained with human clusters. Each dot represents one cell, and the color coding is based on unique cell clusters. (**D**) Percentage bar chart visualization, the consistency of assigned human DRG neurons by the neural-network scoring module. Red indicating the ratio of cells that were not consistently assigned to their defined cell-types, and blue indicating the ratio of consistently assigned cells. (**E**) Cell type purity within each cluster calculated by normalized entropy, Y-axis indicating the estimated purity score. (**F**) hierarchical heatmap visualization of the probabilistic similarity across cell types, color from dark blue to yellow indicating the similarity score from low to high.

Supplementary figure 5. Co-clustering with single-nucleus RNA-seq of human DRG neuron

(**A-B**) Co-clustering of human single soma RNA-seq and human single-nucleus RNA-seq (Nguyen). (**C**) Label transfer from human single soma RNA-seq to human single-nucleus RNA-seq showing the cell type correlation between the two datasets.

Supplementary figure 6. Cross-species comparison of sensory neuron types among human, macaque, and mouse

(**A**) Probabalistic neural network probability scores of mouse (Sharma) DRG neuron types tested on human trained module. (**B**) Probabalistic neural network probability scores of

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

macaque (Kupari, SmartSeq2 dataset) DRG neuron types tested on human trained module. (**C-F**) UMAPs of Conos co-clustering of human and mouse (C-D) or macaque (E-F) DRG neurons. These co-clustering were used for label propagation inferences shown on Fig. 2A & 2B. Each pair UMAPs were plotted with the same coordinates, the upper plot showing human clusters in color and the lower plot showing mouse or macaque, respectively. Small black crosses depict uncolored co-clustered species. For mouse data Sharma cluster nomenclature is used.

Supplementary figure 7. Comparison of marker gene expression across species

Dot plots showing ten specific marker genes in human DRG neuron clusters expressed in human (**A1, A2**), macaque (Kupari) (**B1, B2**) and mouse (Sharma) (**C1, C2**) DRG neuron datasets. The black boxes highlighted the corresponding cell types based on the label transfer and neural-network scoring analysis.

Supplementary figure 8. Expression of additional functional or marker genes in human DRG neurons

Violin plot showing the expression of additional functional or marker genes in human DRG neuron clusters.

Supplementary figure 9. Expression of functional or marker genes in mouse and macaque DRG neurons

Violin plot showing the expression of functional or marker genes in mouse (**A1-A15**, Sharma dataset[16]) and macaque (**B1-B8**, Kupari, SmartSeq2 dataset[18]) DRG neuron clusters.

Supplementary figure 10. Multiplex FISH of human DRG neuron

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

This figure demonstrates the imaging process of multiplex FISH of human DRG neuron and the non-specific lipofuscin autofluorescence. (**A**) Images of entire human DRG sections after performing multiplex FISH using probes for *IL31RA* (red), *OSMR* (green), *SST* (purple), and merged images. (**B-C**) The white box areas (B and C) on merged image in (A) are shown at higher magnification. An example of IL31RA, OSMR and SST triple positive neuron shown in (B) and an example of IL31RA and OSMR positive but SST negative neuron shown in (C). (**D**) An example of lipofuscin autofluorescence (indicated by red arrowhead in the merged image in (A)) indicated by asterisk (*) showed non-specific signals, which are excluded for analysis. Scale bar, 500 μm in (A), 50 μm in (B-D)

Supplementary figure 11. Expression of itch- and pain-sensing related genes in human DRG neurons

Violin plot showing the expression of some itch- and pain-sensing related genes in human DRG neurons (**A-O**).

Supplementary figure 12. Expression of itch receptors in human DRG neurons and identification of potential novel anti-itch targets

(A) Some known itch receptors are highly enriched in human itch populations. (**B**) Some putative itch receptors are not highly enriched in human itch populations. (**C**) Novel GPRRs, ion channels, and other genes highly enriched in human itch populations. They could be novel molecular targets for anti-itch treatment.

Supplementary figure 13. Expression of GPCRs in human DRG neurons

Heatmap showing the expression of GPCRs in human DRG neurons. Genes were ranked by average expression level in all DRG neurons.

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary figure 14. Expression of ion channels in human DRG neurons

Heatmap showing the expression of ion channels in human DRG neurons. Genes were ranked by average expression level in all DRG neurons.

Supplementary figure 15. Expression of chemokine receptors in human DRG neurons

Heatmap showing the expression of chemokine receptors in human DRG neurons. Genes were ranked by average expression level in all DRG neurons.

Supplementary figure 16. Expression of peptides in human DRG neurons

Heatmap showing the expression of peptides in human DRG neurons. Genes were ranked by average expression level in all DRG neurons.

Supplementary figure 17. Expression of *CGRP*, *TRPV1*, *PIEZO2*, *KIT*, and *TRPM8* in mouse CGRP- η and human hPEP.KIT populations

(**A, B**) Violin plot showing the expression of *KIT*, *CALCA*, *PIEZO2*, *TRPV1* and *TRPM8* in mouse GGRP-η population (A) and in human hPEP.KIT population (B). (**C**) Validation of the expression of *TRPM8*, *PIEZO2* and *CGRP* in hPEP.KIT population by multiplex FISH.

Supplementary figure 18. Responses of A-HTMR cooling+ and polymodal hC.LTMR units to temperature changes

(**A** & **B**) Responses of A-HTMR cooling+ and polymodal hC.LTMR units to temperature changes. Each data point represents the mean (± SEM) responses of 5 A-HTMR cooling+ or 5 polymodal

bioRxiv preprint doi: https://doi.org/10.1101/2023.03.17.533207; this version posted March 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

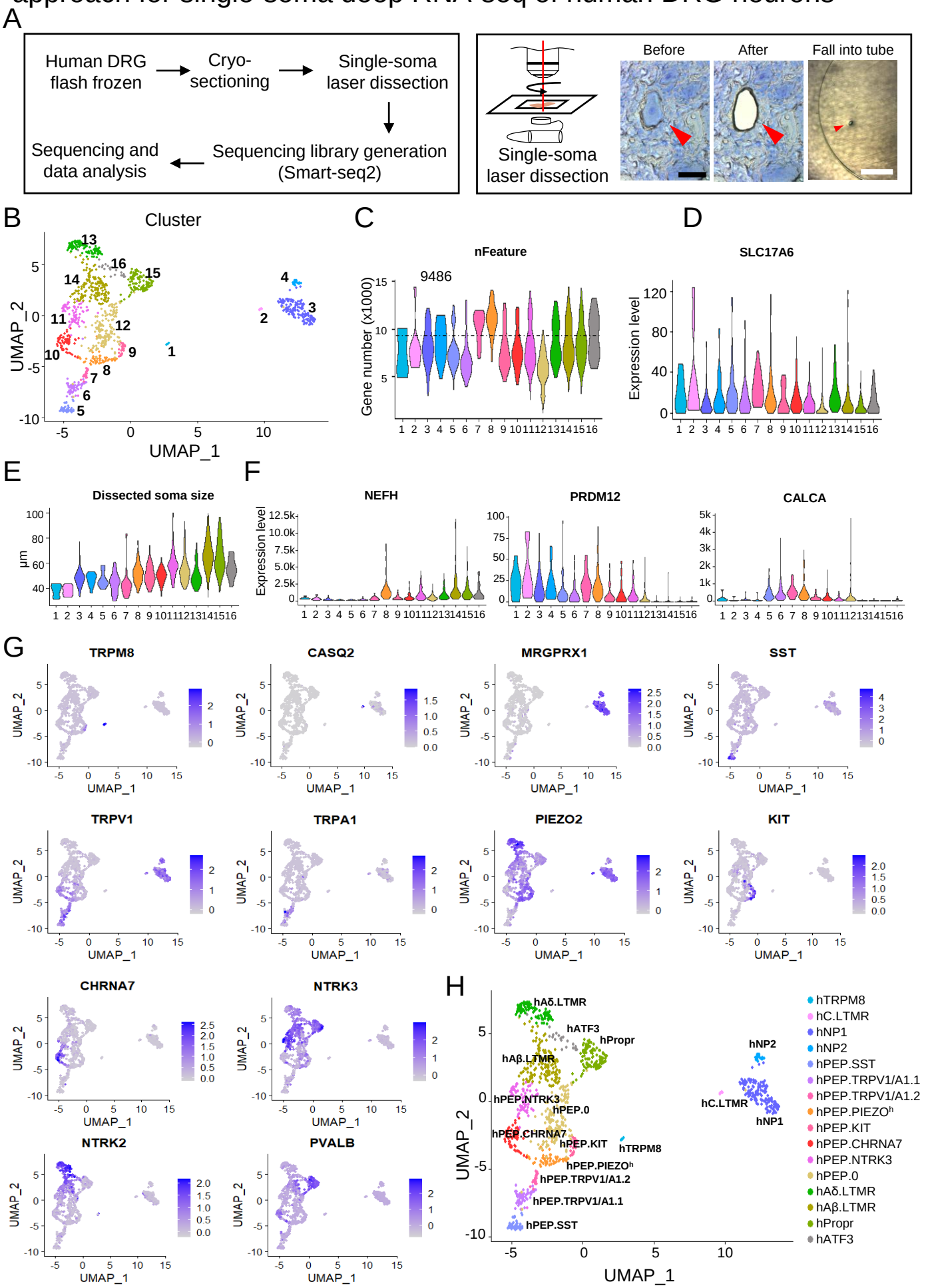
hC.LTMR units, tested in triplicate. The hC.LTMR responses were adjusted for conduction delay based on the latency taken for conduction velocity measurements. (A) The number of action potentials evoked during dynamic (30 to 0°C at 4°C per second) and sustained (0°C for 5.5 second) phases of cooling/cold stimulation is shown. (B) The number of action potentials evoked during dynamic (up to 50°C at 4°C per second) and sustained (~50°C for 6 second) phases of heating/hot stimulation is shown.

Supplementary figure 19. Expression of *TRPM8*, *TRPV1* and *PIEZO2* in hC.LTMR, and physiological recordings of human peripheral nerve fibers

(**A**) Violin plot showing the expression of *TRPM8*, *TRPV1* and *PIEZO2* in hC.LTMR population. (**B**) Validation of the expression of *TRPM8*, *TRPV1* and *PIEZO2* in the hC.LTMR population by multiplex FISH. (**C, D**) Recordings from a hC.LTMR prior to and following menthol (C) and capsaicin (D) applications to the receptive field. (**E**) Spike activity of a hC.HTMR to repeated stimulations of the receptive field, superimposed on an expanded time scale. (**F-I)** Responses of a hC.HTMR to soft brushing (F), coarse brushing (G), hair movement (H), and heating and capsaicin (I). (**J**) The effectiveness of the menthol protocol was confirmed by the emergence of cold perception and spontaneous C-fiber activity, likely from TRPM8+ cold units.

