1 **Single-Soma Deep RNA Sequencing of Human Dorsal Root Ganglion Neurons Reveals**

2 **Novel Molecular and Cellular Mechanisms Underlying Somatosensation**

3 Huasheng Yu[1], Dmitry Usoskin[2*], Saad S. Nagi[3*], Yizhou Hu[2], Jussi Kupari[2], Otmane

4 Bouchatta[3], Suna Li Cranfill[1], Mayank Gautam[1], Yijing Su[1], You Lu[1], James Wymer[4], Max

5 Glanz[4], Phillip Albrecht[5], Hongjun Song[1], Guo-Li Ming[1], Stephen Prouty[6], John Seykora[6], Hao

6 Wu[7], Minghong Ma[1], Frank L Rice[5], Håkan Olausson[3#], Patrik Ernfors[2#], Wenqin Luo[1#]

7 1. Department of Neuroscience, Perelman School of Medicine, University of Pennsylvania, Phil-

8 adelphia, PA 19104, USA.

9 2. Department of Medical Biochemistry and Biophysics, Division of Molecular Neurobiology,

10 Karolinska Institute, Stockholm 17177, Sweden.

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

12 Sweden.

13 4. Department of Biochemistry and Molecular Biology, College of Medicine, University of Flor-

14 ida, Gainesville, FL 32603, USA.

15 5. Neuroscience & Pain Research Group, Integrated Tissue Dynamics, LLC, Rensselaer, NY

16 12144, USA.

17 6. Department of Dermatology, Perelman School of Medicine, University of Pennsylvania, Phil-

18 adelphia, PA 19104, USA.

19 7. Departments of Genetics, Perelman School of Medicine, University of Pennsylvania, Phila-

20 delphia, PA 19104, USA

21 * These authors contributed equally to this work

22 # Correspondence:

23 hakan.olausson@liu.se (Håkan Olausson)

24   patrik.ernfors@ki.se (Patrik Ernfors)

25   luow@pennmedicine.upenn.edu (Wenqin Luo)

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47   **Abstract:**

48   The versatility of somatosensation arises from heterogeneous dorsal root ganglion (DRG) neu-

49   rons. However, soma transcriptomes of individual human DRG (hDRG) neurons – critical in-

50   formation to decipher their functions – are lacking due to technical difficulties. Here, we devel-

51   oped a novel approach to isolate individual hDRG neuron somas for deep RNA sequencing

52   (RNA-seq). On average, >9,000 unique genes per neuron were detected, and 16 neuronal types

53   were identified. Cross-species analyses revealed remarkable divergence among pain-sensing neu-

54   rons and the existence of human-specific nociceptor types. Our deep RNA-seq dataset was espe-

55   cially powerful for providing insight into the molecular mechanisms underlying human

56   somatosensation and identifying high potential novel drug targets. Our dataset also guided the

57   selection of molecular markers to visualize different types of human afferents and the discovery

58   of novel functional properties using single-cell *in vivo* electrophysiological recordings. In sum-

59   mary, by employing a novel soma sequencing method, we generated an unprecedented hDRG

60   neuron atlas, providing new insights into human somatosensation, establishing a critical founda-

61   tion for translational work, and clarifying human species-specific properties.

62

63

64

65

66

67

68

69

Introduction:

The somatosensory system conveys senses, such as 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]. Somatosensation is fundamental to our daily lives, but becomes a devastating human health problem when malfunctioning, such as during chronic pain and itch. Safe and effective drug options for chronic pain and itch are still limited[4-6], and the development of novel treatment strategies is greatly needed.

Most of our current knowledge about the mammalian somatosensory system has been obtained from model organisms, mainly rodents. However, the success rate of translating treatment strategies working in model organisms such as rodents to humans has been low[7,8]. There may be multiple reasons for the lack of success, but a noticeable one is species differences of the genetic, molecular and potentially even cellular makeup of somatosensory neurons between rodents and humans. Some significant differences between rodent and human DRG (hDRG) neurons have been noticed in previous studies, with neuropeptides, ion channels, and other markers not always matching between the species[9,10]. For an example, MRGPRX4, a bile acid receptor for human cholestatic itch, does not have a molecular ortholog in mice[11]. In contrast, TGR5, a receptor identified in mouse DRG neurons for bile acid-induced itch[12], is not expressed in hDRG neurons but instead in the surrounding satellite glial cells[11]. Thus, it is critical to elucidate molecular profiles and cell types of hDRG neurons for understanding human somatosensory mechanisms as well as for translational purposes.

93

94   Single-cell RNA sequencing (RNA-seq) is a powerful tool to qualitatively and quantitatively

95   study transcripts of individual cells (soma and/or nuclei)[13]. Based on the transcriptome, including

96   both transcript identities and expression levels, heterogenous cells can be classified into different

97   types[14]. This approach has been successfully used in varying degrees to study DRG and TG neu-

98   rons in mouse, macaque and other model organisms[15-18], providing comprehensive molecular and

99   cellular atlases for understanding somatosensory neurons and their differentiation. However,

100   there are some unique technical challenges for conducting single-cell RNA-seq of human DRG

101   or TG neurons: 1) Human tissues are more difficult to obtain compared to model organisms, and

102   the quality of human tissues (RNA integrity) is much more variable; 2) In human DRG/TG tis-

103   sues, non-neuronal cells, including satellite glial cells, fibroblasts, and other cell types, greatly

104   outnumber the neuronal cells[9,19]; 3) Satellite glial cells tightly wrap around neuronal somas[9,19],

105   making their physical separation difficult; 4) Human DRG/TG neuronal somas are much larger

106   than those of mouse DRG/TG neuronal somas, so they are prone to damage by enzyme digestion

107   and mechanical forces during single-cell isolation and not compatible with many current se-

108   quencing platforms. In addition, transcriptome changes caused by enzymatic and mechanical

109   dissociation may also affect cell clustering and cell type identification[20]. Due to these difficulties,

110   single-nucleus RNA-seq and 10x Visium spatial transcriptomics have been employed for human

111   DRG/TG neurons[21-24]. Despite novel insights from these studies, the quantity of transcripts in the

112   nucleus is much lower than those in the soma, and the nuclear transcripts may not represent the

113   full transcriptome profile of the whole cell[25], while the 10x Visium spatial transcriptomics lack

114   single-neuron resolution (Fig. 1A). Thus, it is necessary to develop a new method that enables

115   soma RNA-seq of hDRG neurons while preserving the single-cell resolution.

116

117   To meet this challenge, we developed a novel approach by combining laser capture

118   microdissection (LCM) for isolating individual neuronal somas and the Smart-seq2[18] for generat-

119   ing full length RNA-seq libraries (Fig. 1A). We sequenced 1066 hDRG neurons with minimum

120   satellite glial cell contamination from six lower thoracic and lumbar DRGs of three donors, de-

121   tecting an average of >9,000 unique genes per neuron (~ 3-5 times or more than the previous

122   single-nucleus RNA-seq results) and identifying 16 molecularly distinctive neuron types. Cross-

123   species analysis revealed both similarities but considerable differences among human, macaque,

124   and mouse DRG neurons. In addition, we uncovered a set of novel marker genes that can help to

125   distinguish different types of sensory neurons and afferents in hDRG and skin tissues. Based on

126   the molecular profiles, we also predicted novel response properties of human sensory afferents,

127   which were tested and confirmed by single-cell *in vivo* recordings. These results support a close

128   relationship between the molecular profiles uncovered by single-soma RNA-seq, histology, and

129   functional properties of human sensory afferents, highlighting the precision and unique utility of

130   this dataset in guiding functional studies of human somatosensory neurons. In short, we have

131   established a novel approach to conduct single-soma deep RNA-seq of hDRG neurons, which

132   revealed previously unknown neuronal types and functional properties. Given the high number

133   of unique transcripts recovered from each neuron, our dataset is especially powerful for molecu-

134   lar discovery, such as identifying potential high value novel drug targets. We believe that this

135   high-fidelity single-soma RNA-seq dataset will serve as a ground reference for homogenizing

136   RNA-seq data of human DRG/TG neurons using different approaches and for translating animal

137   studies into therapeutic applications.

138

Results

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

Adult human DRG neurons present considerable technical challenges for high-quality single-cell RNA-seq study. To overcome these hurdles, we developed a novel method that utilizes the laser capture microdissection to isolate individual neuronal soma and combine with Smart-seq2 full length RNA-seq library generation method for single-soma high-depth transcriptomic analysis (Fig. 1A). Six hDRGs at the low thoracic (T11-T12) and lumbar (L2-L5) levels from three post-mortem donors (donor information and screening criteria are summarized in the supplementary tables 1 & 2) were procured through NDRI (National Disease Research Interchange). DRGs were extracted and immediately frozen in OCT (within ~ 6 hours after death) at the NDRI pro-curement sites to minimize RNA degradation. Fresh frozen DRGs were cryo-sectioned, mounted onto LCM slides, and briefly stained with a HistoGene[TM] dye for cell visualization (Fig. S1A-C). Individual neuronal somas were discernable under microscope and dissected by a laser (Fig. 1A). Each detached soma dropped into a tube cap containing cell lysis buffer for library preparation (Fig. 1A). Dissected neuronal somas exhibited a similar size distribution as the whole DRG neu-ron population (Fig. S1D-E), suggesting no obvious sampling bias. Sequencing libraries were generated following the Smart-seq2 protocol[26]. In total, 1136 neuronal somas were dissected and passed through the final quality control for sequencing. During preliminary bioinformatic analy-sis, 70 samples with obvious glial cell contamination (dominant expression of APEO, FABP7, and other glia cell markers) were removed, and the remaining 1066 neurons were used for analy-sis in this study. 16 transcriptomic clusters of hDRG neurons were identified by Seurat[27] (Fig. 1B), with an average of 9486 genes detected per cell (Fig. 1C). No obvious batch effects or do-

162   nor differences were observed in the clusters (Fig. S2A-E). All these cells expressed high levels

163   of known peripheral sensory neuronal markers, *SLC17A6* (*VGLUT2*), *SYP* (Synaptophysin), and

164   *UCHL1* (*PGP9.5*) (Fig. 1D, S2F-G).

165

166

Anatomical and molecular features of hDRG neuron clusters

168   Mammalian DRG neurons have some well-known physiological, anatomical, and molecular fea-

169   tures. DRG somatosensory afferents can be identified as A- or C- afferent fibers according to

170   axon conduction velocities[28]. A-fiber DRG neurons usually have large-diameter somas and

171   myelinated axons, while C-fiber DRG neurons have small-diameter somas and non-myelinated

172   axons. The A- and C- afferents can be further divided into peptidergic and non-peptidergic types,

173   based on the expression of one neuropeptide, calcitonin related polypeptide alpha

174   (*CALCA(*CGRP*))*[29]. To determine the broad types of the 16 hDRG neuron clusters, we analyzed

175   their soma sizes and expression of some key molecular markers (Fig. 1E). Clusters 1-16 were

176   arranged from small to large soma sizes. In addition, expression of neurofilament intermediate

177   filament (*INA*), which was enriched in small-diameter DRG neurons, and heavy chain (*NEFH*),

178   which was highly expressed in large-diameter DRG neurons, showed a clear complementary

179   pattern. The combined morphological and molecular features suggested that clusters 1-7 were

180   likely to be unmyelinated, small diameter C-fiber DRG neurons, whereas clusters 8-16 were

181   myelinated, large diameter A-fiber DRG neurons (Fig. 1E). The two groups could be further di-

182   vided based on the expression of *CALCA* (clusters 5-12). Moreover, the PR domain zinc finger

183   protein 12 (*PRDM12*), a transcriptional regulator critical for development of human pain-sensing

184   afferents (nociceptors) and C fibers[30], was expressed in clusters 1-12, further distinguishing be-

8

185   tween C- and A-fiber nociceptors versus A-fiber low-threshold mechanoreceptors (Fig. 1E).

186   Taken all information into consideration, clusters 1-4 were classified as non-peptidergic C-fibers,

187   clusters 5-7 as peptidergic C-fibers, clusters 8-12 as peptidergic A-fibers, clusters 13-15 as low-

188   threshold mechanoreceptors A-fibers (A-LTMRs), and cluster 16 as an unknown A-fiber type

189   (Fig. 1E).

190

191   Based on the expression profiles of top molecular markers, which had either unique expression

192   pattern or well-known functions (Fig. 1F, S2H), we named these 16 hDRG neuron clusters using

193   a nomenclature system with the following format (Fig. 1G): 1) The "h" at the beginning of each

194   cluster name indicated "human"; 2) mouse nomenclature was used for conserved subtypes (i.e.

195   most non-peptidergic neurons and A-LTMRs); 3) human peptidergic neuron types were desig-

196   nated as hPEP.(marker gene). Briefly, for non-peptidergic C-fiber neurons, cluster 1 was named

197   hTRPM8, cluster 2 was C-fiber low-threshold mechanoreceptors (hC.LTMR), cluster 3 was type

198   I non-peptidergic nociceptors (hNP1), and cluster 4 was type II non-peptidergic nociceptor

199   (hNP2). For peptidergic C-fiber neurons, cluster 5 was named hPEP.SST, cluster 6

200   hPEP.TRPV1/A1.1, and cluster 7 hPEP.TRPV1/A1.2. For peptidergic A-fiber neurons, cluster 8

201   was named hPEP.PIEZO^h (the superscript 'h' means 'high'), cluster 9 hPEP.KIT, cluster 10

202   hPEP.CHRNA7, cluster 11 hPEP.NTRK3, and cluster 12 hPEP.0 (no distinctive molecular

203   marker). For A-LTMRs, cluster 13 was named Aδ low-threshold mechanoreceptors

204   (hAδ.LTMR), cluster 14 Aβ low-threshold mechanoreceptors (hAβ.LTMR), and cluster 15 pro-

205   prioceptors (hPropr). Cluster 16 was named hATF3.

206

207   To validate the clusters identified by Seurat analysis, we independently analyzed our data using

208   the graph-based clustering Conos[31] package. Cluster structure revealed by Conos analysis repro-

209   duced the Seurat results (Fig. S3A). As a third method to validate identified clusters, we em-

210   ployed a neural network-based probabilistic scoring module[15,18] that learned human cell type

211   features based on their molecular profiles (Fig. S3B-D). Namely, the accuracy score of our Seu-

212   rat clustering assignment by the learning module was near 90% (Fig. S3B), which meant that

213   most cells were accurately assigned to their corresponding clusters, confirming the accuracy of

214   clustering (Fig. S3C). Moreover, the cell type consistency was validated by probabilistic simi-

215   larity (Fig. S3D). Thus, three independent analysis methods confirmed the robustness of the clus-

216   tering and strongly supported the cell type assignment.