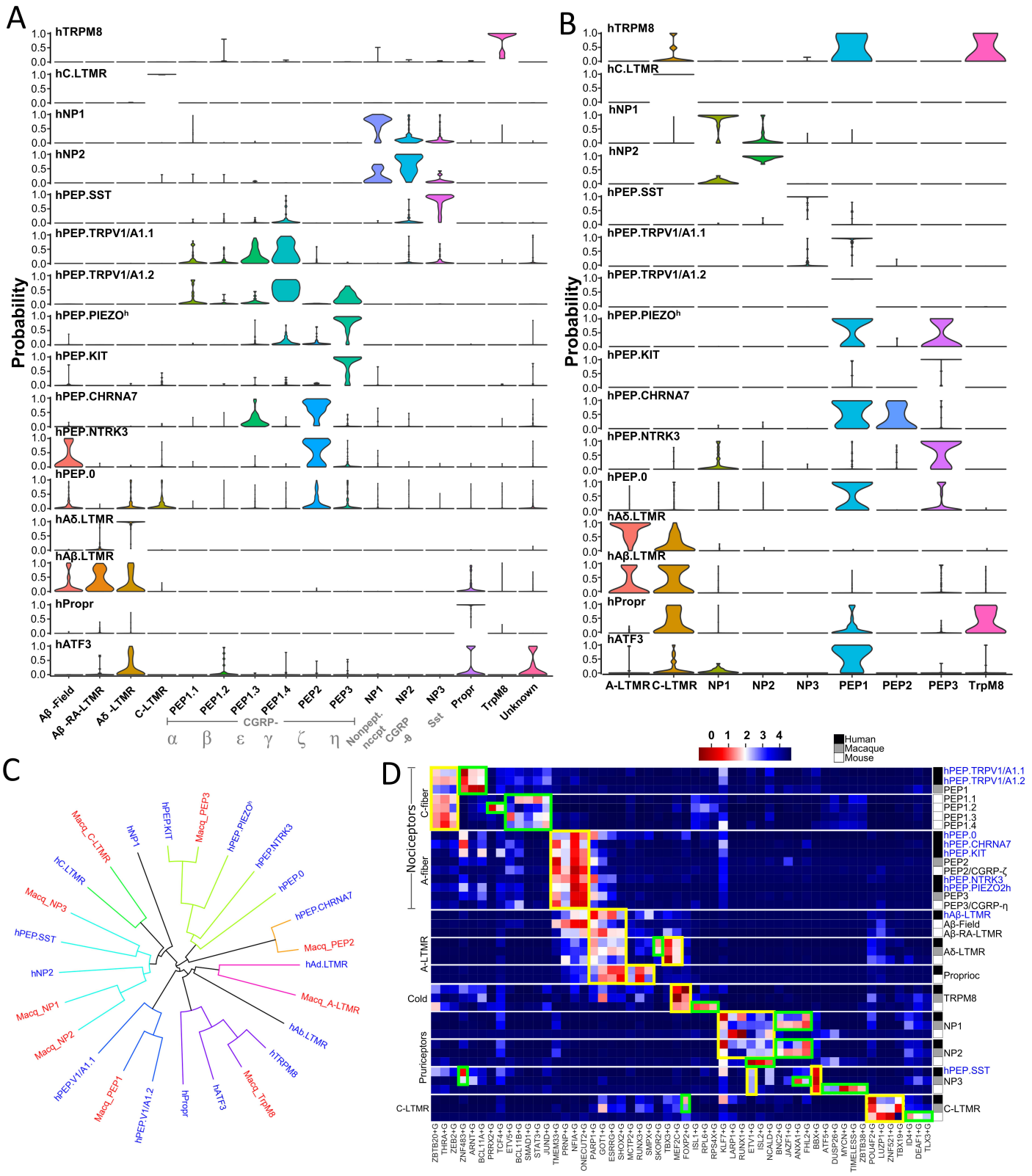
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fig.1 Developing a novel laser capture microdissection (LCM) based approach for single-soma deep RNA-seq of human DRG neurons



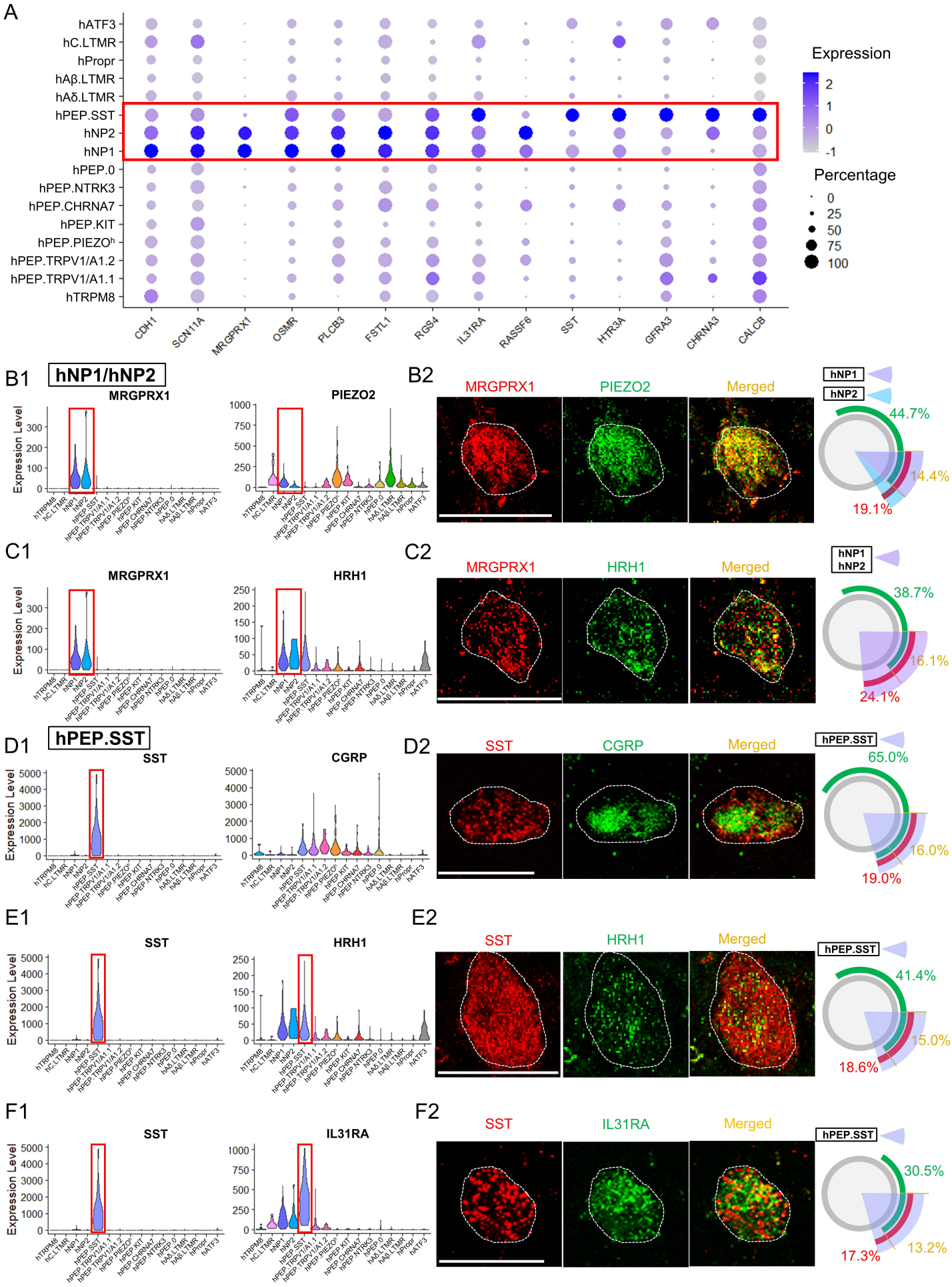
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fig. 2 Cross-species analysis of DRG neurons in human, macaque, and mouse



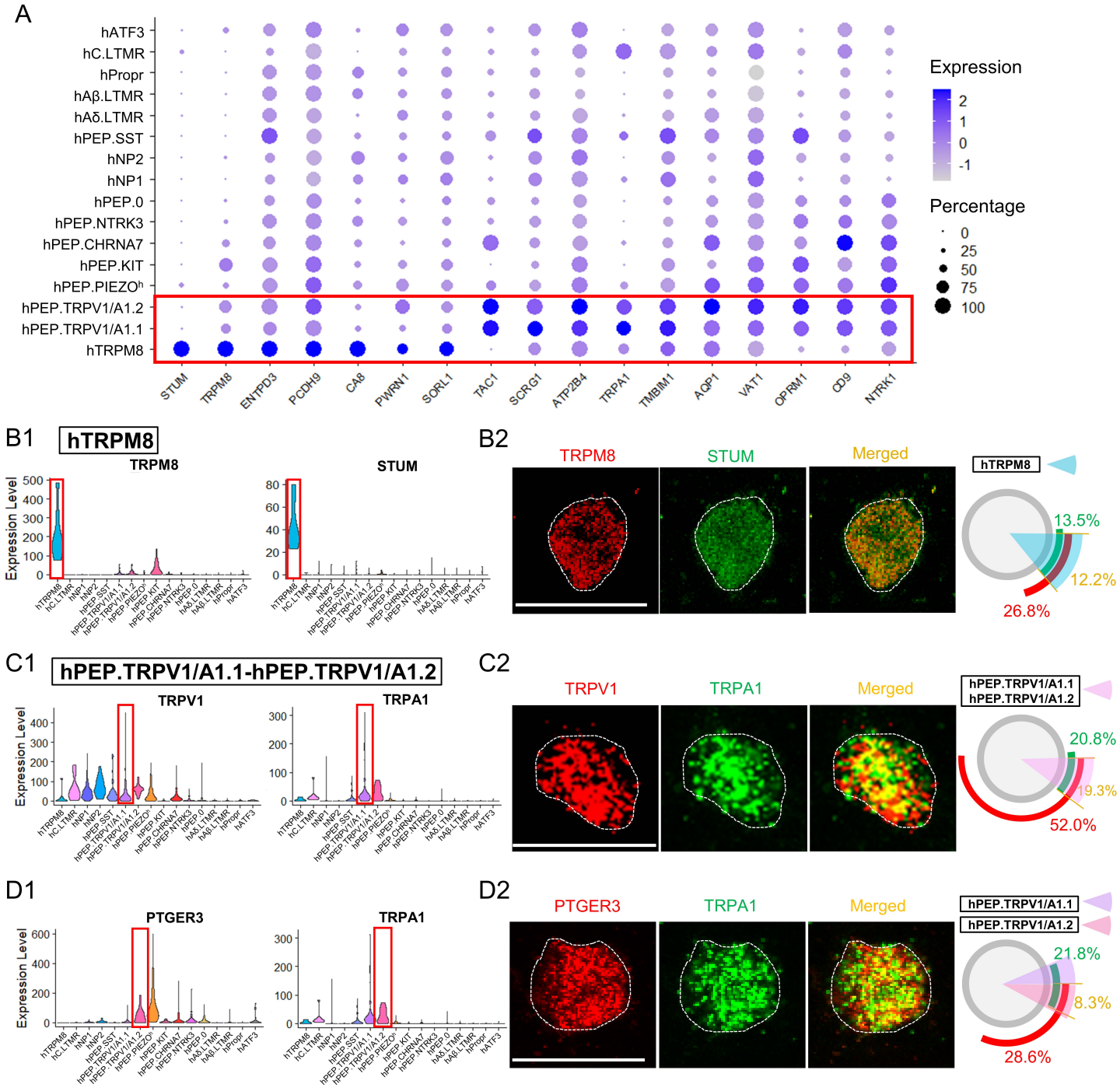
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 3. Markers and functional genes expressed in C-fiber pruriceptors/nociceptors



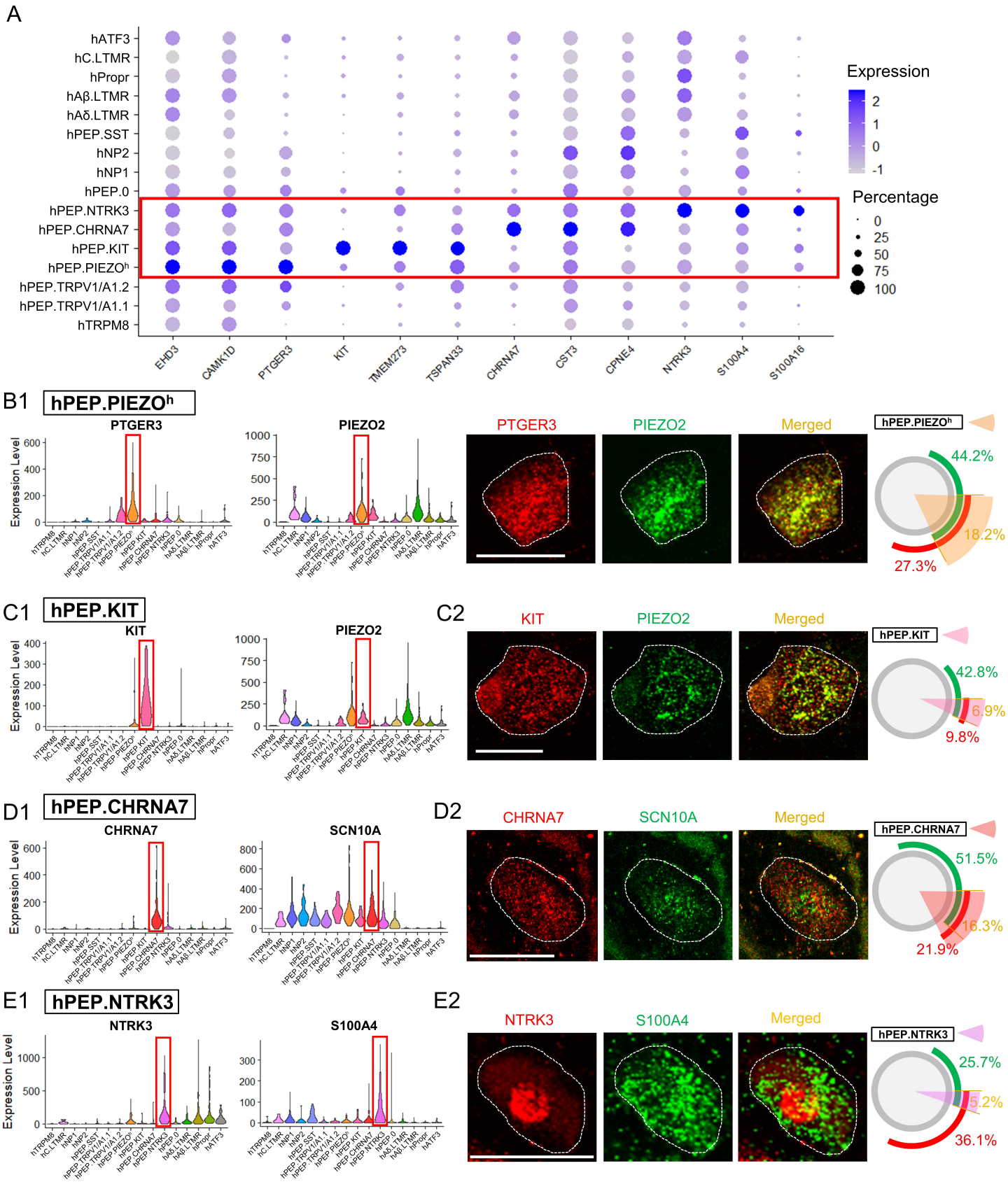
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 4. Markers and functional genes expressed in C-fiber thermoreceptors and nociceptors



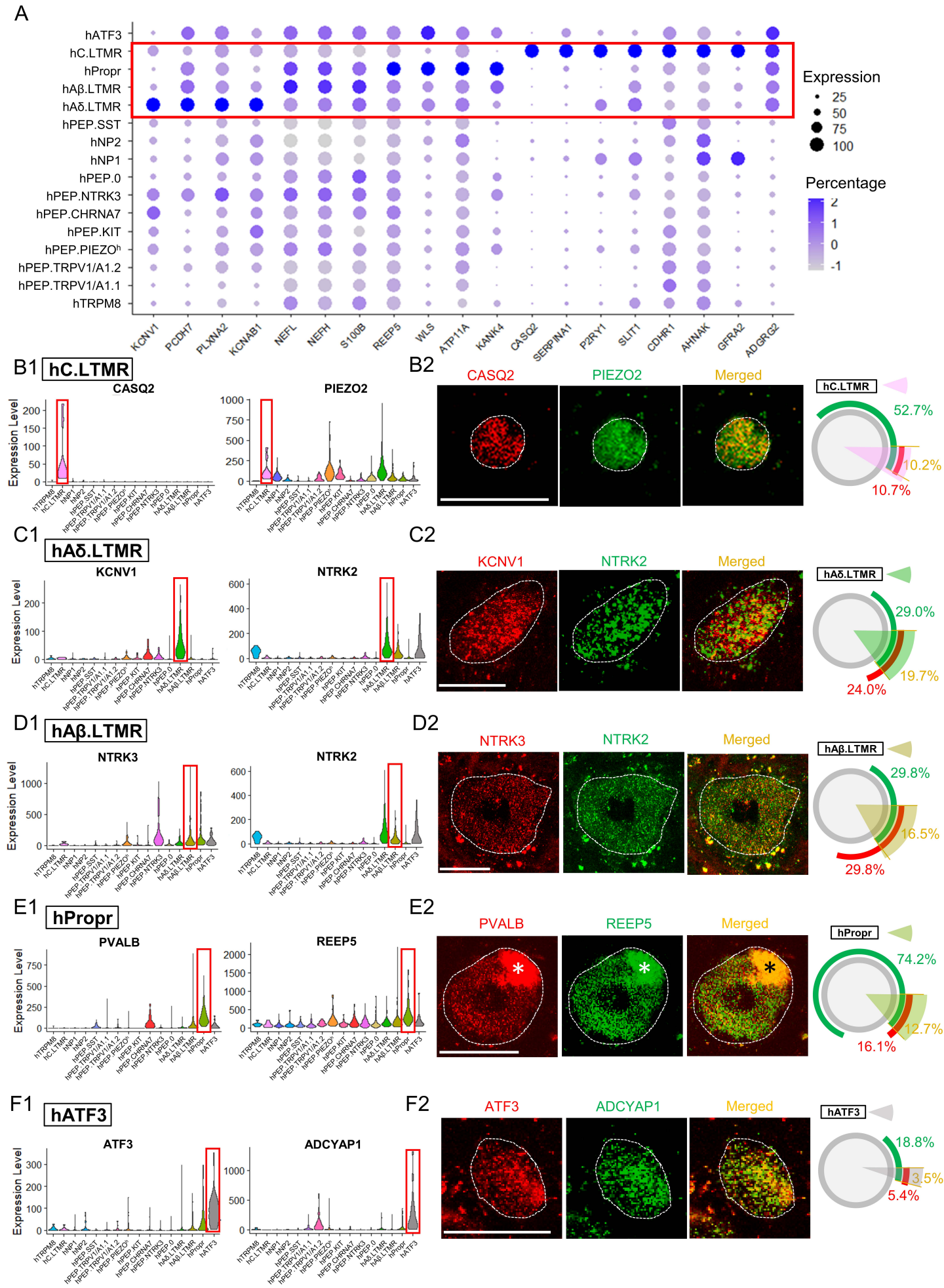
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 5. Markers and functional genes expressed in A-fiber nociceptors



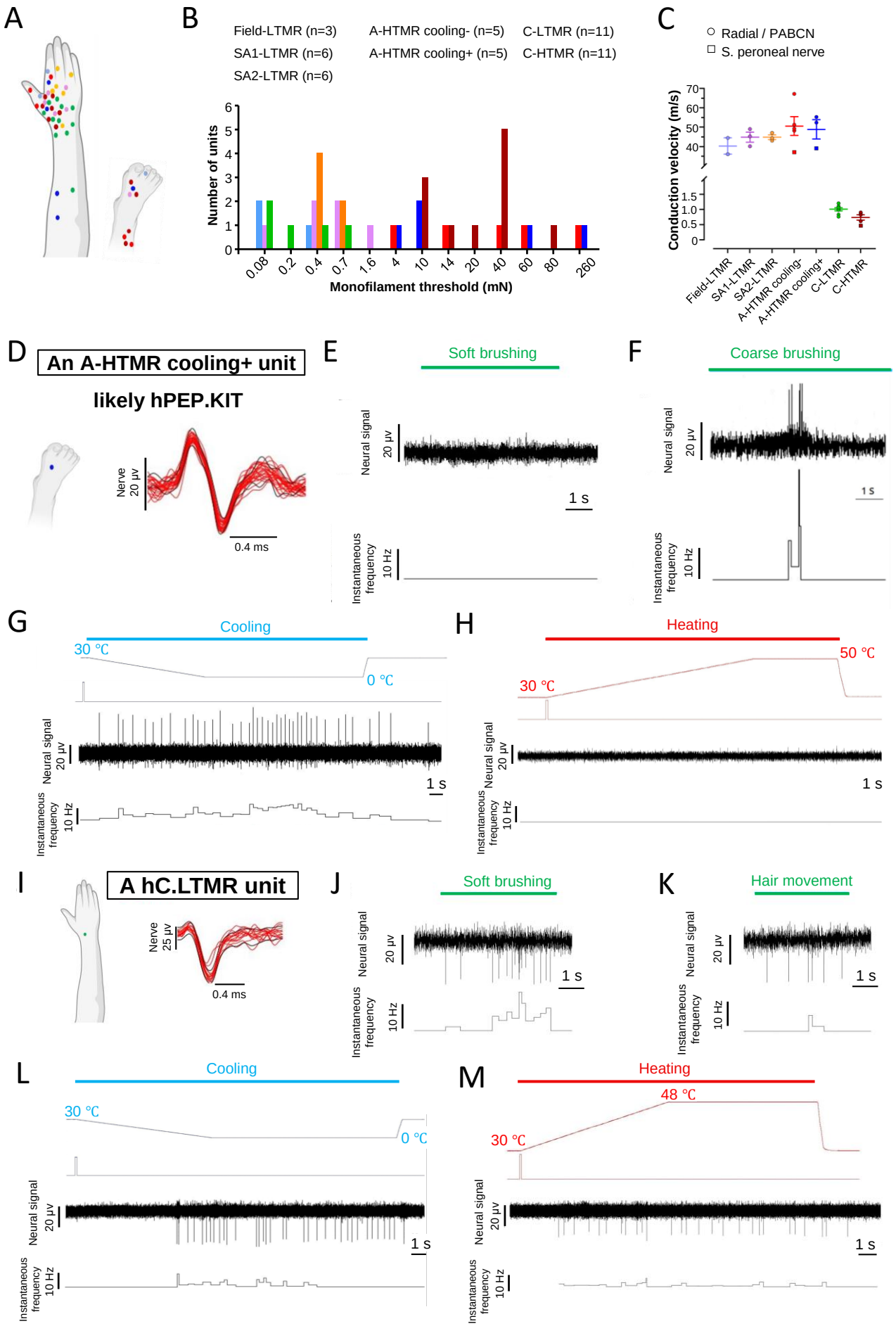
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 6. Markers and functional genes expressed in C- and A-LTMRs



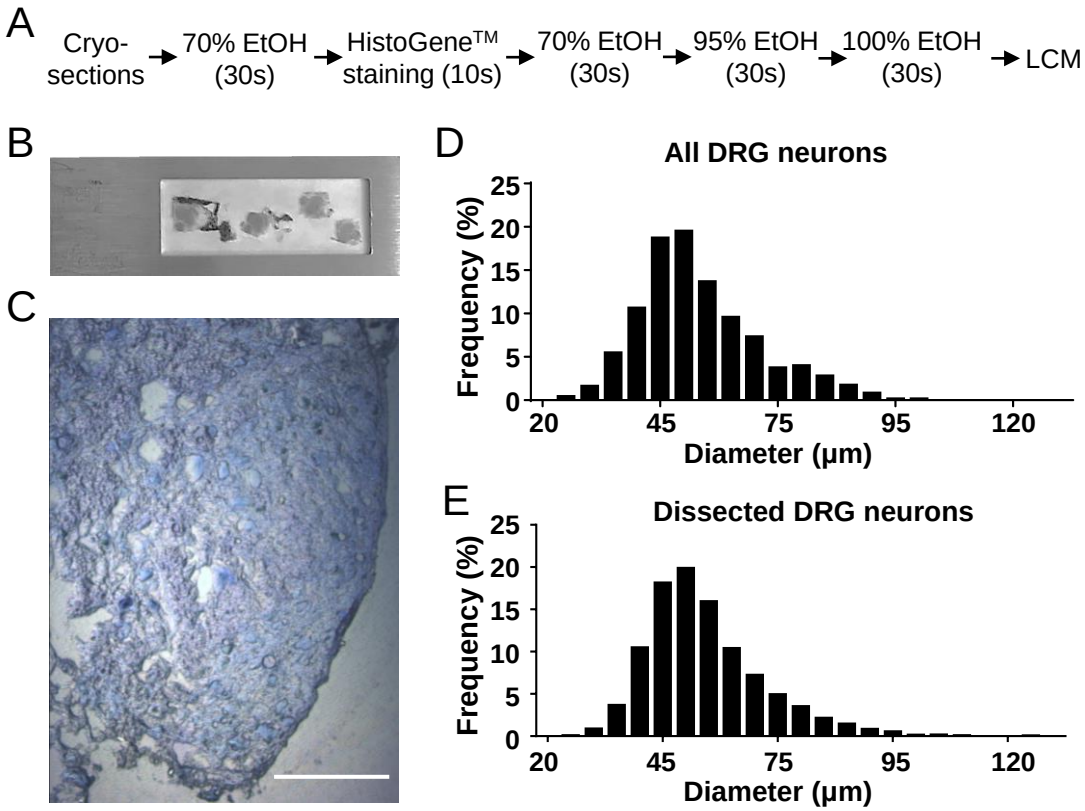
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fig. 7 Novel physiological properties of human peripheral sensory afferents predicted by their molecular profiles