217

218   Using Conos we also performed co-clustering and label propagation with a recently published

219   single-nucleus RNA-seq dataset of hDRG neurons by Nguyen et al[23] (Fig. S4). This analysis

220   showed that while some clusters displayed a one-to-one match between the two datasets, such as

221   hTRPM8, hPEP.SST, hPEP.KIT, and hAd.LTMR, most did not have one-to-one match. This

222   mismatch could be caused by biological variations such as nucleus vs cytoplasm RNA species

223   and quantity[25] and technical differences (the increased resolution obtained by deep sequencing in

224   the present study or variability caused by the different technology platforms[26,32]). This mismatch

225   also indicated that our dataset generated novel molecular profiles and cell types different from

226   the single-nucleus RNA-seq results.

227

228

229   **Cross-species comparison of DRG neuron types**

230   Comparison of hDRG neuron types to those in model organisms helps to uncover the evolutional

231   conservation and divergence of DRG neuron populations, provides clues about functions of

232   hDRG neuron populations, and identifies potential species-specific hDRG neuron populations.

233   Here we performed a cross-species comparison between our human dataset, a mouse dataset

234   (Sharma)[16], and a macaque dataset (Kupari, SmartSeq2 dataset)[18]. To identify corresponding

235   clusters between human and previously published mouse and macaque datasets, we used three

236   different strategies: Conos pairwise co-clustering followed by label propagation (Fig. 2A-B &

237   S5A-B), probabilistic neural network learning (Fig. S5C-D), and machine-learning based hierar-

238   chical clustering of an integrated dataset of human and macaque (Fig. S5E). For characteristics

239   and details of different approaches, see Method section. In all these analyses, human non-PEP

240   DRG neuron cell types showed high correlation to those of mouse and macaque, including

241   hC.LTMR, hNP1, hNP2, hPEP.SST (called NP3 in mouse and monkey), hTRPM8, hAδ.LTMR,

242   hAβ.LTMR, and hProprioceptor neurons (Fig. 2A-B & S5). hPEP.TRPV1/A1.1 and

243   hPEP.TRPV1/A1.2 corresponded to macaque PEP1 (Usoskin nomenclature) and mouse subclass

244   PEP1.4/CGRP-ε. This conclusion was supported by both Conos label propagation analysis (Fig.

245   2A-B) and probabilistic neural network analysis (Fig. S5C-D), suggesting that these clusters to

246   represent C-fiber thermoreceptors and nociceptors. Notably, four types of mouse C-fiber PEP

247   (CGRP) nociceptors have been evidenced[18,33] (nomenclature from Emory & Ernfors/Sharma),

248   but our analysis indicated that mouse clusters PEP1.1/CGRP-α, PEP1.2/CGRP-β, and

249   PEP1.3/CGRP-γ might not be evolutionarily conserved, as they did not have equivalent types in

250   the human datasets (Fig. 2A & S5C). hPEP.CHRNA7 showed high correlation to mouse

251   PEP2/CGRP-ζ and macaque PEP2, while hPEP.KIT corresponded to mouse PEP3/CGRP-η and

252   macaque PEP3, suggesting these clusters functionally belong to A-fiber nociceptors (Fig. 2A-B

253 & S5C-D). Interestingly, hPEP.NTRK3, hPEP.PIEZO^h, and hPEP.0, which made about half of

254 human PEP nociceptors that we sequenced, did not have a one-to-one corresponding cell types in

255 either mouse or macaque (Fig. 2A-B & S5C-D). Since macaque PEP2 cluster in the high-quality

256 Smart-Seg2 macaque dataset contained only three neurons, we focused our interpretation on

257 analyses on mouse homologs for this cell type. hPEP.NTRK3 showed the greatest similarity to

258 mouse PEP2 (CGRP-ζ) by Conos propagation analysis (Fig. 2A). Probabilistic neural network

259 learning revealed similar scores to both mouse PEP2 (CGRP-ζ) and PEP3 (CGRP-η) (Fig. S5C).

260 Overall, it therefore seems that hPEP.NTRK3 represents a convergent mouse PEP2/3-like cell

261 type. hPEP.PIEZO^h showed some similarity to mouse PEP3 (CGRP-η) and macaque PEP1 and

262 PEP3 in co-clustering (Fig. 2A-B). Neural network learning, and hierarchical clustering, indicat-

263 ing this cell type to represent an A-fiber mechanosensory nociceptor (Fig. S5C-E). Thus,

264 hPEP.NTRK3 and hPEP.PIEZO^h represents diverged PEP2- and PEP3-like A-fiber nociceptors,

265 which likely have emerged as human species-specific sensory neuron types. hPEP.0, a type of

266 human PEP A-afferents, showed no similarity to any mouse DRG neurons but some relationship

267 to macaque PEP1 and PEP3 (Fig. 2A-B), suggesting that they might be primate specific PEP

268 nociceptors. A schematic overview of our conclusions based on the above analyses regarding the

269 cell-type homologs across mouse-macaque-human is illustrated in Fig. 2C.

270

271 Transcription factors play critical role in DRG neuron development and differentiation[34]. Thus,

272 we performed a transcription factor associated gene regulatory network analysis (TF-GRNs) us-

273 ing machine learning to identify shared and species-specific TF-GRNs contributing to the simi-

274 larities and differences between sensory neuron subtypes and species (Fig. 2D). Evolutionarily

275 conserved TF-GRNs defining C-fiber nociceptors, A-fiber nociceptors, A-LTMRs, TRPM8, C-

276 fiber pruriceptors/nociceptors (hNP1, hNP2, hPEP.SST), and C-LTMRs were observed (yellow

277 boxes in Fig. 2D) as well as species-specific networks, such as for C-fiber nociceptors, hTRPM8,

278 hNP1, hNP2, hPEP.SST and hC.LTMR (green boxes in Fig. 2D). Among cross-species con-

279 served transcriptional regulators, some were previously known to drive sensory neuron diversifi-

280 cation in mouse, including *ZEB2* in C-fiber nociceptors[35], *SHOX2* in A-LTMRs[36,37], *RUNX3* in

281 proprioceptors[38], *FOXP2* in TRPM8, *RUNX1* in NP1[39], *ZNF52* and *POU4F2* in C-LTMRs[16,40].

282 These transcription factors may contribute to the formation of different DRG neuron cell types

283 and regulate the conserved and species-specific gene expression in each cell population.

284

285

286 **Similarities and differences of top marker genes across species**

287 To reveal molecular differences among the corresponding cell types in human, macaque

288 (Kupari)[18], and mouse (Sharma)[16], we selected the top 10 marker genes from each hDRG neuron

289 population and compared them across species (Fig. S6). In general, the expression patterns of

290 these genes were more similar between human and macaque than between human and mouse,

291 reflecting the evolutionary distances between rodent, non-human primate, and human. Some

292 genes were expressed in the corresponding populations across all three species. For example,

293 *TRPM8* was expressed in C-fiber cold-sensing neurons, and *IL31RA* was expressed in nocicep-

294 tive/pruriceptive population (NP3) as well as its corresponding human population hPEP.SST.

295 Some genes, such as *KCNV1*, a voltage-gated potassium channel, were specific for primate

296 Aδ.LTMR but had low or no expression in the corresponding mouse DRG neurons (Fig. S6).

297 Moreover, some marker genes were specific only for hDRG neuron types. For example,

298 Mechanosensory Transduction Mediator Homolog (*STUM*) was specifically expressed in

13

299   hTRPM8, and calsequestrin 2 (*CASQ2*) was specifically expressed in hC.LTMR (Fig. S6). Genes

300   specifically enriched in hDRG neurons may confer unique physiological and functional proper-

301   ties of the human somatosensory system. The observed differences from the top 10 marker genes

302   represent only the iceberg tip of the overall molecular differences between human and model

303   organism DRG neurons, highlighting the necessity of validating molecular targets in hDRG neu-

304   rons for translational studies.

305

306

307   **Molecular marker expression and validation of C-fiber pruriceptors, thermoreceptors and**

308   **nociceptors**

309   Based on the sequencing results, we established specific marker genes or their combination to

310   identify each type of hDRG neurons (Fig. S7A) and validated their expression using

311   RNASCOPE multiplex fluorescent *in situ* hybridization (FISH) (Fig. S8). We also deduced po-

312   tential functions of each hDRG neuron type based on the cross-species cell type comparison and

313   expression of molecules with known sensory functions.

314

315   hNP1 and hNP2 exclusively expressed *MRGPRX1* (Fig. 3A), a Mas-related GPCR family mem-

316   ber that could be activated by various pruritogens[41]. Similarly, primate specific bile acid receptor

317   *MRGPRX4* for mediating human cholestatic itch[11,42] were detected by our single-soma dataset.

318   Enrichment of *MRGPRX4* and *MRGPRX3* (an orphan GPCR in the same family) in hNP2 but not

319   hNP1 helped to separate these two clusters (Fig. S7A). Other itch-related receptors, such as

320   *HRH1* and *IL31RA* (Fig. 3A-B), were also expressed in both hNP1 and hNP2, suggesting that

321   these two populations function to detect various pruritogens and transmit the sensation of itch.

322   The co-expression of *MRGPRX1* and *HRH1* was validated by multiplex FISH (Fig. 3A). *PIEZO2*

323   was expressed at a higher level in hNP1 than in hNP2 neurons (Fig. S7B), suggesting that hNP1

324   neurons could be more mechanosensitive. Consistent with this molecular feature, recordings

325   from human afferents have revealed that some human histaminergic itch-sensing fibers are

326   mechanosensitive[43]. In mice, NP1 and NP2 afferents are well-characterized by specific expres-

327   sion of two different *Mrgpr* members: NP1 neurons (~20% of total DRG neurons) highly express

328   *Mrgprd*[44], while NP2 neurons (~5% of total DRG neurons) express *MrgprA3*[15]. In humans, how-

329   ever, *MRGPRD* was only expressed in a few NP1 neurons (Fig. S9A), so it was less useful for

330   marking human NP1 population, and *MRGPRA3* gene does not have an orthologue in the human

331   genome. In short, NP1 and NP2 populations likely have conserved physiological functions in

332   mediating itch sensation in both mice and humans. However, some key molecular receptors for

333   detecting pruritogens are different between the species, which may reflect evolutionary adapta-

334   tion to distinctive pruritogens encountered by human and mice in their living environments.

335

336   The hPEP.SST population displayed specific expression of the neuropeptide, somatostatin (*SST*),

337   and an enriched expression of *GFRA3,* a co-receptor of tyrosine kinase RET (Fig. 3B, S7A). We

338   also found another neuropeptide cholecystokinin (*CCK*)[45] enriched in this population (Fig. S7A).

339   The hPEP.SST cluster corresponded to mouse and macaque NP3 population (Fig. 2), which are

340   also marked by the expression of *SST*[15,16]. Given the previously established role of mouse NP3

341   neurons in itch sensation[46] and the high expression of itch-sensing receptors, such as *HRH1*,

342   *IL31RA* (Fig. 3A-B), hPEP.SST afferents could also mediate itch sensation, especially under

343   inflammatory conditions[15]. Neither *PIEZO1* nor *PIEZO2* (Fig. S9A) was detected in hPEP.SST

344   neurons, indicating that these afferents might not be mechanosensitive. Indeed, some human

345 histaminergic itch-sensing fibers are insensitive to mechanical forces[43]. A human feature of

346 hPEP.SST neurons was the co-expression of the peptidergic neuron marker, CGRP, which was

347 barely detected in the corresponding mouse NP3 neurons (Fig. S9B)[16,18].

348

349 The hTRPM8 population was distinguished from other cell types by the specific expression of a

350 novel molecular marker, *STUM,* and high-level expression of *TRPM8* (Fig. 3C). Almost all

351 *STUM*[+] neurons expressed *TRPM8,* which was validated by multiplex FISH (Fig. 3C). Since

352 *TRPM8* is a receptor for cold temperature and cooling chemicals (such as menthol)[47-50], the

353 hTRPM8 population likely functions as cold- and menthol-sensing afferents. Notably, this newly

354 identified maker gene *STUM* was not detected in mouse *TRPM8*[+] neurons[16,33]. In macaque,

355 *STUM* was more broadly expressed in several clusters (Fig. S9C)[18]. Nevertheless, other molecu-

356 lar markers, such as *FOXP2* and *GPR26*, were shared among mouse[16,33], macaque[18], and human

357 TRPM8 cold-sensing neurons (Fig. S9A-C). Intriguingly, some hTRPM8 neurons also expressed

358 a low level of heat-sensing receptor *TRPV1* (Fig. 3D), suggesting that these neurons might also

359 be activated by heat stimuli. Consistently, human physiological recordings have identified cold-

360 sensitive C-fibers that also respond to heat[51]. Thus, some of the neurons in the hTRPM8 popula-

361 tion are likely to be polymodal cold-sensing afferents.

362

363 Two peptidergic C-fiber clusters displayed overlapping high expression of *TRPV1* and *TRPA1*

364 (Fig. 3D), which were therefore named as hPEP.TRPV1/A1.1 and hPEP.TRPV1/A1.2. Since

365 *TRPV1* is activated by noxious heat and capsaicin, and *TRPA1* can be activated by noxious cold

366 and various noxious chemicals[52,53], they are likely to be C-fiber peptidergic thermoreceptors and

367 nociceptors, sensing noxious thermal and chemical stimuli and transmitting pain signals. From

368   the cross-species comparison, hPEP.TRPV1/A1.1 and hPEP.TRPV1/A1.2 were mostly similar to

369   macaque PEP1 and mouse subclass PEP1.4/CGRP-ε (Fig. 2A-C). The exact functional differ-

370   ences between these two quite similar populations are yet to be established, but it is tempting to

371   hypothesize that hPEP.TRPV1/A1.1 innervates the skin while hPEP.TRPV1/A1.2 innervates

372   deep tissues and visceral organs, because CGRP$^+$/TRPV1$^+$ afferents were observed in both the

373   human skin and deep tissues[54]. In contrast to the hPEP.TRPV1/A1.1 population,

374   hPEP.TRPV1/A1.2 neurons expressed PROKR2, and higher level of PTGER3, and prostaglan-

375   din I2 receptor (PTGIR) (Fig. S7E & S9A). These molecular markers have been found to be en-

376   riched in mouse viscera-innervating DRG neurons[55,56].

377

378

379   **Molecular marker expression and validation of A-fiber peptidergic nociceptors**

380   Five clusters of peptidergic A-fiber nociceptors were identified in our dataset. When compared to

381   mouse and macaque peptidergic populations, some clusters showed the greatest divergence (Fig.

382   2A-C), indicating that they might contain human specific sensory neuron types.