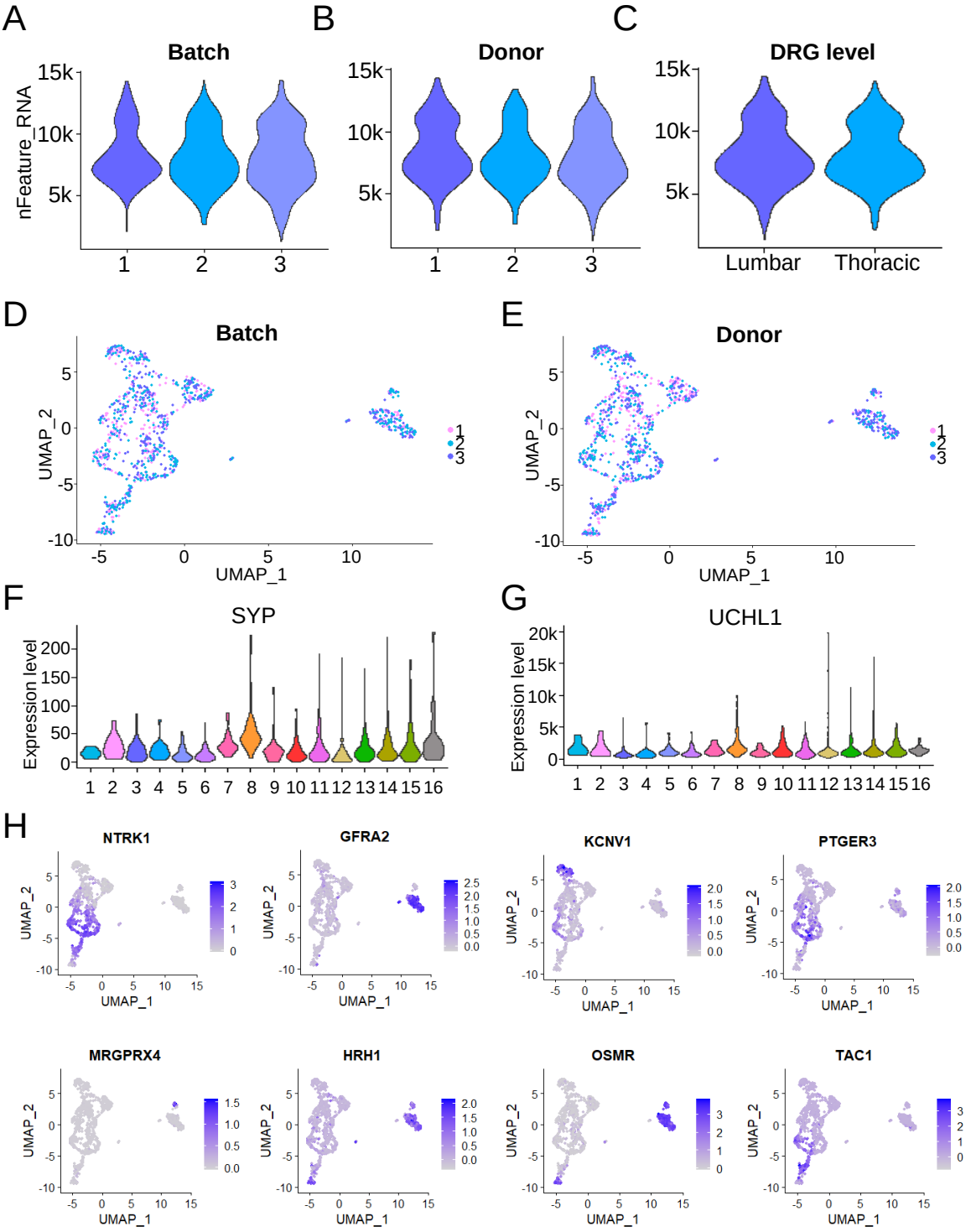
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig. 1 Isolation of human DRG neuronal soma by LCM



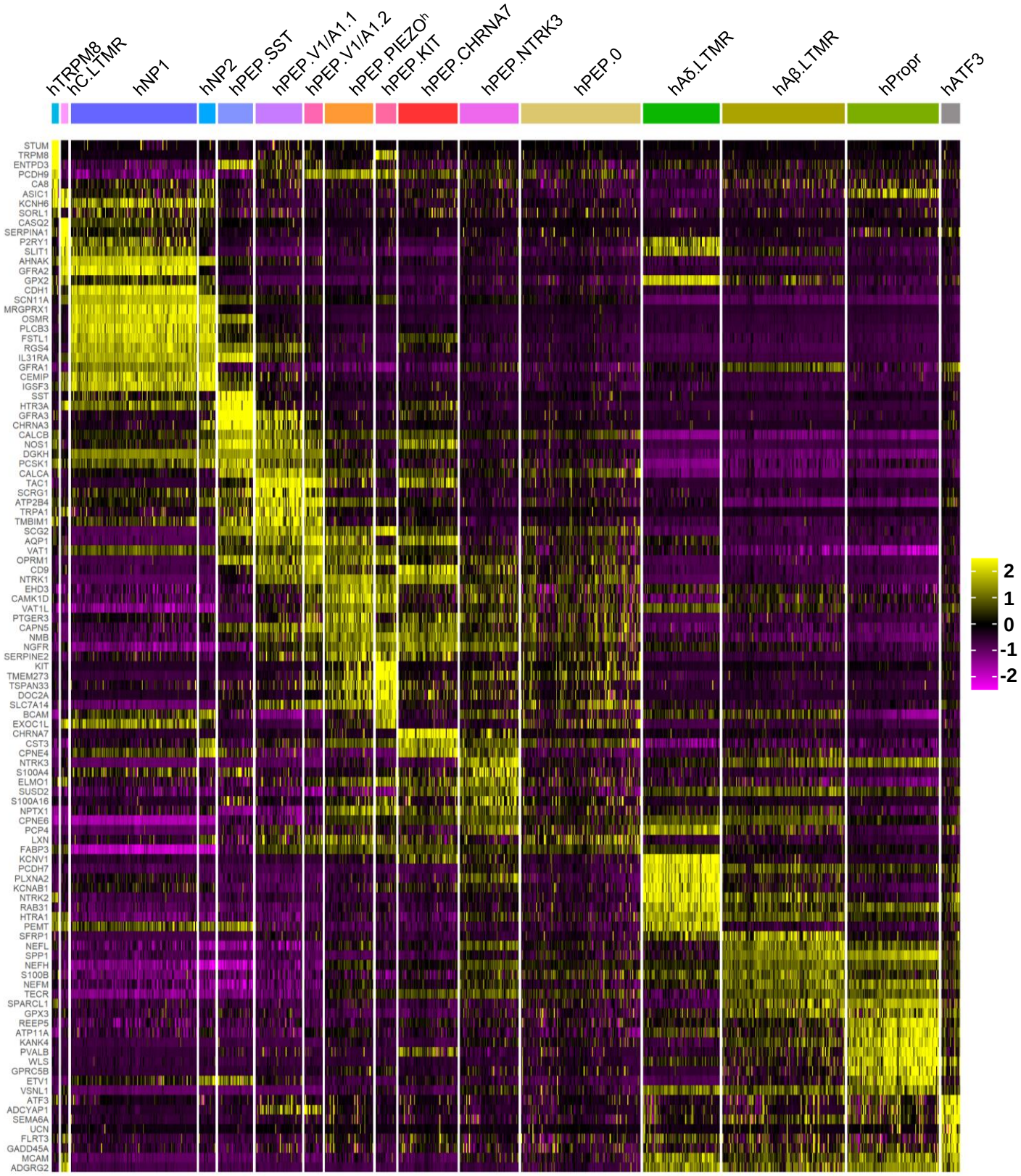
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig. 2 LCM RNA-seq statistics, clustering, and additional marker gene expression



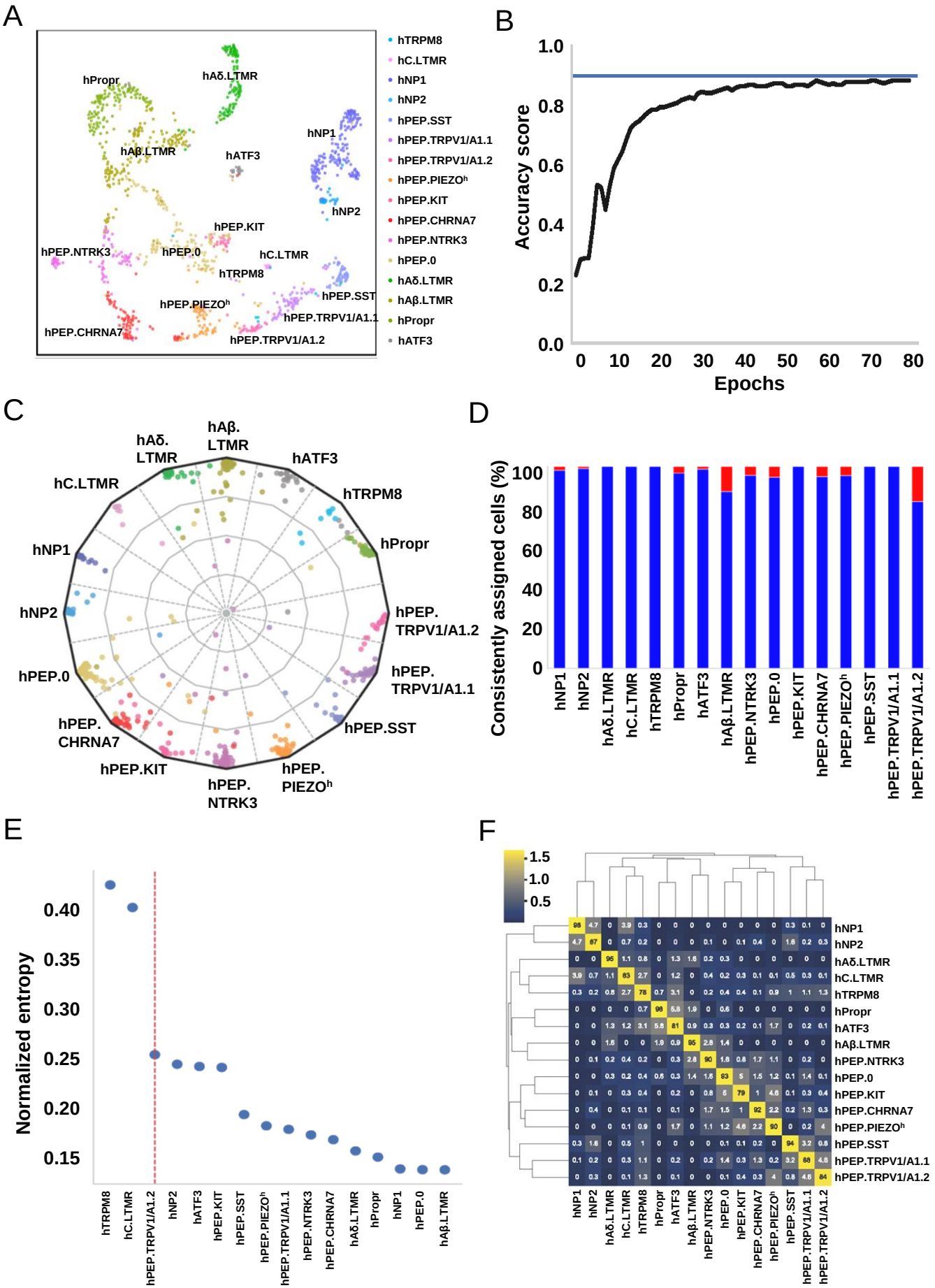
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig. 3 Heatmap of marker genes in human DRG clusters



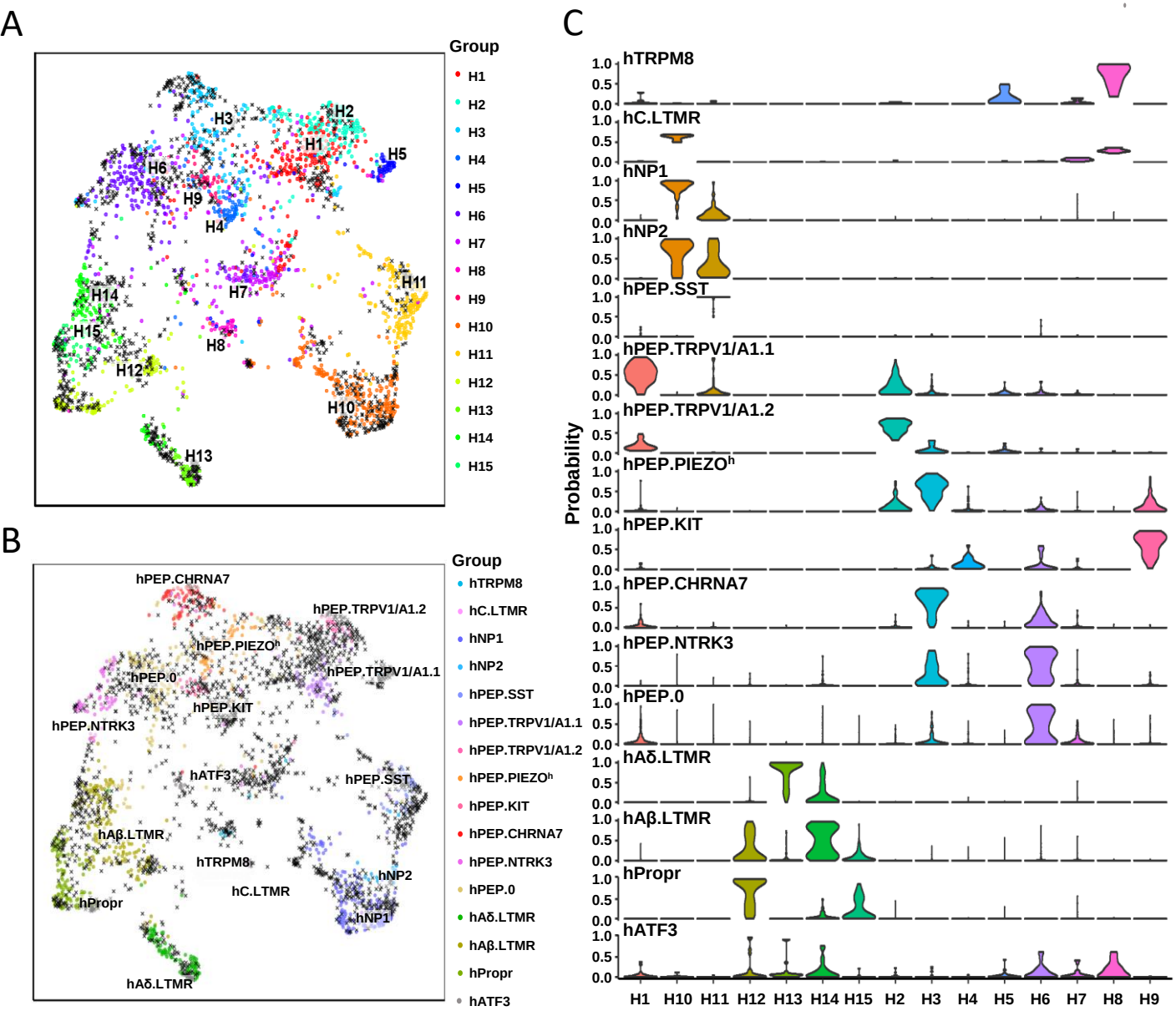
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig. 4 Validation of human DRG neuron clustering