383

384   The hPEP.PIEZO^h cluster was named because they expressed the highest number of *PIEZO2*

385   transcripts among all PEP clusters (Fig. 4A), an expression level comparable to the most

386   mechanosensitive afferents, hC.LTMR and hAδ.LTMR. The hPEP.PIEZO^h neurons also ex-

387   pressed a relatively high levels of *PIEZO1* transcripts, though the overall expression of *PIEZO1*

388   in hDRG neurons was low (Fig. S9A). This cluster could be identified by its high expression of

389   *PTGER3* but not *TRPA1* (Fig. 4A & 3D). No mouse DRG neuron population was highly

390   matched for the hPEP.PIEZO^h cluster. Nevertheless, the expression of a few unique marker

17

391   genes in this cluster could provide clues into its functions. The adrenoreceptor, *ADRA2C*, a spe-

392   cific molecular marker found in human sensory fibers innervating blood vessels[57-60], was mainly

393   detected in this cluster (Fig. S9A). In addition, the hPEP.PIEZO^h population expressed *GPR68*, a

394   membrane receptor sensing flow and shear forces in the vascular endothelia cells[61] (Fig. S9A).

395   Given the well-established functions of *PIEZO1* and *PIEZO2* in mediating neuronal sensing of

396   blood pressure and the baroreceptor reflex[62] and the expression of known function of *ADRA2C*

397   and *GPR68*, we proposed that some hPEP.PIEZO^h afferents innervate blood vessels and sense

398   the blood pressure or flow. This cluster also expressed a high level of *PTGIR* (Fig. S9A). Mouse

399   *PTGIR*$^+$ DRG neurons innervate the bladder[55], and *PIEZO2* expressed in these neurons is re-

400   quired for sensing the bladder pressure to coordinate urination[63]. Thus, some hPEP.PIEZO^h af-

401   ferents might also innervate bladder and play a similar role. Take all into consideration,

402   hPEP.PIEZO^h neurons might function in sensing mechanical forces from blood vessels and in-

403   ternal organs. Given that there was not a clear mouse DRG CGRP$^+$ population with high *PIEZO2*

404   expression from single-cell RNA-seq datasets, we speculated that hPEP.PIEZO^h neurons are

405   either human specific or greatly expanded in humans. A fundamental difference between human

406   and mouse is their body sizes (humans are more than 2000 times larger than mice[64]). This addi-

407   tion or expansion of hPEP.PIEZO^h neurons is likely to meet the challenge of mediating sensation

408   from internal organs and blood vessels in much larger human bodies.

409

410   The hPEP.KIT cluster had the specific expression of a receptor tyrosine kinase, *KIT*, and a medi-

411   um to low expression level of *PIEZO2* (Fig. 4B). In mouse DRG neurons, *Kit* is more broadly

412   expressed and found in four peptidergic clusters[16] (Fig. S9B). Cross-species analysis suggest that

413   this cluster mainly corresponded to the mouse PEP3/CGRP-η and macaque PEP3 population

414 (Fig. 2A-C), which are A-fiber high threshold mechanoreceptors (HTMRs), forming circumfer-

415 ential endings around hair follicles and mediating hair pulling pain[65,66]. A more recent study sug-

416 gests that mouse CGRP-η neurons mediates mechanosensation from the distal colon[67]. Thus, the

417 hPEP.KIT cluster likely functions as a population of fast-conducting mechano-nociceptors.

418

419 The third peptidergic A-fiber cluster, hPEP.CHRNA7, featured abundant expression of *CHRNA7*

420 but almost no expression of *PIEZO2* (Fig. 4C). Cross-species analysis revealed that this cluster

421 corresponds to mouse PEP2/CGRP-ζ and macaque PEP2 populations (Fig. 2A-C), which are also

422 marked by the unique expression of *CHRNA7*[16,18]. Interestingly, this cluster also expressed

423 *PVALB* (Fig. 1F), a molecular marker commonly used for proprioceptors in mouse and human.

424 Retrograde tracing from the mouse gastrocnemius muscle labeled *CHRNA7*[+] DRG neurons[55],

425 suggesting that hPEP.CHRNA7 may contain muscle-innervating A-fiber nociceptive sensory

426 afferents.

427

428 hPEP.NTRK3 is a population of peptidergic A-fibers with high expression of *NTRK3* and

429 *S100A4* (Fig. 4D). Neurons in this cluster expressed a low level of *PIEZO2* (Fig. 4B). Finally,

430 hPEP0 is a population of peptidergic A-fibers that expressed *CALCA* and a moderate level of

431 *PIEZO2* but lack of other specific marker genes. Potential function of hPEP.NTRK3 and hPEP0

432 are currently unclear. They could be some types of A-fiber mechano-nociceptor[68].

433

434 **Molecular marker expression and validation of C-LTMRs, A-LTMRs and an ATF3 popu-**

435 **lation**

436   hC.LTMR is the putative human C-tactile nerve fibers. All cross-species neuron cluster compari-

437   son methods unequivocally identified C-LTMRs as conserved across all three species. Neverthe-

438   less, the specific marker gene *CASQ2* for hC.LTM (Fig. 5A) is not detected in either mouse or

439   macaque C-LTMRs[15,18]. Conversely, mouse C.LTMR cells are characterized by exclusive ex-

440   pression of tyrosine hydroxylase (TH) and *SLC17A8* (*VGLUT3*)[15] (Fig. S9B), which are barely

441   expressed in hDRG neurons (Fig. S9A). Thus, the molecular markers to identify C.LTMRs are

442   different between human and mouse. On the other hand, human, mouse, and macaque C.LTMRs

443   all had conserved expression of GFRA2, another co-receptor of RET, and a zinc finger transcrip-

444   tion factor ZNF521 (Fig. S9A-C). Multiplex FISH confirmed that $CASQ2^+$ cells expressed high

445   levels of *PIEZO2* (Fig. 5A). hC.LTMRs likely mediate innocuous affective touch sensation[70-72].

446

447   A-LTMRs were featured by large diameter somas and high expression of *NTRK2* and *NTRK3*

448   (Fig. 5B-C) but a lack of expressions of *SCN10A* and *PRDM12* (Fig. 4C & 1E), two genes highly

449   associated with human nociception[30,73,74]. We identified 4 clusters in this category. hAδ.LTMR

450   was named based on its high expression level of *NTRK2* and *PIEZO2*, but low level of *NTRK3*

451   (Fig. 5A-C), a molecular feature similar to the mouse Aδ-LTMRs (Fig. S9B). *KCNV1* was en-

452   riched in this cluster and could serve as a novel molecular marker for identifying this population

453   in hDRG neurons (Fig. 5B). hAβ.LTMR, for tactile touch, expressed higher level of *NTRK3* but

454   a lower level of *NTRK2* compared to hAδ.LTMRs (Fig. 5C). hPropr, for limb position sensing,

455   expressed a high level of a proprioceptor marker *PVALB* and *REEP5* (Fig. 5D).

456

457   We also identified a cluster named as hATF3 (Fig. 5E), which contained mainly large diameter

458   neurons and strongly corresponded the "unknown" cluster first identified in normal male mice by

20

459 the Sharma RNA-seq dataset[16]. This cluster in both human and mouse datasets showed a very

460 conserved molecular profile (Fig. S7A). One of the markers it expressed, *ATF3*, is a transcrip-

461 tional factor associated with sensory afferent injury, indicating that these cells may represent

462 neurons that are undergoing (or previously underwent) stress/damage/insult (Fig. 5E). *ATF3*[+]

463 neurons also expressed high level of neuropeptide *ADCYAP1* (Fig. 5E). Since there were no

464 medical records indicating obvious nerve injuries in our human donors, and since the mouse data

465 came from naïve mice, we speculate that these neurons might reflect a very low level of sporadic

466 sensory afferent injury accumulated from daily lives. On the other hand, we cannot exclude the

467 possibility that *ATF3*[+] cluster represent a population of normal DRG neurons.

468

469

470 **The single-soma deep RNA-seq dataset provides novel insights into the molecular and cel-**

471 **lular mechanisms underlying human sensation of itch and pain**

472 Given the high number of unique transcripts per neuronal soma, our dataset is uniquely powerful

473 for molecular discovery. We identified a much higher number of membrane proteins, such as

474 GCPRs, ion channels, chemokine receptors, as well as neuropeptides, compared to 10x Visium

475 Spatial RNA-seq or single-nucleus RNA-seq datasets (Fig. 6A). These discoveries provided nov-

476 el insights into understanding molecular and cellular bases of physiological recording and psy-

477 chological experiment results regarding human itch and pain sensation.

478

479 Physiological recordings have identified at least two groups of C-fibers responding to different

480 pruritogens. One group responds to cowhage, a plant triggering intense itch in human[75], and

481 comprises mechano-sensitive polymodal units[43]. The other group responds to histamine with

21

482   sustained discharges but is mechanically insensitive[43]. However, the molecular and cellular bases

483   for these observations remain unclear. To understand these physiological properties, we analyzed

484   the expression of histamine and cowhage receptors, and the mechanoreceptor *PIEZO2* in the

485   three potential itch populations, hNP1, hNP2 and hPEP.SST. *PIEZO2* was highly expressed in

486   hNP1, with low expression in hNP2 and almost no detectable expression in hPEP.SST (Fig. 5A).

487   Protease-activated receptor *F2RL1*, the receptor mediating cowhage induced itch in humans[75],

488   was exclusively expressed in hNP1 (Fig. 6B). Thus, the hNP1 population likely contained C-

489   fiber pruriceptive afferents sensitive to both cowhage and mechanical stimuli. For histamine re-

490   ceptors, *HRH1* was expressed in all three itch populations, *HRH2* was expressed in hNP2 and

491   hPEP.SST but not in hNP1, *HRH3* had low expression in hNP1 and hPEP.SST, while *HRH4* was

492   not detected (Fig. 6B-C). Thus, hNP2 and hPEP.SST clusters are good candidates for histamine-

493   sensitive but mechano-insensitive itch-sensing C-fibers.

494

495   When intracutaneously applied into human skin, some chemicals tend to trigger itch, such as

496   histamine and prostaglandin E2 (PGE2), while others preferentially induce pain, such as seroto-

497   nin and bradykinin[76]. To explore the potential underlying mechanisms, we examined the expres-

498   sion of all known receptors for these chemicals in our dataset. We found that the receptors of

499   itch-inducing chemicals are enriched in itch sensing populations. For example, histamine recep-

500   tors *HRH1* and *HRH2*, and PGE2 receptor *PTGER2* were enriched in hNP1, hNP2 and

501   hPEP.SST putative itch-sensing neurons (Fig. S10A1). In contrast, serotonin receptors, *HTR1B*

502   and *HTR1F*, and bradykinin receptor *BDKRB2* were enriched in putative nociceptive populations

503   (Fig. S10A2). The differential expression patterns of these receptors might explain the different

504   sensory experience induced by these chemicals.

505

506

Utility of single-soma deep RNA-seq dataset for discovering novel drug targets and for obtaining insights into molecular targets of existing drugs

Our deep sequencing dataset can also provide novel insights into the mechanisms of existing clinical drugs or for identifying potential novel drug targets. The expression patterns of promising drug targets, such as GPCRs, ion channels, chemokine and cytokine receptors, and neuropeptides, were analyzed (Fig. S11-S14). Here we analyzed itch-related receptors and molecules as an example. A series of itch-sensing receptors have been identified in model organisms[3,46,79], but which of these targets can be translated remains to be determined. A subset of these itch receptors were indeed enriched in hNP1, hNP2 and hPEP.SST populations, such as the chloroquine receptor *MRGPRX1*, bile acid receptor *MRGPRX4*, histamine receptor *HRH1*, leukotriene receptor *CYSLTR2*, and interleukin receptors *IL31RA* and *OSMR* (Fig. 6B)[16,18]. However, some suspected itch receptors did not exhibit expression enrichment in human itch populations. For example, *HRH4* was reported to mediate histamine-dependent itch in mice[80], but expression of *HRH4* was barely detected in human itch-sensing populations (Fig. 6C), suggesting that *HRH4* might not be directly involved in human histaminergic itch. The same was found true for *IL7R* and *TLR7*[79], which were proposed to mediate non-histaminergic itch in mice (Fig. 6C). In addition to these known players, we identified other membrane receptors and signaling molecules in human itch populations (Fig. 6D), including MRGPRX3, EDNRA, PTGDR, HTR3A, CHRNA3, KCNG4, and PLCB3, which could represent novel anti-itch targets.

526

527   Gabapentin was originally developed for treating epilepsy and more recently used in the treat-

528   ment of neuropathic pain and chronic itch[81]. It inhibits neurotransmitter release by acting on α2δ-

529   1 and α2δ-2 voltage-dependent calcium channels *CACNA2D1* and *CACNA2D2*[81]. We found that

530   both receptors were broadly expressed in hDRG neurons (Fig. S10B), suggesting that one poten-

531   tial mechanism by which gabapentin could provide clinical benefit is through inhibiting synaptic

532   transmission of primary afferents, as has been shown in mice[82].

533

534   Opioids and derivatives activate opioid receptors to modulate pain and itch. Agonists of the μ-

535   opioid receptor (*OPRM1*) alleviate pain but elicit itch in humans and model organisms, whereas

536   agonists of the κ-opioid receptor (*OPRK1*) inhibit itch in humans[83,84]. Unlike mouse DRG neu-

537   rons, which did not5 display cell-type enriched expression patterns of opioid receptors (Fig.

538   S10C), we found that opioid receptors in hDRG neurons were present in some but not other neu-

539   ron populations. Transcripts of δ-opioid receptor (*OPRD1*) were preferentially expressed in itch

540   populations hNP1 and hNP2, while *OPRM1* was enriched in all hPEP clusters (Fig. S10C). Since

541   opioid receptors are inhibitory GPCRs, our results suggest that activation of *OPRM1* could di-

542   rectly inhibit human nociceptive afferents, while *OPRD1* could be a molecular target for inhibit-

543   ing itch transmission. On the other hand, *OPRK1* was barely detected in our dataset, suggesting

544   that *OPRK1* agonists may relieve itch through indirect or central mechanisms (Fig. S10C).

545

546

547   **Immunostaining of sensory fibers in the human skin using molecular markers identified by**

548   **single-soma deep RNA-seq dataset**

24

549   Although human peripheral sensory afferents can be identified and visualized using the pan neu-

550   ronal marker antibody against PGP9.5, different types of human sensory afferents cannot be dis-

551   tinguished due to a lack of specific molecular markers. Developing an antibody panel to label

552   different types of human sensory afferents would be invaluable for basic research, translational

553   studies, and diagnostics. To start this direction and to test whether some molecular markers iden-

554   tified by the single-soma RNA-seq are useful for labeling specific types of human sensory fibers,

555   we conducted immunostaining of sensory afferents with leg skin biopsies from three normal

556   adult donors (see Methods for detailed donor information).

557

558   *SST* is specifically expressed in hPEP.SST neurons (Fig. 3B). Consistent with this molecular

559   prediction, we found SST$^+$ sensory fibers in the human skin sections in the dermis, the epider-

560   mis-dermis junction, entering the epidermis (Fig. S15A-C), and near the hair follicle (Fig. S15D-

561   F). In addition, double immunostaining revealed that SST$^+$ sensory fibers were CGRP$^+$ and made

562   up a subset of CGPR$^+$ sensory fibers (Fig. S15A-F). These results confirmed that human SST$^+$

563   sensory afferents had CGRP proteins and belonged to the PEP but not NP afferents. Since CGRP

564   involves in neuro-immune interactions and other physiological functions[85], addition of this neu-

565   ropeptide in human SST$^+$ afferents suggests a potential gain of new functions during evolution.

566

567   hPEP.KIT neurons specifically express *KIT* transcripts. Strong KIT signals were observed in

568   continuous regenerative cells at the basal layers of the skin, sweat ducts, and hair follicles, serv-

569   ing as a positive control. A few KIT$^+$ and PGP9.5$^+$ sensory fibers were observed around hair fol-

570   licles (Fig. S15G-I). Consistent with our RNA-seq results, KIT$^+$ sensory fibers were also positive

571   for CGRP and NEFH (Fig. S15J-O), suggesting that they were peptidergic A fibers. In short,

572   these results support that our high-fidelity hDRG neuron single-soma transcriptome dataset is

573   useful for selecting specific molecular markers to label and visualize different types of peripheral

574   somatosensory afferents.

575

576

577   **A novel strategy to study functions of human somatosensory afferents using single-soma**

578   **deep RNA-seq dataset-informed microneurography recordings**

579   Molecules profiles and physiological properties of hDRG neurons are intrinsically linked: the

580   expressed molecules form the physical basis for physiological properties. Physiological record-

581   ings of human somatosensory afferents have been conducted for more than half a century, which

582   generated important parts of our knowledge on the human somatosensory system. Physiological

583   properties of somatosensory afferents generate an important foundation for informing molecular

584   cloning of critical receptors, such as TRPV1 and PIEZO2. Due to a lack of clear molecular com-

585   positions of human somatosensory afferents, the reverse experimental direction, from molecular

586   profiles of human sensory afferents to inform physiological recordings, has not been possible.

587   With the single-soma deep RNA-seq dataset, we were finally in a position to do so. Here we fo-

588   cused on two populations of human sensory afferents to demonstrate the feasibility and power of

589   using single-soma deep RNA-seq informed microneurography recordings to study the function

590   of human afferent subtypes.

591

592   As mentioned previously, the hPEP.KIT population was highly correlated to the mouse

593   PEP3/CGRP-η cluster[15,18] (Fig. 2A-C), a population of TRPV1-negative A-fiber fast-conducting

594   hair pull-sensitive mechano-nociceptors[86]. Our single-soma deep RNA-seq data revealed that

595   hPEP.KIT neurons express NEFH and PIEZO2, but not TRPV1 (Fig. 7A), suggesting that they

596   were A-fiber mechano-nociceptors without heat/capsaicin sensitivity. Interestingly, our dataset

597   also revealed that hPEP.KIT neurons expressed a low level of *TRPM8* (Fig. 7A), a feature that

598   has not been reported in mice. Multiplex FISH validated the co-expression of *KIT*, *PIEZO2* and

599   *TRPM8* in hDRG neurons (Fig. S16A). Altogether, these molecular profiles suggest that

600   hPEP.KIT neurons may be A-fiber HTMRs, responsive to cooling but not heating. Thus, we hy-

601   pothesized that there exist some previously uncharacterized fast-conducting A-fiber HTMRs in

602   the human skin, which can be activated by cooling but not heating stimuli.

603

604   To test this idea, using the *in vivo* electrophysiological technique of microneurography, single-

605   unit axonal recordings were performed from the radial, antebrachial, and peroneal nerves of

606   healthy participants (Fig. 7B-D). The A-HTMRs (n=10) were identified by their insensitivity to

607   soft-brush stroking while responding to a rough brush and displaying high indentation thresholds

608   ($\geq$4 mN); further, they had Aβ-range conduction velocities (>30 m/s, Fig. 7B-F). Remarkably, a

609   subtype of these heat-insensitive A-HTMRs (n=5) units responded to cooling (Fig. 7G-H). Com-

610   pared to mechanically evoked responses, those evoked by cooling although relatively modest but

611   reproducible with tested in triplicates for each recording (Fig. S16B). Furthermore, these cool-

612   ing-evoked responses persisted during the sustained phase (Fig. S16B). These observations con-

613   firmed our prediction that some human cutaneous A-HTMRs are cold-sensitive but heat-

614   insensitive. Based on the transcriptome dataset and current knowledge about molecular receptors

615   for heat and cold, hPEP.KIT seems to be the only population that has the molecular basis for this

616   distinct combination of physiological properties. Thus, we propose that the molecularly defined

617 hPEP.KIT population is correlated to the physiologically defined A-HTMRs with cold but not

618 heat sensitivity (Fig. 7I).

619

620 Human C-LTMRs are readily found during microneurography recordings from the upper limb

621 including the distal regions[87], but they seem to sparsely innervate the distal lower limb[88]. This is

622 consistent with our sequencing results wherein the hC.LTMRs constituted a small population of

623 neurons in lower thoracic and lumbar level DRGs (Fig. 1G). Unexpectedly, our sequencing data

624 revealed that human hC.LTMRs had almost no expression of the cold and menthol receptor

625 *TRPM8* (Fig. 8A) though C.LTMRs display sensitivity to cooling in both humans and mice[89,90].

626 Another unexpected finding was that hC.LTMRs expressed *TRPV1* (Fig. 8A). Multiplex FISH

627 confirmed that $CASQ2^+$ hC.LTMRs were TRPM8$^-$ but TRPV1$^+$ (Fig. S17A). This expression

628 pattern suggested that hC.LTMRs might respond to heating and capsaicin, a novel physiological

629 property that has not been discovered in rodent or non-human primate models, but is predicted

630 not to respond to menthol. In human microneurography, the C-LTMRs (n=11) were identified by

631 their soft-brush sensitivity, low indentation threshold (in this case, ≤0.7 mN), and slow conduc-

632 tion velocity (~1 m/s, n=11, Fig. 7B-C, 8B-D). Consistent with mouse C-LTMRs, these human

633 counterparts responded to hair movement (Fig. 8D) and dynamic cooling (Fig. 8E). In two of

634 them, after having confirmed the cooling response, we applied menthol to their individual recep-

635 tive fields resulting in a cold sensation, but the recorded C-LTMRs, while still responsive to me-

636 chanical and thermal stimuli, were not activated by menthol (Fig. 8F). Remarkably,

637 microneurography recordings showed that a subset of human C-LTMRs responded to dynamic

638 heating (5 out of 11 units or 45%, example shown in Fig. 8G). This proportion of heat-sensitive

639 C-LTMRs aligns with findings in the rabbit (8 out of 18 units or 44%)[91]. In three C-LTMRs,

28

640 after having confirmed the heating response, we applied capsaicin to their receptive fields. Con-

641 sistent with the sequencing results, all three were activated by capsaicin (Fig. 8H). Relative to

642 mechanical stimulation, the C-LTMR responses to dynamic temperature changes were compara-

643 tively modest. However, these responses were reproducible, having been tested in triplicates for

644 each modality and recording (Fig. S16B-C). Interestingly, the C-LTMR has different cooling

645 response properties compared to A-HTMR cooling+ fibers. The cooling-evoked responses rapid-

646 ly diminished in C-TLMR but persisted during the sustained phase in A-HTMR cooling+ fibers.

647 The existence of a polymodal (mechano-heat-cold) C-LTMR type is novel and confirms the se-

648 quencing predictions. Furthermore, the cooling response must be mediated through a non-

649 TRPM8 mechanism.

650

651 For comparison, our recordings also identified C-HTMRs (n=11) that did not respond to soft-

652 brush stroking and hair movement (they responded to a rough brush), had high indentation

653 thresholds (in this case, ≥10 mN), and slow conduction velocities (~1 m/s, Fig. 7B-D & S17B-E).

654 Based on their temperature responses (tested in 9 units), a mix of C-fiber mechano-heat (n=6), C-

655 fiber mechano-cold (n=2), and C-fiber polymodal (mechano-heat-cold, n=1) subtypes were iden-

656 tified. An example of heating and capsaicin responses of a C-fiber mechano-heat nociceptor is

657 shown in Fig. S17F-G. Collectively, these results highlight the accuracy and utility of our single-

658 soma deep RNA-seq dataset of hDRG neurons in guiding and informing the functional character-

659 ization of human somatosensory afferents and the power of combining the two approaches. We

660 believe that this signifies a promising new research direction to link molecular and physiological

661 types and to discover novel functional properties of human somatosensory afferents.

662

663

Discussion

Despite the high relevance to human health, 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 recovered more than 9000 unique transcripts per neuron over 1000 adult hDRG neurons. The sequenced hDRG neurons were clustered into 16 molecular groups, displaying similarities to and remarkable differences from macaque and mouse DRG neurons. As exemplified in the study, our dataset provides important and novel insights into human pain and itch sensory phenomena, explains mechanisms of drug effects, and represents a rich resource to identify new molecular targets for modulating the activity of itch- and pain-sensing primary afferents. Moreover, our dataset successfully guided histological labeling (Fig. S15) and physiological recording (Fig. 7, 8 & S17) of different types of human sensory afferents, leading to discovery of their new anatomical features and physiological properties, and serving as a common ground to connect them all together (Fig. 7I).

677

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. Strategies, including spatial transcriptomics and single-nucleus RNA-seq[22,23,92], 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. Importantly, though nuclear and soma transcripts are intrinsically linked and overlapping to some extent, the soma contains much more transcripts both quan-

30

686   titatively and qualitatively. Soma transcripts are also one step closer than nuclear transcripts for

687   functions. Thus, soma transcripts are preferable, if available, for cell type clustering and func-

688   tional interpretation. In this study, we developed a new strategy by combining fresh frozen

689   hDRG tissues, cryo-section, laser capture microdissection of individual neuronal soma, and

690   Smart-seq2 deep sequencing. Fresh frozen tissue and cryosection techniques minimized

691   transcriptomic changes during the tissue transportation and single-cell isolation process. LCM

692   allowed isolation of DRG neuronal soma with minimal contamination from surrounding non-

693   neuronal cells, while preserving information about cellular morphology, and the Smart-seq2 pro-

694   tocol enabled a high recovery rate of mRNA molecules. Although LCM has been used for isolat-

695   ing a group of neurons for RNA-seq[6,7] or single TG neurons for genomic DNA analysis[93], the

696   successful application of LCM for single-cell transcriptomic analysis has not been achieved be-

697   fore. Thus, we have established a new method for single-cell RNA-seq of adult hDRG neurons.

698   Our approach should be readily applicable to other human neurons with large soma sizes, such as

699   other peripheral ganglia neurons, motor neurons, etc.

700

701   Different single-cell RNA-seq approaches, including single-nucleus RNA-seq[21,23,92], spatial

702   transcriptomics[22], and our LCM-based single-soma RNA-seq, have generated four datasets of

703   transcriptome profiles and cell type clusters of hDRG neurons. These datasets and results overlap

704   to some extent but also exhibit some significant differences (an example shown in Fig. S4). The

705   observed differences are likely caused by both biological and technical features associated with

706   each method. Given the high sequencing depth of transcripts from the neuronal soma, our ap-

707   proach is powerful for molecular discovery, especially for functional molecules expressed at a

708   low level. For example, our approach detected the specific expression of *MRGPRX1 and*

31

709 *MRGPRX4*, two important pruritogen-sensing GPCRs, in hNP1 and/or hNP2 neurons, while the

710 previous datasets barely or not detected these transcripts[22,23]. Our analysis of the expression of

711 receptors, ion channels, and neuropeptides, in human itch-sensing DRG neurons have also identi-

712 fied a set of potential new targets for modulating activities of these sensory afferents (Fig. 6).

713

714 In addition, our sequencing and microneurography results raise many interesting questions re-

715 garding the molecular receptors and afferent types involved in human cold and mechanical pain

716 sensation. For example, human hC.LTMRs rely on a non-TRPM8 cold receptor or another

717 mechanism for their cooling sensitivity[88]. We noticed a low expression level of *TRPA1*, which

718 also has cold sensitivity[94,95], within this population (Fig. S7E). Whether *TRPA1* or some current-

719 ly unknown cold receptors mediate cooling sensitivity in this population will be of interest for

720 future research. The discovery of the hPEP.KIT population indicated a potential role for PIEZO2

721 in human mechano-nociception. This population also responded to cooling, which is a curious

722 property of a large-diameter myelinated nociceptor. Our discovery of TRPM8 expression in the-

723 se neurons provides a molecular explanation for this unique property. In addition, a likely mo-

724 lecular type for C-HTMRs, which display a mixture of responsiveness to mechanical forces and

725 temperature, is the hPEP.TRPV1/A1.1 population. Since hPEP.TRPV1/A1.1 neurons have no

726 expression of *PIEZO2* or *PIEZO1* (Fig. S9A), a non-*PIEZO* mechanoreceptor may exist in these

727 neurons for mediating mechanical pain sensation. This is consistent with reports that patients

728 with *PIEZO2* loss-of-function mutations still have normal mechanical pain threshold and sensi-

729 tivity[68,96]. One suggested mechanical pain channel, TMEM120A (TACAN), is broadly expressed

730 in all types of hDRG neurons (Fig. S9A). This expression pattern does not support its purported

731 role as a mechanical pain receptor in humans. Both C-LTMRs and A-HTMRs responded to cool-

32

732 ing in human microneurography. While the responses were comparatively modest in contrast to

733 mechanically evoked responses, they were consistent on a trial-to-trial basis. This suggests that

734 the requisite circuit may already be in place, which could potentially have implications for un-

735 derstanding thermal hypersensitivities in pathological states. Indeed, there is indirect evidence

736 from human psychophysics and targeted pharmacology, pointing to the role of C-LTMRs in me-

737 diating acute cold allodynia[97] and the role of A fibers in signaling chronic cold allodynia[98]. In

738 short, our dataset serves as an important atlas for understanding sensory properties of hDRG af-

739 ferents and somatosensation more generally.

740

741 Insights into the relationship of the neuronal types between the mouse and human are essential,

742 as such knowledge is important for translation of pre-clinical discoveries and for inferring func-

743 tions of human sensory afferents. The three different analysis strategies used in this study com-

744 plement each other as they have different strength and weakness (see Methods for details). To

745 our satisfaction, despite the different advantages and disadvantages, results from these methods

746 were highly consistent with each other (Fig. 2 & S5). Our results suggested that many broader

747 functional groups of DRG sensory afferents are conserved across species, despite noticeably mo-

748 lecular differences (Fig. 2 & S6). The greatest divergence between mouse and hDRG neurons

749 was observed among C- and A-fiber nociceptors. Mice contain four C-fiber nociceptors (PEP1

750 neuron subtypes), two expressing TRPV1 but not TRPA1 or PIEZO2 (PEP1.1 and PEP1.2) and

751 another two which expresses all three channels (PEP1.3 and PEP1.4) (Fig. 2A & 2C). While not

752 much is known about mouse PEP1.2/CGRP beta, PEP1.1 represents a noxious heat sensory type,

753 terminating with free nerve endings in epidermis of hairy skin[99]. PEP1.3/CGRP gamma neurons

754 innervate mainly internal organs and are a silent nociceptor type, becoming active during in-

33

755   flammation[100]. No functional studies have been performed on PEP1.4 in the mouse. These neu-

756   rons are speculated to represent C-HTMRs which are functionally known to exist in the mouse[101].