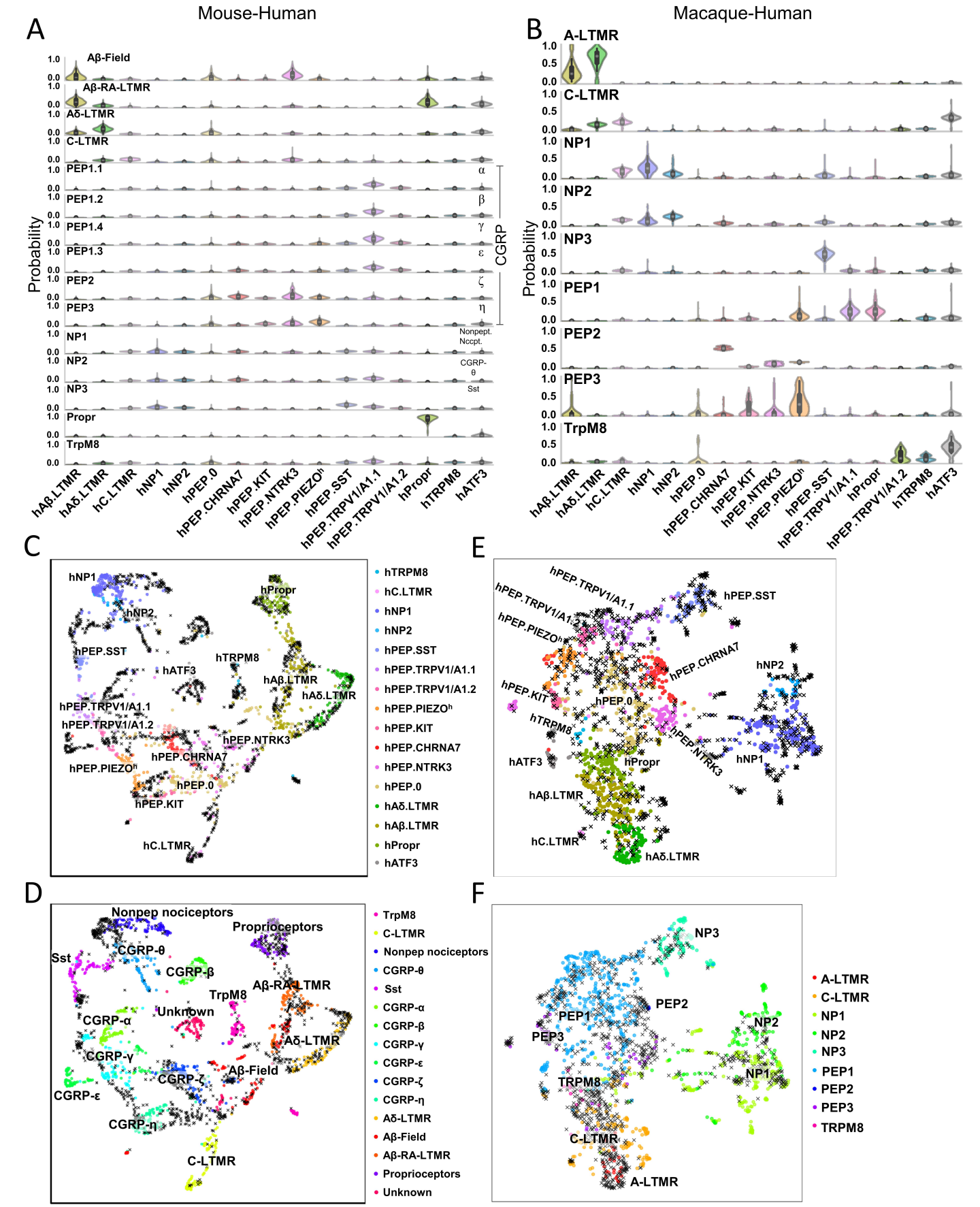
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig 5. Co-clustering with single-nucleus RNA-seq of human DRG neuron



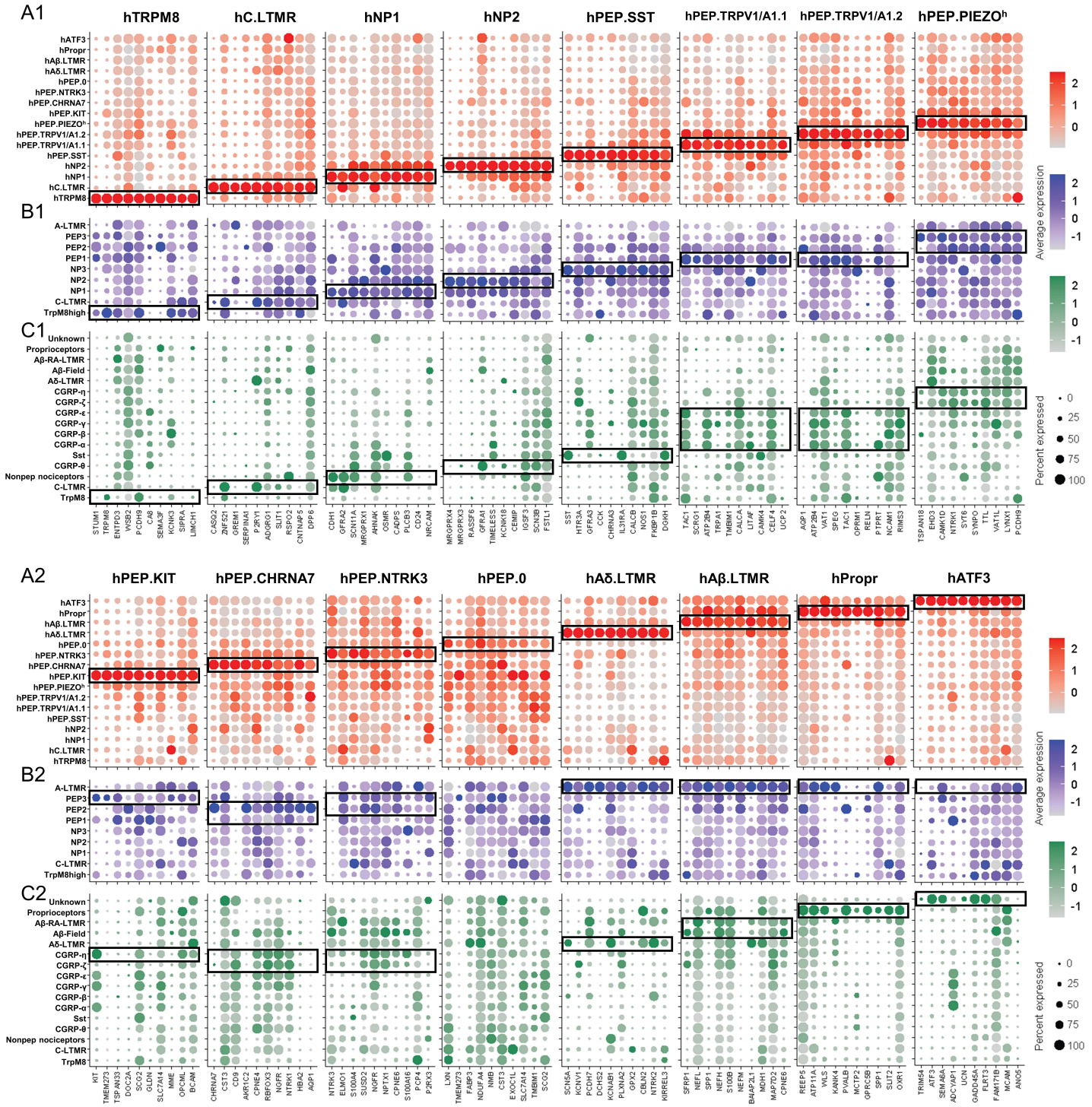
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig 6. Cross-species comparison of sensory neuron types among human, macaque, and mouse



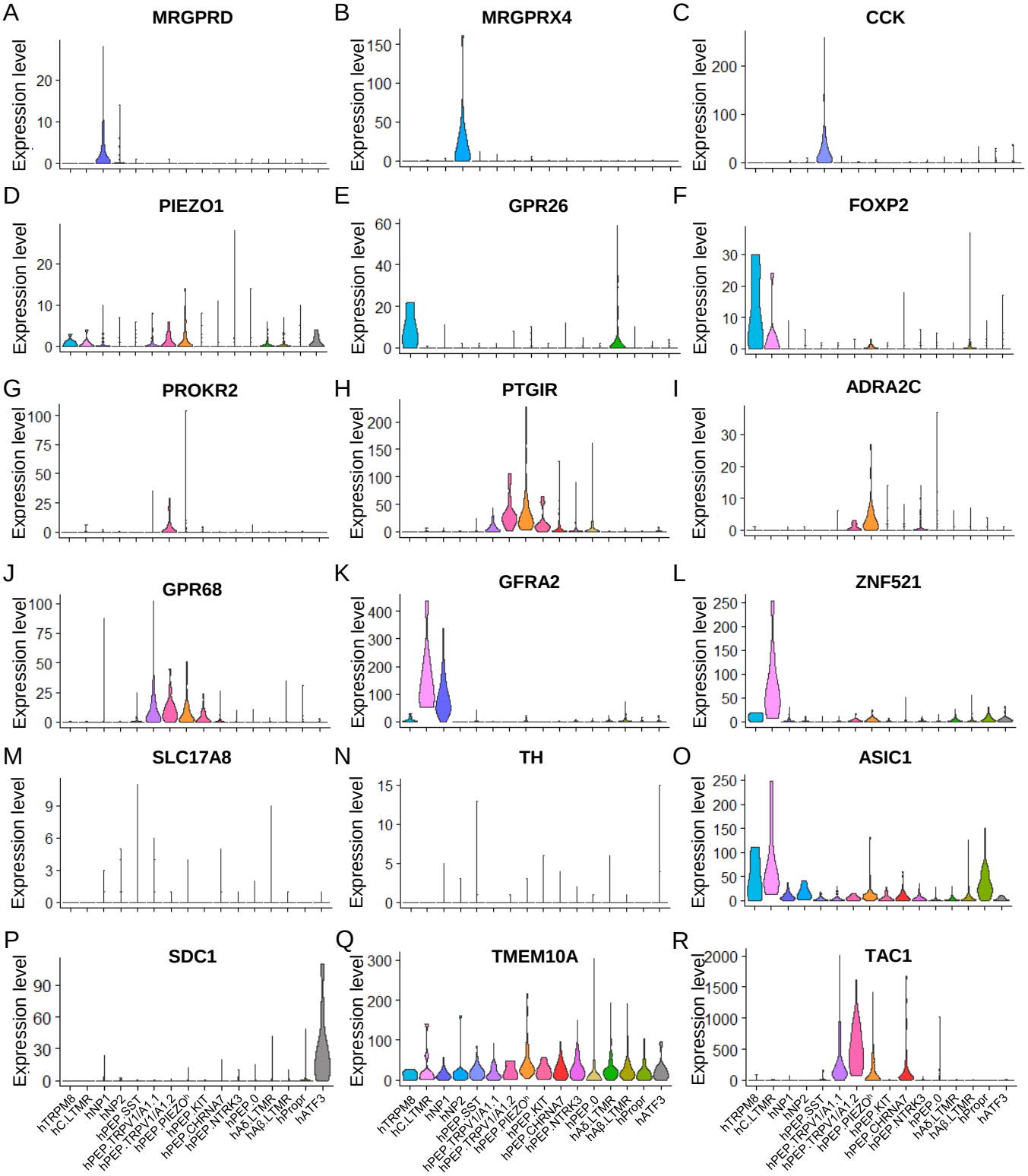
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig. 7 Comparison of marker gene expression across species