757   In humans, it seems that this diversity of C-fiber nociceptors has conflated into two types, the

758   hPEP.TRPV1/A1.1 and hPEP.TRPV1/A1.2 (Fig. 2A & 2C), expressing both TRPV1/TRPA1

759   with or without PIEZO2, respectively. If similar to mice, different neuronal populations in hu-

760   man are somewhat specialized according localization of peripheral target, one of these types may

761   represent deeply innervating silent nociceptors, known to be present in human[102], and the other

762   involved in slow noxious thermal and burning cutaneous sensation and perhaps joint pain. One

763   thing to note is that our dataset contains transcriptomes of hDRG neurons from the lower thorac-

764   ic and lumbar levels, which mainly convey sensations from the leg and lower back. Thus, it re-

765   mains an open question whether humans have a population of internal organ-specific C-

766   nociceptors, like the mouse PEP1.3 cluster. This question will be resolved when sequencing

767   more hDRG neurons from the thoracic level.

768

769   For A-fiber nociceptors, only two types have been identified in mice. The PEP2/CGRP-zeta

770   mechano-heat nociceptive population expresses PIEZO2 and low levels of TRPV1[16,33], con-

771   sistent with heat activation only at very high temperatures[99], and the PEP3/CGRP-eta mechano-

772   nociceptive conveying fast and sharp pain, including hair pulling[44,103]. In this study, we identified

773   five types A-fiber nociceptors in hDRG neurons (Fig. 1G, 2A & 2C). Two populations have

774   clearly conserved features to mouse: hPEP.CHRNA7 showed molecular similarities to mouse

775   and macaque PEP2/CGRP-ζ, and hPEP.KIT to mouse and macaque PEP3/CGRP-η. The other

776   three populations displayed much greater divergence (i.e. hPEP.NTRK3, hPEP.PIEZO^h, and

777   hPEP.0). The hPEP.PIEZO^h cluster is particularly interesting since these neurons express very

34

778   high levels of *PIEZO2* and *PIEZO1,* unlike any neuron types in the mouse. Our results suggest

779   that fast pain and interoception might have been an evolutionary preference in humans as com-

780   pared to rodents, an effect that may relate to significantly increased human body sizes.

781

782   Finally, in this study, we sequenced over 1000 DRG neurons from 6 DRGs at the thoracic (T11-

783   T12) and lumbar (L2-L5) levels of three Caucasian human donors with no obvious somatosenso-

784   ry or systematic diseases and substance uses (Supplementary tables 1 & 2). It is about ~150 to

785   200 neurons per DRG, which represent 1-2% of the total neurons within a hDRG. Thus, sequenc-

786   ing more neurons might give an improved representation. The existence of different types of RA

787   and SA Aβ-LTMRs and different types of proprioceptors in humans is well known, but our cur-

788   rent dataset did not have the resolution yet to separate them (Fig. 1G). With increased number of

789   neurons sequenced, we anticipate discovering an even greater heterogeneity among hDRG neu-

790   rons. In addition, DRGs at different spinal levels innervate different peripheral target tissues[29].

791   Sequencing DRG neurons from different spinal levels will help to uncover molecular and cellu-

792   lar mechanisms underlying physiological functions. Moreover, increased sampling from addi-

793   tional donors representing different demographics would also be critical for investigating sex,

794   race, and age-related differences. Last but not the least, pathological conditions greatly alter the

795   transcriptomic landscape[104,105]. Systematic comparison of molecular and cellular changes be-

796   tween donors at the baseline condition (like the screening criteria of this study) and those with

797   chronic itch or chronic pain would be of great value, if samples are available, for understanding

798   pathological mechanisms and for identifying molecular targets for effective treatments.

799

800

801

802

803

804

805

806

807

808

809

810

811 **Reference:**

812 1   Nascimento, A. I., Mar, F. M. & Sousa, M. M. The intriguing nature of dorsal root
813   ganglion neurons: Linking structure with polarity and function. *Prog Neurobiol* **168**, 86-
814   103, doi:10.1016/j.pneurobio.2018.05.002 (2018).

815 2   Purves, D. A., George J.; Fitzpatrick, David; Katz, Lawrence C.; LaMantia, Anthony-
816   Samuel; McNamara, James O.; Williams, S. Mark. The Major Afferent Pathway for
817   Mechanosensory Information: The Dorsal Column-Medial Lemniscus System.
818   *Neuroscience. 2nd edition* (2001).

819 3   Azimi, E., Xia, J. & Lerner, E. A. Peripheral mechanisms of itch. *Itch-Management in*
820   *Clinical Practice* **50**, 18-23 (2016).

821 4   Mills, S. E. E., Nicolson, K. P. & Smith, B. H. Chronic pain: a review of its epidemiology
822   and associated factors in population-based studies. *Br J Anaesth* **123**, e273-e283,
823   doi:10.1016/j.bja.2019.03.023 (2019).

824   5   Manglik, A. *et al.* Structure-based discovery of opioid analgesics with reduced side

825   effects. *Nature* **537**, 185-190, doi:10.1038/nature19112 (2016).

826   6   Yosipovitch, G., Rosen, J. D. & Hashimoto, T. Itch: From mechanism to (novel)

827   therapeutic   approaches.   *J   Allergy   Clin   Immunol*   **142**,   1375-1390,

828   doi:10.1016/j.jaci.2018.09.005 (2018).

829   7   Burma, N. E., Leduc□Pessah, H., Fan, C. Y. & Trang, T. Animal models of chronic pain:

830   advances and challenges for clinical translation. *Journal of neuroscience research* **95**,

831   1242-1256 (2017).

832   8   Hill, R. NK1 (substance P) receptor antagonists–why are they not analgesic in humans?

833   *Trends in pharmacological sciences* **21**, 244-246 (2000).

834   9   Haberberger, R. V., Barry, C., Dominguez, N. & Matusica, D. Human Dorsal Root

835   Ganglia. *Front Cell Neurosci* **13**, 271, doi:10.3389/fncel.2019.00271 (2019).

836   10   Middleton, S. J. *et al.* Studying human nociceptors: from fundamentals to clinic. *Brain*

837   **144**, 1312-1335 (2021).

838   11   Yu, H. *et al.* MRGPRX4 is a bile acid receptor for human cholestatic itch. *Elife* **8**, e48431

839   (2019).

840   12   Alemi, F. *et al.* The TGR5 receptor mediates bile acid–induced itch and analgesia. *The*

841   *Journal of clinical investigation* **123**, 1513-1530 (2013).

842   13   Tang, F. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* **6**,

843   377-382, doi:10.1038/nmeth.1315 (2009).

844   14   Hwang, B., Lee, J. H. & Bang, D. Single-cell RNA sequencing technologies and

845   bioinformatics pipelines. *Exp Mol Med* **50**, 96, doi:10.1038/s12276-018-0071-8 (2018).

846   15   Usoskin, D. *et al.* Unbiased classification of sensory neuron types by large-scale single-

37

847   cell RNA sequencing. *Nature neuroscience* **18**, 145-153 (2015).

848   16   Sharma, N. *et al.* The emergence of transcriptional identity in somatosensory neurons.

849   *Nature* **577**, 392-398 (2020).

850   17   Wang, K. *et al.* Single-cell transcriptomic analysis of somatosensory neurons uncovers

851   temporal development of neuropathic pain. *Cell research* **31**, 904-918 (2021).

852   18   Kupari, J. *et al.* Single cell transcriptomics of primate sensory neurons identifies cell

853   types associated with chronic pain. *Nature communications* **12**, 1-15 (2021).

854   19   Stuart, T. *et al.* Comprehensive integration of single-cell data. *Cell* **177**, 1888-1902.

855   e1821 (2019).

856   20   Mattei, D. *et al.* Enzymatic dissociation induces transcriptional and proteotype bias in

857   brain cell populations. *International journal of molecular sciences* **21**, 7944 (2020).

858   21   Jung, M. *et al.* Cross-species transcriptomic atlas of dorsal root ganglia reveals species-

859   specific programs for sensory function. *Nature Communications* **14**, 366 (2023).

860   22   Tavares-Ferreira, D. *et al.* Spatial transcriptomics of dorsal root ganglia identifies

861   molecular signatures of human nociceptors. *Science translational medicine* **14**, eabj8186

862   (2022).

863   23   Nguyen, M. Q., von Buchholtz, L. J., Reker, A. N., Ryba, N. J. & Davidson, S. Single-

864   nucleus transcriptomic analysis of human dorsal root ganglion neurons. *Elife* **10**, e71752

865   (2021).

866   24   Bhuiyan, S. *et al.* Harmonized cross-species cell atlases of trigeminal and dorsal root

867   ganglia. *bioRxiv*, 2023.2007. 2004.547740 (2023).

868   25   Ding, J. *et al.* Systematic comparison of single-cell and single-nucleus RNA-sequencing

869   methods. *Nature biotechnology* **38**, 737-746 (2020).

870 26 Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nature protocols* **9**, 171-181 (2014).

872 27 Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587. e3529 (2021).

874 28 Abraira, V. E. & Ginty, D. D. The sensory neurons of touch. *Neuron* **79**, 618-639, doi:10.1016/j.neuron.2013.07.051 (2013).

876 29 Luo, L. *PRINCIPLES OF NEUROBIOLOGY*. 265-267 (Garland Science, 2016).

877 30 Chen, Y.-C. *et al.* Transcriptional regulator PRDM12 is essential for human pain perception. *Nature genetics* **47**, 803-808 (2015).

879 31 Barkas, N. *et al.* Joint analysis of heterogeneous single-cell RNA-seq dataset collections. *Nature methods* **16**, 695-698 (2019).

881 32 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).

883 33 Zeisel, A. *et al.* Molecular architecture of the mouse nervous system. *Cell* **174**, 999-1014. e1022 (2018).

885 34 Hippenmeyer, S. *et al.* A developmental switch in the response of DRG neurons to ETS transcription factor signaling. *PLoS biology* **3**, e159 (2005).

887 35 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).

889 36 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).

892 37 Abdo, H. *et al.* Dependence on the transcription factor Shox2 for specification of sensory

893   neurons conveying discriminative touch. *European Journal of Neuroscience* **34**, 1529-
894   1541 (2011).

895   38   Levanon, D. *et al.* The Runx3 transcription factor regulates development and survival of
896   TrkC dorsal root ganglia neurons. *The EMBO journal* **21**, 3454-3463 (2002).

897   39   Chen, C.-L. *et al.* Runx1 determines nociceptive sensory neuron phenotype and is
898   required for thermal and neuropathic pain. *Neuron* **49**, 365-377 (2006).

899   40   Vermeiren, S., Bellefroid, E. J. & Desiderio, S. Vertebrate sensory ganglia: common and
900   divergent features of the transcriptional programs generating their functional
901   specialization. *Frontiers in Cell and Developmental Biology* **8**, 587699 (2020).

902   41   Liu, Q. & Dong, X. The role of the Mrgpr receptor family in itch. *Pharmacology of itch*,
903   71-88 (2015).

904   42   Meixiong, J., Vasavda, C., Snyder, S. H. & Dong, X. MRGPRX4 is a G protein-coupled
905   receptor activated by bile acids that may contribute to cholestatic pruritus. *Proc Natl*
906   *Acad Sci U S A* **116**, 10525-10530, doi:10.1073/pnas.1903316116 (2019).

907   43   Namer, B. *et al.* Separate peripheral pathways for pruritus in man. *Journal of*
908   *neurophysiology* **100**, 2062-2069 (2008).

909   44   Dong, P. *et al.* TRPC3 is dispensable for β-alanine triggered acute itch. *Scientific reports*
910   **7**, 1-12 (2017).

911   45   Rehfeld, J. F. Cholecystokinin and the hormone concept. *Endocrine Connections* **10**,
912   R139-R150 (2021).

913   46   McNeil, B. & Dong, X. Peripheral mechanisms of itch. *Neuroscience bulletin* **28**, 100-
914   110 (2012).

915   47   Colburn, R. W. *et al.* Attenuated cold sensitivity in TRPM8 null mice. *Neuron* **54**, 379-

916   386 (2007).

917   48 Bautista, D. M. *et al.* The menthol receptor TRPM8 is the principal detector of
918   environmental cold. *Nature* **448**, 204-208 (2007).

919   49 Dhaka, A. *et al.* TRPM8 is required for cold sensation in mice. *Neuron* **54**, 371-378
920   (2007).

921   50 Samour, M. S., Nagi, S. S. & Mahns, D. A. Cav3. 2-expressing low-threshold C fibres in
922   human hairy skin contribute to cold allodynia—a non-TRPV1-and non-TRPM8-
923   dependent phenomenon. *Pain* **156**, 1566-1575 (2015).

924   51 Campero, M., Baumann, T., Bostock, H. & Ochoa, J. Human cutaneous C fibres
925   activated by cooling, heating and menthol. *The Journal of physiology* **587**, 5633-5652
926   (2009).

927   52 Caterina, M. J. *et al.* The capsaicin receptor: a heat-activated ion channel in the pain
928   pathway. *Nature* **389**, 816-824 (1997).

929   53 Bandell, M. *et al.* Noxious cold ion channel TRPA1 is activated by pungent compounds
930   and bradykinin. *Neuron* **41**, 849-857 (2004).

931   54 Defaye, M. *et al.* Gut-innervating TRPV1+ neurons drive chronic visceral pain via
932   microglial P2Y12 receptor. *Cellular and Molecular Gastroenterology and Hepatology*
933   **13**, 977-999 (2022).

934   55 Yang, F.-C. *et al.* Genetic control of the segregation of pain-related sensory neurons
935   innervating the cutaneous versus deep tissues. *Cell reports* **5**, 1353-1364 (2013).

936   56 Liu, S. *et al.* A neuroanatomical basis for electroacupuncture to drive the vagal–adrenal
937   axis. *Nature* **598**, 641-645 (2021).

938   57 Albrecht, P. J. *et al.* Pathologic alterations of cutaneous innervation and vasculature in

939   affected limbs from patients with complex regional pain syndrome. *Pain* **120**, 244-266

940   (2006).

941 58 Morelli, C. *et al.* Identification of a population of peripheral sensory neurons that

942   regulates blood pressure. *Cell reports* **35**, 109191 (2021).

943 59 Albrecht, P. J. *et al.* Excessive peptidergic sensory innervation of cutaneous arteriole–

944   venule shunts (AVS) in the palmar glabrous skin of fibromyalgia patients: Implications

945   for widespread deep tissue pain and fatigue. *Pain medicine* **14**, 895-915 (2013).

946 60 Bowsher, D. *et al.* Absence of pain with hyperhidrosis: a new syndrome where vascular

947   afferents may mediate cutaneous sensation. *PAIN®* **147**, 287-298 (2009).

948 61 Xu, J. *et al.* GPR68 senses flow and is essential for vascular physiology. *Cell* **173**, 762-

949   775. e716 (2018).

950 62 Zeng, W.-Z. *et al.* PIEZOs mediate neuronal sensing of blood pressure and the

951   baroreceptor reflex. *Science* **362**, 464-467 (2018).

952 63 Marshall, K. L. *et al.* PIEZO2 in sensory neurons and urothelial cells coordinates

953   urination. *Nature* **588**, 290-295 (2020).

954 64 Perlman, R. L. Mouse models of human disease: an evolutionary perspective. *Evolution,*

955   *medicine, and public health* **2016**, 170-176 (2016).

956 65 Willimann, B. & Trüeb, R. M. Hair pain (trichodynia): frequency and relationship to hair

957   loss and patient gender. *Dermatology* **205**, 374-377 (2002).

958 66 von Buchholtz, L. J. *et al.* Decoding cellular mechanisms for mechanosensory

959   discrimination. *Neuron* **109**, 285-298. e285 (2021).

960 67 Wolfson, R. L. *et al.* DRG afferents that mediate physiologic and pathologic

961   mechanosensation from the distal colon. *Cell* **186**, 3368–3385 (2023).

962   68   Nagi, S. S. *et al.* An ultrafast system for signaling mechanical pain in human skin.
963   *Science advances* **5**, eaaw1297 (2019).

964   69   Schmidt, R. F. *Encyclopedia of Pain*. pp 2171–2173 (Spriner 2007).

965   70   Löken, L. S., Wessberg, J., McGlone, F. & Olausson, H. Coding of pleasant touch by
966   unmyelinated afferents in humans. *Nature neuroscience* **12**, 547-548 (2009).

967   71   Olausson, H., Wessberg, J., McGlone, F. & Vallbo, Å. The neurophysiology of
968   unmyelinated tactile afferents. *Neuroscience & Biobehavioral Reviews* **34**, 185-191
969   (2010).

970   72   Huzard, D. *et al.* The impact of C-tactile low-threshold mechanoreceptors on affective
971   touch and social interactions in mice. *Science Advances* **8**, eabo7566 (2022).

972   73   Xue, Y., Chidiac, C., Herault, Y. & Gaveriaux-Ruff, C. Pain behavior in SCN9A (Nav1.
973   7) and SCN10A (Nav1. 8) mutant rodent models. *Neuroscience Letters* **753**, 135844
974   (2021).

975   74   Duan, G. *et al.* A SCN10A SNP biases human pain sensitivity. *Molecular pain* **12**,
976   1744806916666083 (2016).

977   75   Reddy, V. B., Iuga, A. O., Shimada, S. G., LaMotte, R. H. & Lerner, E. A. Cowhage-
978   evoked itch is mediated by a novel cysteine protease: a ligand of protease-activated
979   receptors. *Journal of Neuroscience* **28**, 4331-4335 (2008).

980   76   Schmelz, M. *et al.* Chemical response pattern of different classes of C-nociceptors to
981   pruritogens and algogens. *Journal of neurophysiology* **89**, 2441-2448 (2003).

982   77   Dahlhamer, J. *et al.* Prevalence of Chronic Pain and High-Impact Chronic Pain Among
983   Adults - United States, 2016. *MMWR Morb Mortal Wkly Rep* **67**, 1001-1006,
984   doi:10.15585/mmwr.mm6736a2 (2018).

985   78   Tajiri, K. & Shimizu, Y. Recent advances in the management of pruritus in chronic liver
986   diseases. *World J Gastroenterol* **23**, 3418-3426, doi:10.3748/wjg.v23.i19.3418 (2017).

987   79   Dong, X. & Dong, X. Peripheral and central mechanisms of itch. *Neuron* **98**, 482-494
988   (2018).

989   80   Bell, J., McQueen, D. & Rees, J. Involvement of histamine H4 and H1 receptors in
990   scratching induced by histamine receptor agonists in BalbC mice. *British journal of*
991   *pharmacology* **142**, 374-380 (2004).

992   81   Calandre, E. P., Rico-Villademoros, F. & Slim, M. Alpha2delta ligands, gabapentin,
993   pregabalin and mirogabalin: a review of their clinical pharmacology and therapeutic use.
994   *Expert review of neurotherapeutics* **16**, 1263-1277 (2016).

995   82   Peng, C. *et al.* miR-183 cluster scales mechanical pain sensitivity by regulating basal and
996   neuropathic pain genes. *Science* **356**, 1168-1171 (2017).

997   83   Cowan, A., Kehner, G. B. & Inan, S. Targeting itch with ligands selective for κ opioid
998   receptors. *Pharmacology of itch*, 291-314 (2015).

999   84   Ko, M. H., Song, M., Edwards, T., Lee, H. & Naughton, N. The role of central μ opioid
1000   receptors in opioid-induced itch in primates. *Journal of Pharmacology and Experimental*
1001   *Therapeutics* **310**, 169-176 (2004).

1002   85   Chu, C., Artis, D. & Chiu, I. M. Neuro-immune interactions in the tissues. *Immunity* **52**,
1003   464-474 (2020).

1004   86   Ghitani, N. *et al.* Specialized mechanosensory nociceptors mediating rapid responses to
1005   hair pull. *Neuron* **95**, 944-954. e944 (2017).

1006   87   Xu, S. *et al.* Mechanoreceptive Aβ primary afferents discriminate naturalistic social touch
1007   inputs at a functionally relevant time scale. *bioRxiv*, 2023.2007. 2022.549516 (2023).

1008   88   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).

1011   89   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).

1014   90   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).

1016   91   Shea, V. & Perl, E. Sensory receptors with unmyelinated (C) fibers innervating the skin of the rabbit's ear. *Journal of neurophysiology* **54**, 491-501 (1985).

1018   92   Yang, L. *et al.* Human and mouse trigeminal ganglia cell atlas implicates multiple cell types in migraine. *Neuron* **110**, 1806-1821. e1808 (2022).

1020   93   Wang, K., Lau, T. Y., Morales, M., Mont, E. K. & Straus, S. E. Laser-capture microdissection: refining estimates of the quantity and distribution of latent herpes simplex virus 1 and varicella-zoster virus DNA in human trigeminal Ganglia at the single-cell level. *Journal of virology* **79**, 14079-14087 (2005).

1024   94   Bautista, D. M. *et al.* TRPA1 mediates the inflammatory actions of environmental irritants and proalgesic agents. *Cell* **124**, 1269-1282 (2006).

1026   95   Macpherson, L. J. *et al.* Noxious compounds activate TRPA1 ion channels through covalent modification of cysteines. *Nature* **445**, 541-545 (2007).

1028   96   Chesler, A. T. *et al.* The role of PIEZO2 in human mechanosensation. *New England Journal of Medicine* **375**, 1355-1364 (2016).

1030   97   Shaikh, S., Nagi, S. S., McGlone, F. & Mahns, D. A. Psychophysical investigations into

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

98   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).

99   Qi, L. *et al.* A DRG genetic toolkit reveals molecular, morphological, and functional diversity of somatosensory neuron subtypes. *bioRxiv*, 2023.2004. 2022.537932 (2023).

100   Prato, V. *et al.* Functional and molecular characterization of mechanoinsensitive "silent" nociceptors. *Cell reports* **21**, 3102-3115 (2017).

101   Boada, M. D. & Woodbury, C. J. Physiological properties of mouse skin sensory neurons recorded intracellularly in vivo: temperature effects on somal membrane properties. *Journal of neurophysiology* **98**, 668-680 (2007).

102   Buffington, C. T. Visceral pain in humans: lessons from animals. *Current pain and headache reports* **5**, 44-52 (2001).

103   Arcourt, A. *et al.* Touch receptor-derived sensory information alleviates acute pain signaling and fine-tunes nociceptive reflex coordination. *Neuron* **93**, 179-193 (2017).

104   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).

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

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

1054   107   Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21
1055   (2013).

1056   108   Hu, Y. *et al.* Neural network learning defines glioblastoma features to be of neural crest
1057   perivascular or radial glia lineages. *Science Advances* **8**, eabm6340 (2022).

1058   109   Shrikumar, A., Greenside, P. & Kundaje, A. in *International conference on machine*
1059   *learning.* 3145-3153 (PMLR).

1060   110   Huynh-Thu, V. A., Irrthum, A., Wehenkel, L. & Geurts, P. Inferring regulatory networks
1061   from expression data using tree-based methods. *PloS one* **5**, e12776 (2010).

1062   111   Aibar, S. *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nature*
1063   *methods* **14**, 1083-1086 (2017).

1064   112   Vallbo, A., Olausson, H., Wessberg, J. & Kakuda, N. Receptive field characteristics of
1065   tactile units with myelinated afferents in hairy skin of human subjects. *The Journal of*
1066   *physiology* **483**, 783-795 (1995).

1067   113   Vallbo, Å., Olausson, H. & Wessberg, J. Unmyelinated afferents constitute a second
1068   system coding tactile stimuli of the human hairy skin. *Journal of neurophysiology* **81**,
1069   2753-2763 (1999).

1070   114   Wasner, G., Schattschneider, J., Binder, A. & Baron, R. Topical menthol—a human
1071   model for cold pain by activation and sensitization of C nociceptors. *Brain* **127**, 1159-
1072   1171 (2004).

1073

1074

1075

1076

1077

1078

1079

1080

1081

1082

1083

1084

1085

1086

1087

1088

1089

1107

1108

1109

1110

1111

1112

1113 **Materials and Methods:**

1114 *Human tissues and subjects*

1115 Human DRG tissues were procured from National Disease Research Interchange (NDRI). The

1116 research application was approved by the NDRI Feasibility Committee (RLUW1 01). Six DRGs

1117 between T11 to L5 of three human donors aged from 23- 61 years were used in this study. The

1118 dissected DRG tissues from human donors were immediately imbedded in OCT, shipped to the

1119 Luo lab on dry ice and stored in -80 □ until use. The information of DRGs and dis-identified

1120 donors and screening criteria are summarized in the Supplementary Tables 1 & 2. As determined

1121 by the University of Pennsylvania IRB, this study was exempted from the human subject re-

1122 quirements for the Luo lab.

1123

1124   The human skin biopsies were extracted from three healthy donors at the college of medicine,

1125   University of Florida. This tissue procurement was approved by the university IRB (protocols

1126   IRB201500232 and IRB202300291). The information of human skin biopsies and dis-identified

1127   donors is summarized in the Supplementary Table 3. These three donors are members of one

1128   family and have no noticeable abnormal somatosensation or peripheral neuropathy. All partici-

1129   pants were provided written informed consent and signed the document.

1130

1131   *In vivo* recordings of peripheral sensory afferents of healthy human subjects were performed at

1132   Linköping University, Sweden. These subjects were recruited through social media. All partici-

1133   pants were provided written informed consent before the start of the experiment. The study was

1134   approved by the ethics committee of Linköping University (dnr 2020-04426) and complied with

1135   the revised Declaration of Helsinki.

1136

1137   *Laser capture microdissection of hDRG neurons*

1138   The hDRGs imbedded in OCT were cryosectioned (Cryostat Leica Cm1950) into 20 μm sections

1139   and mounted onto Arcturus PEN Membrane Frame Slides (Applied biosystems, LCM0521). One

1140   of every five consecutive sections was collected for laser capture microdissection to avoid re-

1141   peated dissection of the soma from the same neuron in different sections. The slides were stored

1142   in -80 ☐ until further use.

1143

1144   On the day of laser capture microdissection, the slides were transferred to the SBDRC laser cap-

1145   ture microdissection (LCM) core on dry ice. Before dissection, the section was briefly stained

1146   with RNase free Arcturus™ Histogene™ staining solution (Applied biosystems, 12241-05) for

1147   better visualization of neuronal soma: 70% cold EtOH for 30s; Histogene™ staining for 10s; 70%

1148   cold EtOH for 30s; 95% cold EtOH for 30s; 100% cold EtOH for 30s and air dry for 2min. Then,

1149   the slide was put onto laser capture microdissection microscope system (Leica LMD6) for the

1150   neuronal soma dissection. The laser was calibrated, and the laser intensity was adjusted to

1151   achieve best dissection efficiency. The dissected individual neuronal soma was collected in the

1152   cap of a 200 µl PCR tube containing 4 µl lysis buffer[26]. The sequencing library was generated

1153   following Smart-seq2 workflow[26]. The libraries passing through all quality controls were select-

1154   ed for the final sequencing.

1155

1156   *Sequencing and sequence alignment*

1157   The libraries were pooled together (384 samples for one batch) and sequenced on NovaSeq 6000

1158   platform by the Children's Hospital of Philadelphia (CHOP) Center for Applied Genomics

1159   (CAG). Raw sequencing data was demultiplexed with bcl2fastq2 v.2.20 (Illumina) followed by

1160   Tn5 transposon adapter sequences trimming with Cutadapt[106]. The processed reads were then

1161   aligned to human genome (GRCh38 GENCODE as the reference genome, and GRCh38.104.

1162   GTF as the annotation) using STAR v.2.7.9a49[107]. STAR quantMode GeneCounts was used to

1163   quantify unique mapped reads to gene counts.

1164

1165   *Analysis of single-soma RNA-seq data of hDRG neurons using Seurat and R*

1166   R (version 4.1.2) and Seurat (version 4.0.5) were used for the single-cell RNA-seq analysis. Six

1167   objects were created from the individual biological replicates. The data were normalized

1168   (NormalizeData) after which 4500 most variable features were selected (FindVariableFeatures).

1169   To mitigate batch effects between replicates, we used Seurat's integrated analysis approach that

1170   transforms datasets into a shared space using pairwise correspondences (or "anchors")[19]. An-

1171   chors were first defined using FindIntegrationAnchors (dims☐=☐1:30) and the data were then

1172   integrated (IntegrateData) and scaled (ScaleData), followed by principal component analysis

1173   (PCA) (RunPCA, npcs☐=☐50). For clustering, the final parameters were: RunUMAP, reduc-

1174   tion☐=☐pca, dims☐=☐1:25; FindNeighbors, reduction☐=☐pca, dims☐=☐1:25; FindClusters,

1175   resolution☐=☐3.4. Highly similar clusters without clearly distinguishable markers were merged

1176   to produce the final 16 clusters.

1177

1178   *Analysis of single-soma RNA-seq data of hDRG neurons using Conos*

1179   For Conos[31] analysis, single-soma hDRG data were integrated using CCA space

1180 $buildGraph(k☐=☐8, k.self=3, space☐=☐"CCA", ncomps=30, n.odgenes=2000, ver-

1181   bose=TRUE, snn=T, snn.k☐=☐10). For human single-soma and human single-nucleus dataset,

1182   co-integration was performed as $buildGraph(k☐=☐8, k.self=3, space☐=☐"CCA",

1183   n.odgenes=2000, verbose=TRUE, snn=T, snn.k☐=☐10). For Conos co-clustering mouse

1184   (Sharms) dataset was downsampled to max 300 cells per cluster, co-integration was performed as

1185 $buildGraph(k☐=☐8, k.self=3, space☐=☐"CCA", n.odgenes=2000, verbose=TRUE,

1186   snn=T,snn.k☐=☐10). Macaque (Kupari, SmartSeq2 dataset) was used for interspecies analysis.