(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

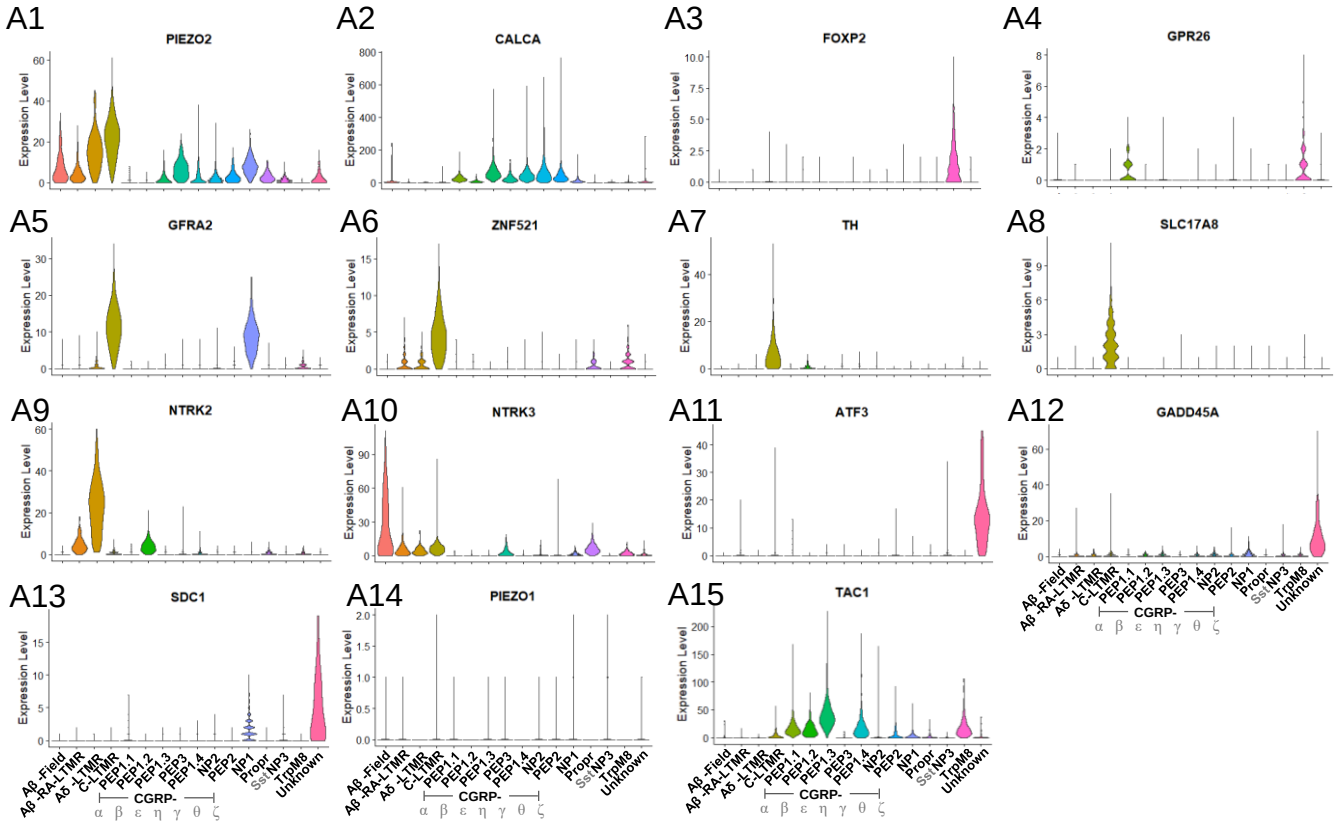
Supplementary Fig. 8 Expression of additional functional or marker genes in human DRG neurons



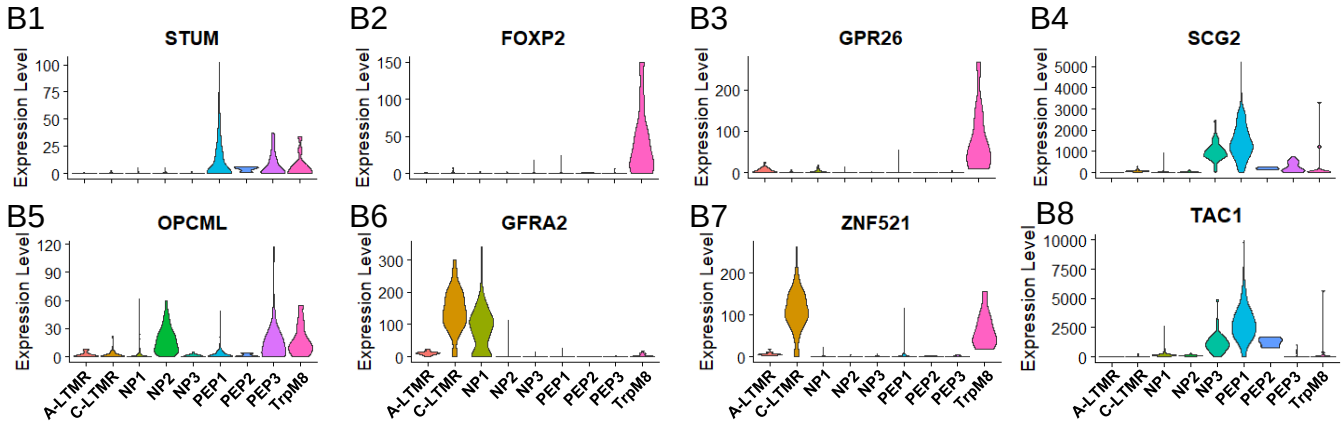
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig. 9 Expression of functional or marker genes in mouse and macaque DRG neurons

Mouse (Sharma)

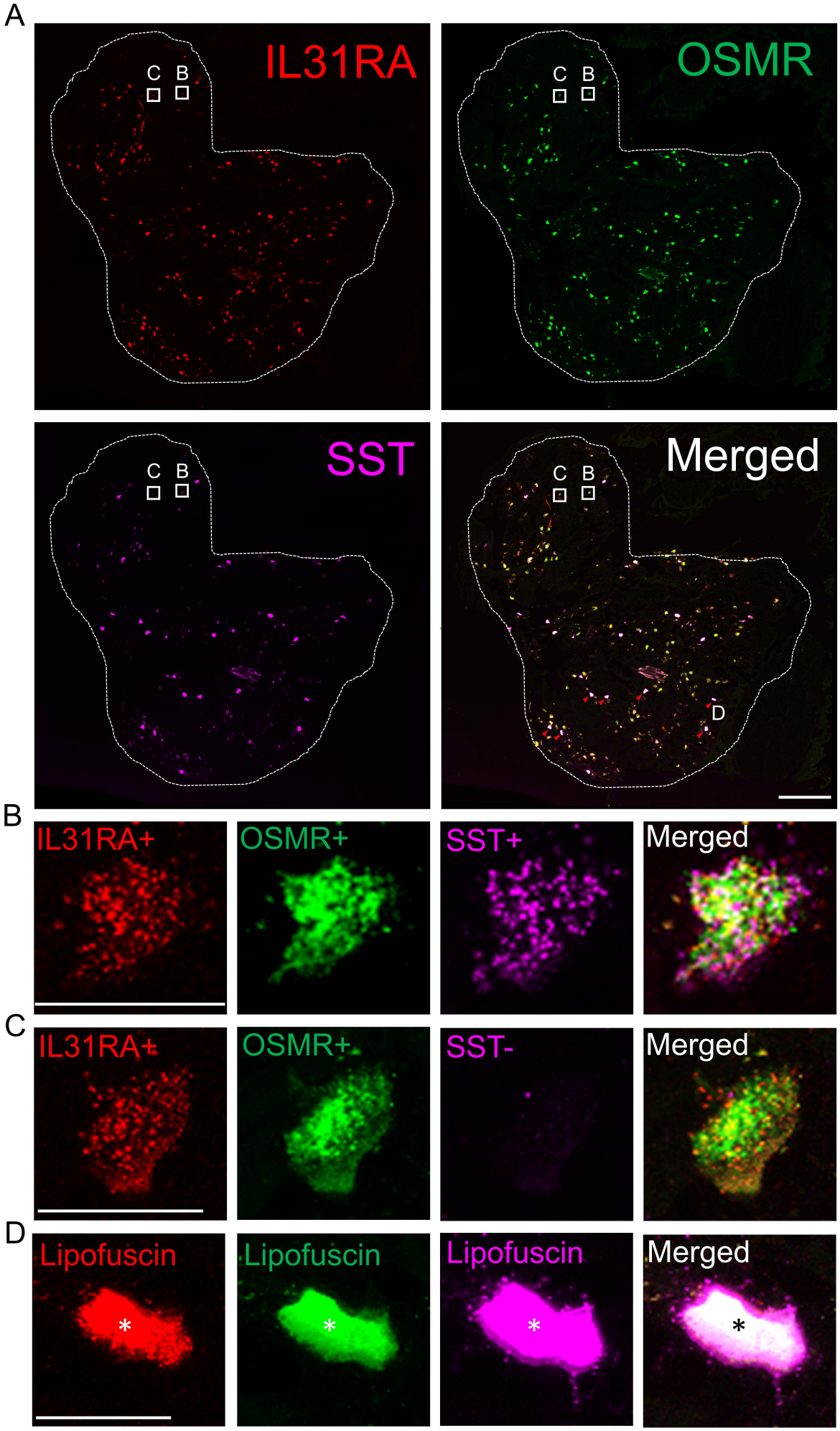


Macaque (Kupari, SmartSeq2)



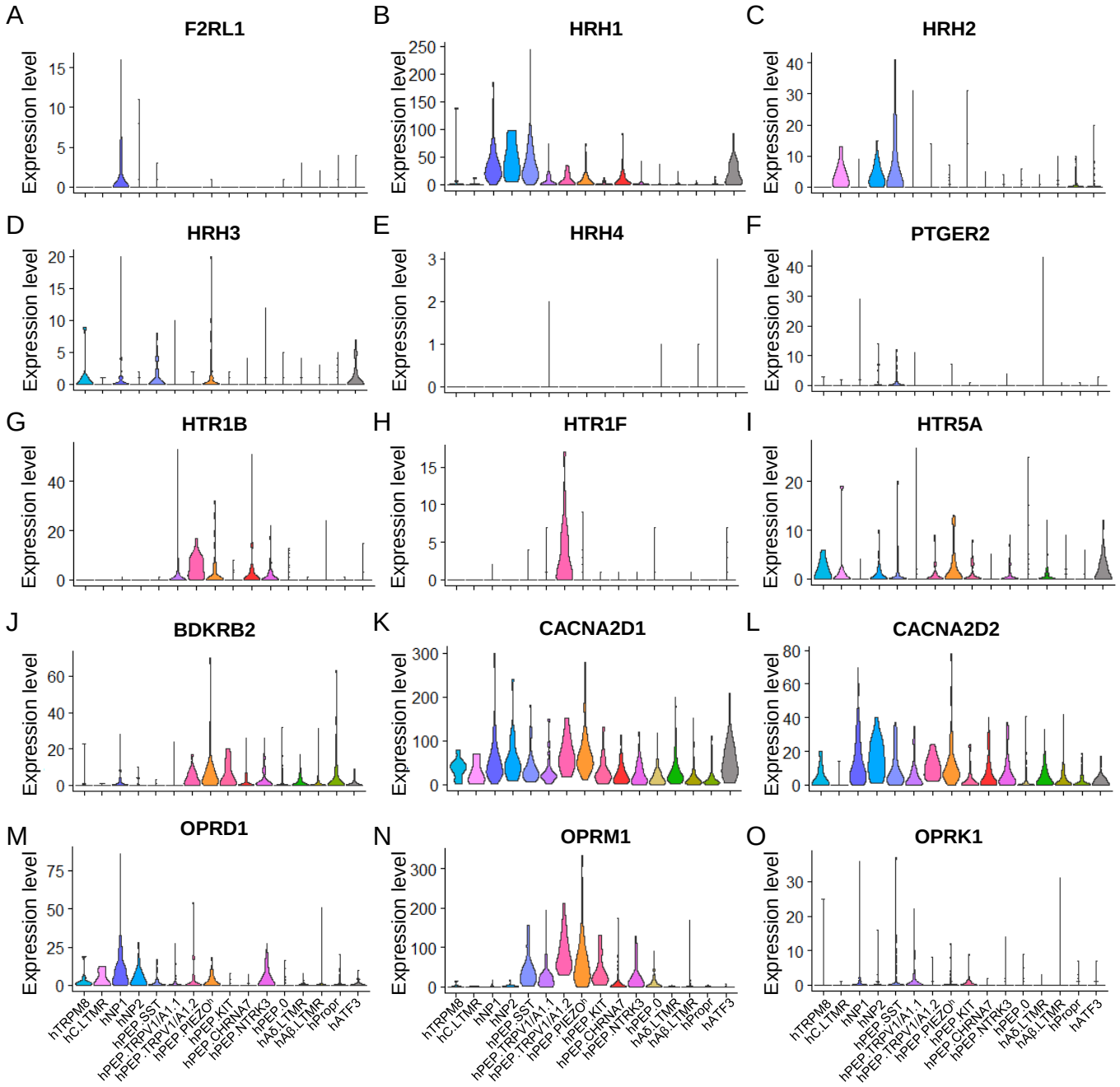
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig. 10 Multiplex FISH of human DRG neuron