1187   For Conos co-clustering macaque/human graph was built as $buildGraph(k☐=☐4, k.self=3,

1188   space☐=☐"CCA", ncomps = 30, n.odgenes=2000, snn=F, snn.k☐=☐10). For all UMAP plots in

1189   Conos graphs were embeded as: $embedGraph(method☐=☐"UMAP", spread=15, seed = 3).

1190   Label propagation ($propagateLabels) was run using method "diffusion".

1191

Methods used to elucidate cross-species cluster relationships

We used four different interspecies analysis approaches. First, Conos, which uses graph-based dataset integration and was developed to co-cluster and compare datasets originating from different RNA-seq platforms and species (Fig. 2A-B, S5A-B). Second, probabilistic neural network analysis which is a variant of machine learning in which learning module is trained with one dataset and then testing other datasets for pattern recognition and probability output[18] (Fig. S5C-D). Third, neural network based hierarchical clustering analysis (Fig. S5E). In the hierarchical clustering analysis, each query neuron types, either human or macaque, is assigned weights of the sensory-type associated patterns by a neural network, which was trained with gene patterns including both 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. Finally, we performed transcription factor associated gene regulatory network analysis across all three species using genes network modules presumably driven by individual transcription factors (Fig.2D). These three methods have different strengths and weaknesses. Conos finds shared principal components between integrated datasets, but some species-specific features may be lost and lead to impaired statistical sufficiency during integration and furthermore can be affected by the number of principal components and nearest neighbor distance. Machine learning is based on gene expression and is not supervised by shared latent space (i.e. common principal components). Each single reference dataset is used to train the machine learning module and then tested by the other dataset. Thus, an advantage of this method is that at the stage of machine learning, datasets are not integrated, and hence, probability calculations are not affected by principal components or nearest distance. Conversely, its disadvantage is that since it emphasizes the features of each cell type, the learning accuracy and

1215   reliability depend on the robustness of the reference or training dataset. For the third approach,

1216   machine learning-based hierarchical clustering, we extracted the weight of cell type-specific fea-

1217   tures to construct the latent space covering all cell types across species, whether shared or not.

1218   With this strategy, we tried to obtain sufficient latent space, as compared to Conos, by training

1219   and predicting every dataset independently and furthermore, parameterization was used to find

1220   the most robust hierarchical clustering.

1221

1222   *Cell type probabilistic similarity estimation across cell types and the data integration across*

1223   *species*

1224   The assessment of cell type purity, the probabilistic similarity, and cell-type integration across

1225   species are performed using packages in a machine learning based single-cell analysis toolkit-

1226   scCAMEL, released separately at https://sccamel.readthedocs.io/.

1227

1228   *Probabilistic similarity estimation across cell types*

1229   The calculation of cell-type probabilistic score has been described in SWAPLINE package[108].

1230   Briefly, a vanilla neural network model was built for cell-type classification. To train the model,

1231   we removed the cell cycle–related genes and then computed the most variable features. In addi-

1232   tion, we ranked the marker genes for each cell type by two heuristics for the cell-type specificity

1233   of both fold change and enrichment score change. Subsequently, the ranked marker genes and

1234   the most variable genes were merged, log-transformed, and scaled by min-max normalization for

1235   learning models. The frame of the neural network model and the parameters have been described

1236   in the SWAPLINE package. The learning accuracy of the neural-network classifier was inspect-

1237   ed against epoch numbers and was estimated by k-fold cross-validation (k = 3). The learning rate

1238   and learning epochs were selected according to the maximum point of the learning curve reach-

1239   ing the accuracy plateaus. The probabilistic scores from mouse and macaque species against hu-

1240   man reference were visualized in violin plot.

1241

1242   *Data integration across species*

1243   For the integration task, we applied interpretable neural-network learning. First, we took one

1244   dataset from the dataset pool. We trained a neural-network classifier by learning the transcrip-

1245   tional features of each cell-types in this dataset and then calculated the trained cells' probabilistic

1246   scores against all cell-types. Subsequently, we used all other datasets as query datasets and calcu-

1247   lated the probabilistic score of every cell in each query dataset via the trained classifier. Then, we

1248   took another dataset from the dataset pool and repeated the training and prediction. We repeat the

1249   training and prediction till every dataset has been used as a training reference for the predictions.

1250   Here, we consider that the probabilistic score of each cell reflects the weighted gene patterns

1251   representing each trained cell-type. Thus, we merged the probabilistic scores of all cells from all

1252   trained and predicted datasets for the principal component analysis. The most significant princi-

1253   pal components were determined by the elbow method and subsequently used as the latent space

1254   for further downstream analysis. The tree plot was constructed with the parameter of 11 principal

1255   components, 90 nearest neighbors, and correlation metric. The trained cell-type similarity was

1256   calculated with the correlation distance and the average/UPGMA linkage and visualized in the

1257   hierarchical heatmap.

1258

1259   In parallel, we normalized the gene expression by interpretable learning. We transformed the

1260   gene symbols of each species into the nomenclature in Homo sapiens. We estimated the features'

1261   weights in each reference cell-type by using the DeepLift algorithm[109]. The gene expression of

1262   each cell that has been learned or predicted in one trained reference dataset, was inferred by the

1263   matrix multiplication between the features' weights and the cell-type probabilistic scores. And

1264   the final gene*cell expression matrix was calculated by the average of non-empty values across

1265   all datasets. Using this normalized expression matrix, we enriched the mostly co-expressed genes

1266   via spearman correlation. These co-expressed genes were used for inferring the TF associated

1267   gene patterns via a modified GENIE3, as described in[110,111]. The result was visualized as a hier-

1268   archical heatmap.

1269

1270   *Multiplex FISH, confocal microscopy imaging, and quantification*

1271   OCT embedded freshly dissected human lumbar or thoracic DRG tissues were cryosectioned at

1272   20 μm thickness and mounted on glass slides. The slides were stored in −80 °C to preserve

1273   RNA integrity until use. RNAscope Fluorescent Multiplex Reagent Kit and RNASCOPE probes

1274   for the targeted genes (Advanced Cell Diagnostics Inc.) were used for *multiplex FISH*.

1275   RNAscope *in situ* hybridization was performed in accordance with the manufacturer's instruc-

1276   tions. In brief, fresh frozen hDRG sections were fixed, dehydrated, and treated with protease.

1277   The sections were then hybridized with the respective target probe for 2 hr at 40°C, followed by

1278   two to three rounds of signal amplification. The sections were then mounted under coverslips,

1279   sealed with nail polish, and stored in the dark at 4°C until imaged. A Leica SP5 confocal micro-

1280   scope was used to capture images and ImageJ was used for image analysis. In some DRG neu-

1281   rons, accumulation of lipofuscin in part of cells caused strong autofluorescence in all channels.

1282   These signals were considered as non-specific background (labeled by asterisk) were excluded

1283   for analysis. (See Fig. S9 for examples). The percentage of each cluster over all DRG neurons

1284 could be a little bit overestimated due to the following two reasons: 1) Some marker genes or

1285 marker gene combination may also label a small subset of other cell types; 2) An underestima-

1286 tion in quantification of total neuronal numbers because some cells have neither multiple FISH

1287 signals nor DAPI (4′,6-diamidino-2-phenylindole) nucleus staining signals.

1288

1289 *Human skin biopsy extraction, processing, and immunostaining*

1290 Dermal skin punch biopsies were performed as described herein. Briefly, 1 cc of lidocaine was

1291 injected subdermally at each biopsy location (Supplementary table 3). A total of six, 3 mm der-

1292 mal skin punch biopsies were performed on each patient. Excised skin was immediately placed

1293 in 1.5 mL eppendorf tubes containing 4 degree Celsius 4% paraformaldehyde (PFA) solution that

1294 was freshly prepared on the same day of the skin biopsy procedure. Biopsy tissue was fixed in 4%

1295 PFA (dissolved in PBS) for exactly four hours at 4 degrees, followed by 2 X 30 minutes washes

1296 in phosphate buffered saline (PBS) solution, and then cryoprotected using 1XPBS, 30% sucrose

1297 at 4 degrees. These tissues in cold 1XPBS, 30% sucrose were overnight shipped to the lab of

1298 Integrated Tissue Dynamics LLC.

1299

1300 The skin biopsies were mounted in OCT and cryosectioned into 14 μm sections. Adjacent sec-

1301 tions were collected by continuous slides. Immunofluorescence in this study was performed us-

1302 ing combinations of mouse monoclonal anti-human PGP9.5 (Protein Gene Product, CedarLane,

1303 Burlington, Canada, 31A3, [source UltraClone Ltd, Isle of Wight, UK]; 1:200), sheep polyclonal

1304 anti-human CGRP, mouse monoclonal anti-human NEFH (Sigma [ab142]; 1:400), rabbit anti-

1305 human SST (ImmunoStar 20067, Hudson, WI, USA), and anti-human KIT. Slides were

1306 preincubated in 1% bovine serum albumin and 0.3% Triton X-100 in PBS (PBS-TB) for 30

1307 minutes and then incubated with primary antibodies diluted in PBS-TB overnight in a humid

1308 atmosphere at 4°C. Slides were then rinsed in excess PBS for 30 minutes and incubated for 2

1309 hours at room temperature with the appropriate secondary antibodies diluted in PBS-TB. Follow-

1310 ing secondary antibody incubation, the sections were rinsed for 30 minutes in PBS and

1311 coverslipped under 90% glycerol in PBS. Images were collected using a 20X objective on an

1312 Olympus BX51-WI microscope equipped with conventional fluorescence filters (Cy3:528–553

1313 nm excitation, 590–650 nm emission; Cy2/Alexa 488: 460–500 nm excitation, 510–560 nm

1314 emission), a Hamamatsu ER, DVC high-speed camera, linear focus encoder, and a 3-axis motor-

1315 ized stage system interfaced with Neurolucida software (MBF Bioscience, Essex, VT, USA).

1316

1317 *In vivo electrophysiological recording of human peripheral sensory fibers*

1318 Single-unit axonal recordings (microneurography) were performed from the right posterior

1319 antebrachial cutaneous, radial, or superficial peroneal nerve of 41 healthy participants (19 males

1320 and 22 females; 19 to 41 years). Participants were comfortably seated in an adjustable chair with

1321 legs and arms stretched out (and hand pronated), supported by vacuum pillows, and covered in a

1322 blanket if they reported as feeling cold.

1323

1324 Under real-time ultrasound guidance (LOGIQ P9, GE Healthcare, Chicago, IL, USA), the target

1325 nerve was impaled with an insulated tungsten recording electrode (FHC Inc., Bowdoin, ME,

1326 USA). Adjacent to that, an uninsulated reference electrode was inserted just under the skin. A

1327 high-impedance preamplifier (MLT185 headstage) was attached to the skin near the recording

1328 electrode and used together with a low-noise high-gain amplifier (FE185 Neuro Amp EX,

1329 ADInstruments, Oxford, UK). Once the electrode tip was intrafascicular, single LTMRs were

1330 searched for by soft-brush stroking, and single HTMRs were searched for by coarse-brush strok-

1331 ing, pinching, and hair tugging in the fascicular innervation zone while making minute electrode

1332 adjustments.

1333

1334 All recorded afferents were mechanically responsive and divided into subtypes based on estab-

1335 lished criteria[68,112,113]. Mechanical threshold and receptive field size were determined using

1336 Semmes-Weinstein monofilaments (nylon fiber; Aesthesio, Bioseb, Pinellas Park, FL, USA).

1337 Mechanical threshold was defined as the weakest monofilament to which the unit responded to in

1338 at least 50% of trials. Hair deflection was tested with a small pair of forceps, carefully avoiding

1339 skin contact while manipulating the hair. Further, force measurements were performed to ensure

1340 that no skin/hair pulling occurred. Conduction velocity of the recorded afferent was estimated

1341 from latency responses to surface electrical stimulation of the receptive field (FE180 Stimulus

1342 Isolator, ADInstruments, Oxford, UK). Electrically and mechanically evoked spikes were com-

1343 pared on an expanded time scale to confirm they originated from the same unit. Thermal respon-

1344 siveness was tested by placing a Peltier probe (7.4 x 12.2 mm, T09, QST.Lab, Strasbourg, France)

1345 onto the receptive field. After recording baseline activity for at least 30 s (with the thermode in

1346 contact with the receptive field) at a neutral temperature of 30°C, a series of cooling (down to

1347 0°C) and warming (up to 50°C) stimuli were delivered at 30-s intervals. If needed, the thermode

1348 was mounted on a stand for better stability.

1349

1350 To test *TRPV1* expression, capsaicin (Capsina 0.075%, Bioglan AB, Malmö, Sweden) was topi-

1351 cally applied to the receptive field. After 1 minute, the skin was wiped clean, and the emergence

1352 of any spontaneous spiking activity from the recorded afferent was monitored. *TRPM8* expres-

1353   sion was tested by placing an ethanol-soaked gauze pad (90% ethanol as control) onto the recep-

1354   tive field followed by menthol solution (400 mg of 40% L-menthol dissolved in 90% ethanol,

1355   Sigma-Aldrich, Inc., Schnelldorf, Germany[114]). The gauze pad was covered with an adhesive

1356   film to prevent the evaporation of ethanol. After 5 minutes, the skin was wiped clean and the

1357   emergence of any spontaneous spiking activity from the recorded afferent was monitored. Dur-

1358   ing these procedures, we documented the participants' verbal descriptions of what they felt, and

1359   if there was no obvious sensation, the procedure was repeated.

1360

1361   Neural activity was sampled at 20 kHz and recorded using the ADInstruments data acquisition

1362   system (LabChart software v8.1.24 and PowerLab 16/35 hardware PL3516/P, Oxford, UK), then

1363   exported to Spike2 (v10.13, Cambridge Electronic Design Ltd., Cambridge, UK). Recorded ac-

1364   tion potentials were carefully examined offline on an expanded time scale. Threshold crossing

1365   was used to distinguish action potentials from noise with a signal-to-noise ratio of at least 2:1,

1366   and spike morphology was confirmed by template matching. Recordings were discarded if mul-

1367   tiple units were present or if spike amplitudes were not distinct from the noise, preventing secure

1368   action potential identification.

1369

1370   **Figure Generation Software**

1371   Figures were generated in Powerpoint (Microsoft Office) and GraphPad Prism (v9, GraphPad

1372   Software Inc. La Jolla, CA, USA). Some cartoons were made partially in BioRender (BioRender,

1373   2022, RRID:SCR_018361).