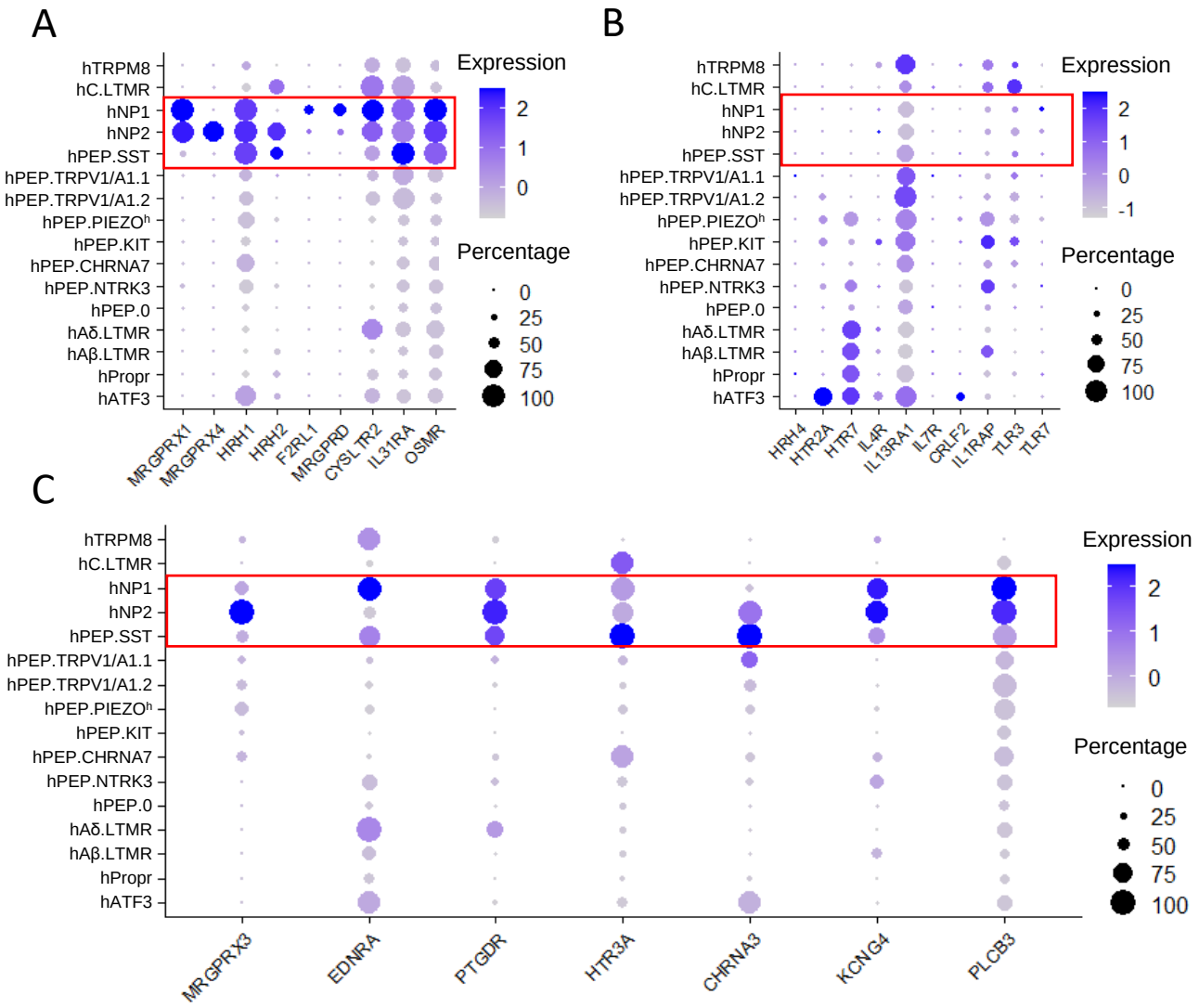
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig. 11 Expression of itch- and pain-sensing related genes in human DRG neurons



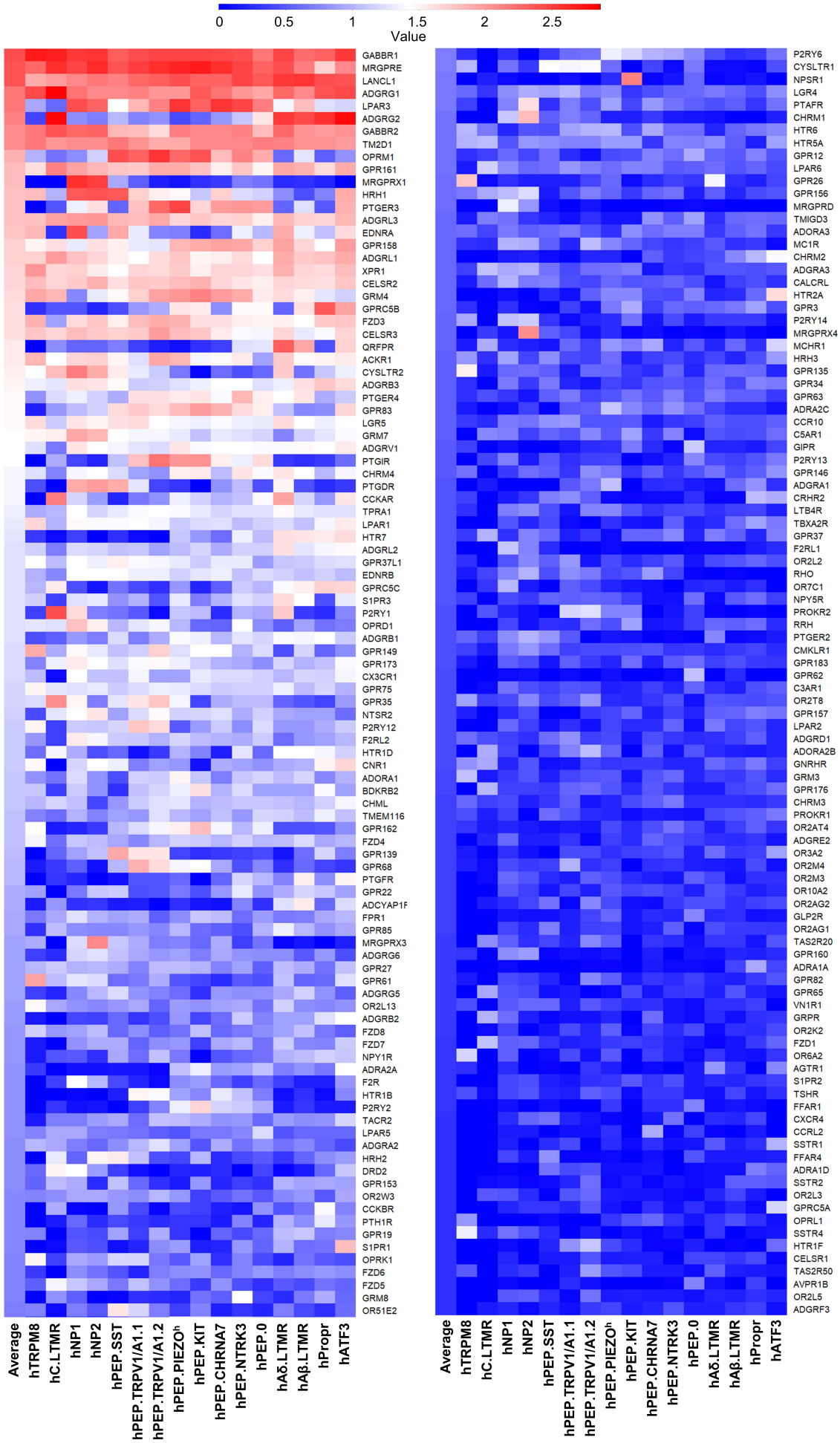
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig. 12 Expression of itch receptors in human DRG neurons and identification of potential novel anti-itch targets



(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig. 13 Expression of GPCRs in human DRG neurons

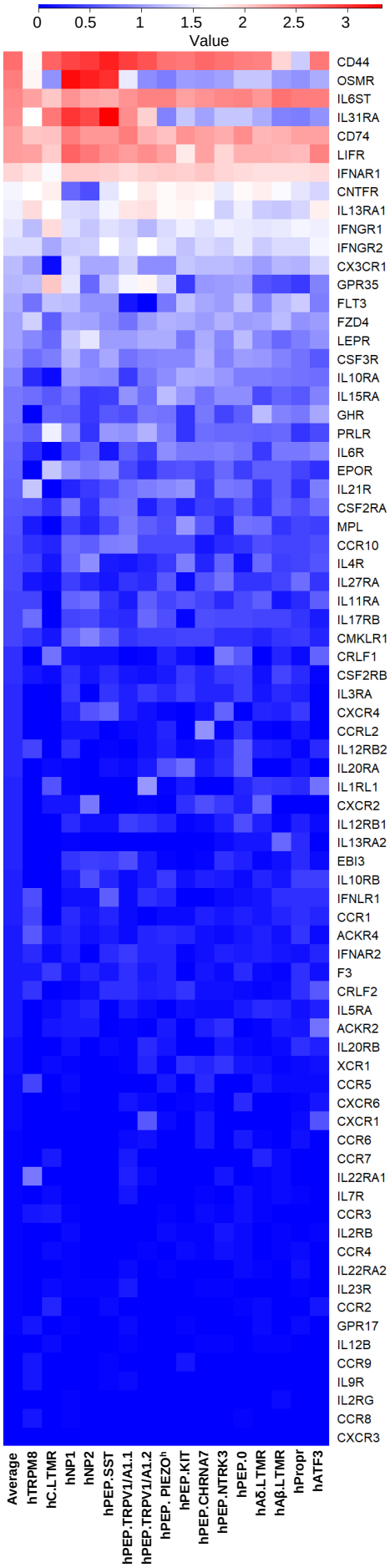


(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig. 14 Expression of ion channels in human DRG neurons

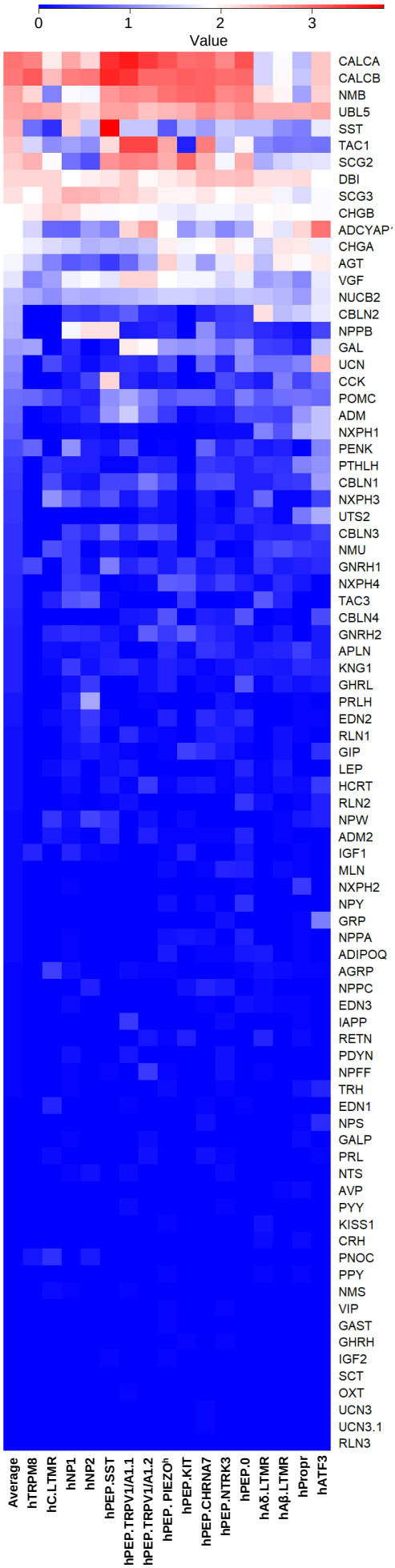
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig. 15 Expression of chemokine receptors in human DRG neurons



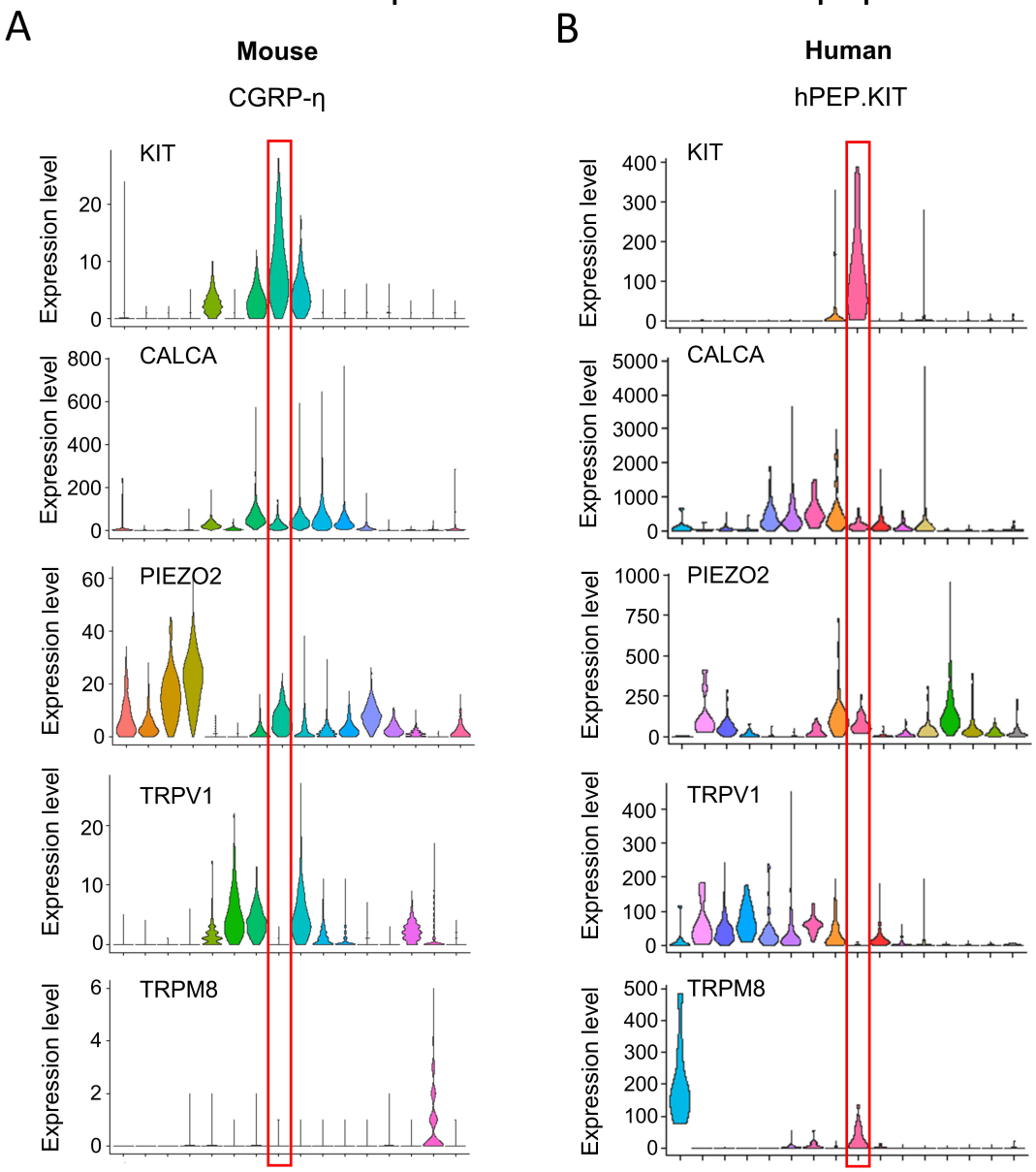
(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

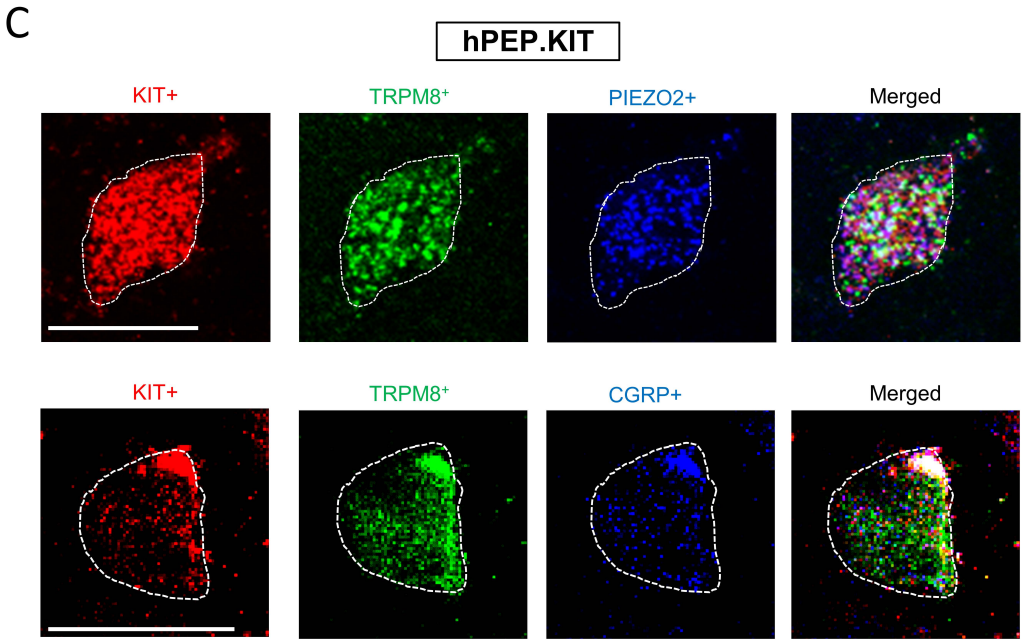
Supplementary Fig. 16 Expression of peptides in human DRG neurons



(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

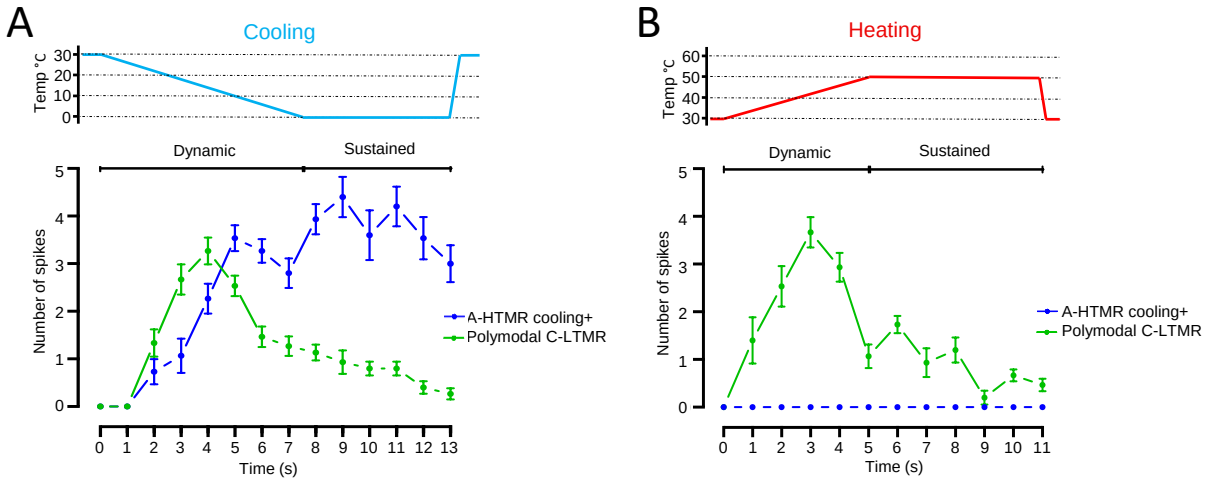
Supplementary Fig. 17 Expression of CGRP, TRPV1, PIEZO2, KIT, and TRPM8 in mouse CGRP- η and human hPEP.KIT populations

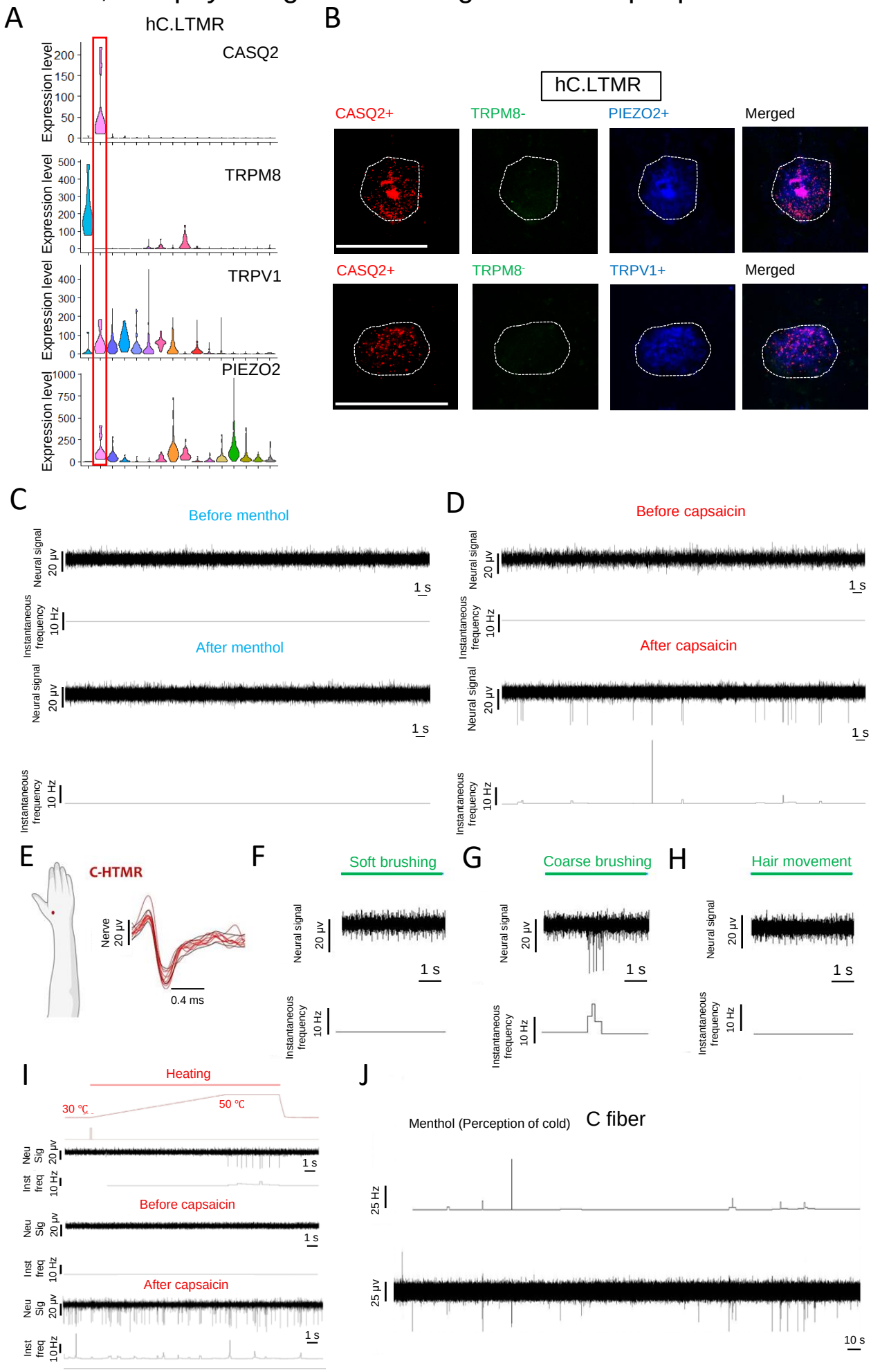




(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig. 18 Responses of A-HTMR cooling+ and polymodal hC.LTMR units to temperature changes



(which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Fig. 19 Expression of TRPM8, TRPV1 and PIEZO2 in hC.LTMR, and physiological recordings of human peripheral nerve fibers