1374

1375   **Data availability**

1376   The raw and processed datasets for the single-soma sequencing of hDRG neurons reported in this

1377   study will be deposited into Broad Institute Single cell portal

1378   (https://singlecell.broadinstitute.org/single_cell) once the manuscript is accepted for publication.

1379   Macaque (Kupari) data is available at

1380   https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE165569

1381   Mouse (Zeisel) DRG data is available at

1382   http://loom.linnarssonlab.org/clone/Mousebrain.org.level6/L6_Peripheral_sensory_neurons.loom.

1383   Mouse (Sharma) DRG data is available at

1384   https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE139088

1385

1386   **Code availability**

1387   Any custom code will be deposited to Github once the manuscript is accepted for publication.

1388   All analyses are based on previously published code and software.

1389

1390

1391

1392

1393

1394

1395

1396

1397

1398

1399

1400

1401

1402

1403

1404

1405

1406

1407

1408

1409 **Figure 1. Developing a novel laser capture microdissection (LCM) based approach for sin-**

1410 **gle-soma deep RNA-seq of hDRG neurons**

1411 (**A**) A diagram shows the overall workflow of this study. (left) A cartoon illustrates features as-

1412 sociated with different strategies for single-cell RNA-seq of hDRG neurons. (Middle) One real

1413 example of the laser dissection of a hDRG neuron soma is shown. (Right) Analyses and experi-

1414 ments conducted in this study are summarized. Scale bar, 50 μm (cell) and 500 μm (cap)

1415 (**B**) UMAP plot showing the clusters of 1066 hDRG neurons. (**C-D**) Violin plots showing total

1416 number of detected genes (C) and the expression of neuronal marker *SLC17A6* (D). (**E**) The

1417 grouping clusters based on the soma size, and the expression of *INA*, *NEFH*, *PRDM12*, and

1418 *CALCA*. (**F**) UMAPs showing some canonical marker gene expression in each cluster. (**G**)

1419 UMAP plot with names of each cluster.

1420

1421 **Fig. 2 Cross-species analysis of DRG neurons in human, macaque, and mouse**

1422　(**A, B**) Conos label propagation from mouse (Sharma) (**A,** combined Sharma & Usoskin nomen-

1423　clature) and macaque (B) to hDRG neuron clusters showing the cell type correlation. For

1424　UMAPs for correspondent co-integration, from which these results were inferred, see Fig. S5A-

1425　B. (**C**) Summary of correspondence of DRG neuron clusters among three species. Solid lines

1426　depict clear match, and dashed lines represent partial similarity. (**D**) Heatmap visualization of

1427　cross-species-conserved and species-specific transcription factor associated gene patterns across

1428　mouse, macaque, and human. Species are color coded in the right column. Yellow boxes, con-

1429　served, green boxes, species-specific gene regulatory networks.

1430

1431　**Figure 3. Marker gene expression and validation in C-fiber pruriceptors, thermoreceptors,**

1432　**and nociceptors**

1433　(**A-D**) Marker genes for specific labeling of each cluster and validation by multiplex FISH for

1434　hNP1 and hNP2 (A), hPEP.SST(B), hTRPM8 (C), and hPEP.TRPV1/A1.1 and

1435　hPEP.TRPV1/A1.2 (D). The fluorescent images show detected transcripts in one example hDRG

1436　neuron (cell body outlined by the white dashed line. Circle charts next to images show quantifi-

1437　cation: the arcs indicate the percentage of neurons positive for the given marker gene over all

1438　sampled DRG neurons. The sector shaded areas indicate the approximate percentage of each cell

1439　type over the total quantified hDRG neurons. N=2, A (199 neurons total), B (220 neurons total),

1440　C (156 neurons total), D (202 neurons total). Scale bar, 50 μm.

1441

1442　**Figure 4. Marker gene expression and validation in A-fiber nociceptors**

1443　(**A-D**) Marker genes for specific labeling of each cluster and validation by multiplex FISH for

1444　hPEP.PIEZO2[h] (A), hPEP.KIT (B), hPEP.CHRNA7 (C), and hPEP.NTRK3 (D). The fluorescent

1445   images show the detected transcripts in one example hDRG neuron (cell body outlined by the

1446   white dashed line). Circle charts next to images show quantification: the arcs indicate the per-

1447   centage of neurons positive for the given marker gene over all sampled DRG neurons. The sector

1448   shaded areas indicate the approximate percentage of each cell type over the total quantified

1449   hDRG neurons. N=2, A (165 neurons total), B (173 neurons total), C (196 neurons total), D (191

1450   neurons total). Scale bar, 50 μm.

1451

1452   **Figure 5. Marker gene expression and validation in C- and A-LTMRs**

1453   (**A- D**) Marker genes for specific labeling of each cluster and validation by multiplex FISH for

1454   hC.LTMR (A), hAδ.LTMR (B), hAβ.LTMR (C), hPropr (D) and hATF3 (E). The fluorescent

1455   images show the detected transcripts in one example hDRG neuron (cell body outlined by the

1456   white dashed line). Circle charts next to images show quantification: the arcs indicate the per-

1457   centage of neurons positive for the given marker gene over all sampled DRG neurons. The sector

1458   shaded areas indicate the approximate percentage of each cell type over the total quantified

1459   hDRG neurons. N=2, A (205 neurons total), B (183 neurons total), C (188 neurons total), D (198

1460   neurons total), E (202 neurons total). Scale bar, 50 μm.

1461

1462   **Figure 6. Single-soma deep RNA-seq dataset is powerful for molecular discovery**

1463   (**A**) Comparison of the total detected number of GPCRs, ion channels, chemokine receptors, and

1464   neuropeptides in single-soma, spatial, and single-nucleus RNA-seq datasets. (**B-C**) Expression of

1465   putative itch receptors in hDRG neurons. Some receptors are highly enriched (B), while some

1466   receptors are barely detected (C) in human itch populations. (**D**) Novel GPCRs, ion channels,

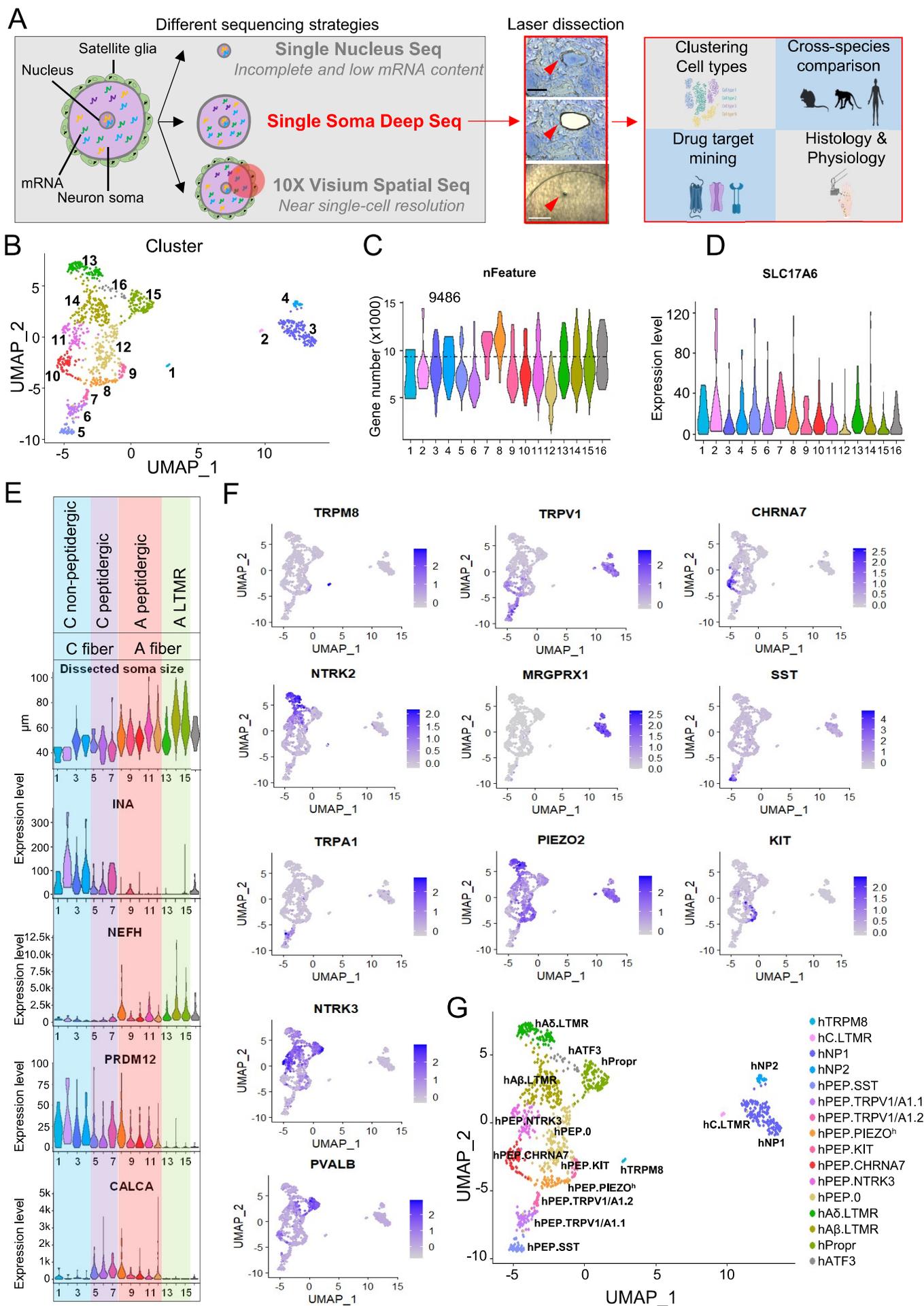1467   and other genes enriched in human itch populations.

1468

1469 **Figure 7. Molecular profile-informed single-unit microneurography recordings revealed**

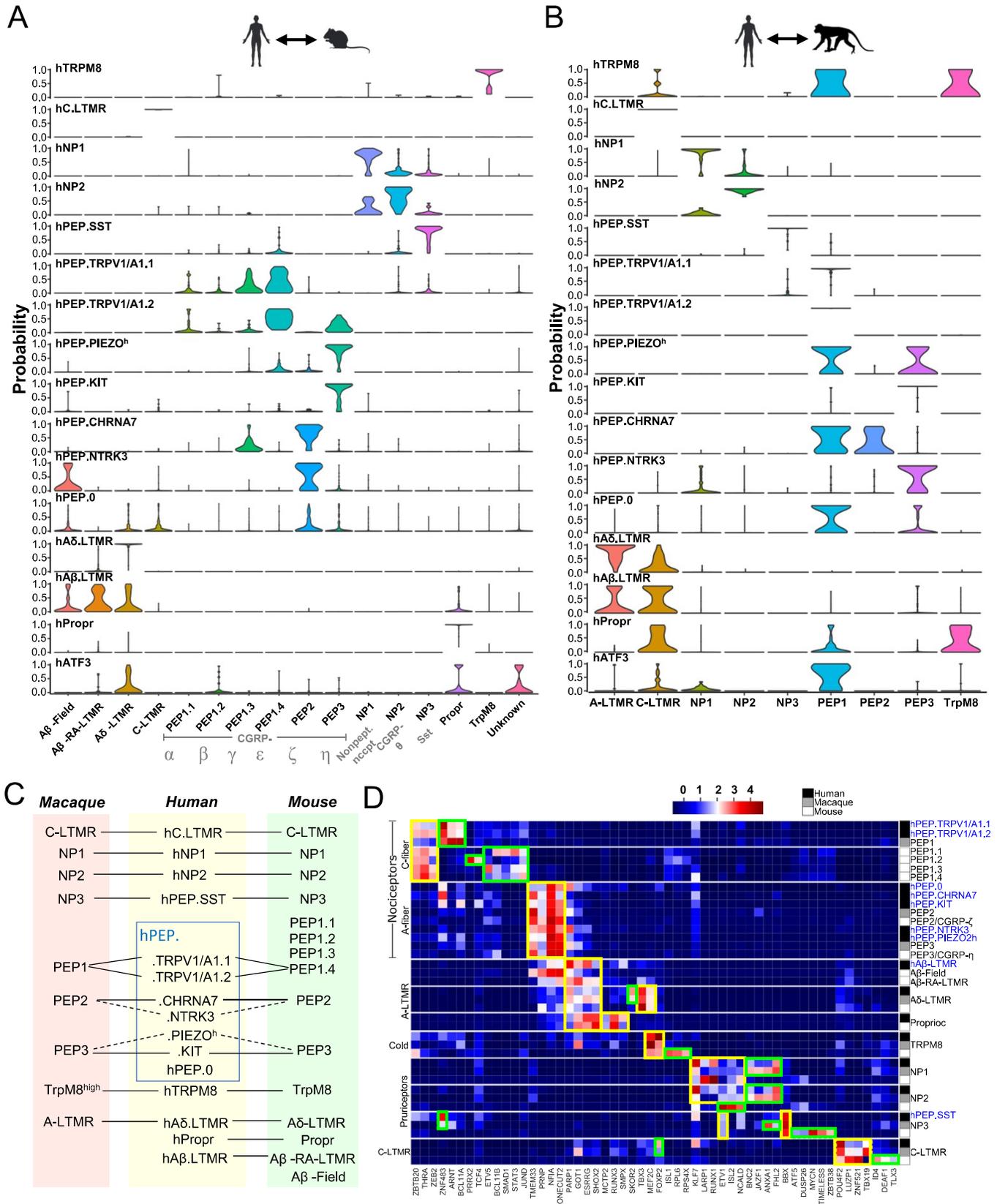1470 **novel physiological properties of a population of human A-HTMRs.**

1471 (**A**) Predicted physiological properties of hPEP.KIT sensory afferents based on single-soma deep

1472 RNA-seq data. (**B**) Receptive field locations of recorded single afferents from superficial per-

1473 oneal (S. peroneal), posterior antebrachial cutaneous (PABCN), and radial nerve recordings (n =

1474 47). (**C**) Distribution of mechanical (monofilament) thresholds for HTMRs and LTMRs in the

1475 recorded samples. (**D**) Individual and mean (±SEM) conduction velocities of different HTMR

1476 and LTMR types in response to surface electrical stimulation from upper and lower limbs (Field-

1477 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;

1478 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

1479 ± 0.05 m/s, n=8; C-HTMR: 0.7 ± 0.08 m/s, n=5). (**E**) Spike activities of a putative hPEP.KIT

1480 unit (A-HTMR cooling+) in response to repeated stimulations of the receptive field, superim-

1481 posed on an expanded time scale. (**F-H**) Responses of an A-HTMR cooling+ unit to soft and

1482 coarse brushing (F), heating (G) and cooling (H). (**I**) Schematic showing the link between the

1483 histologically identified KIT$^+$/CGRP$^+$/NEFH$^+$ sensory afferents and the molecularly defined

1484 hPEP.KIT population, likely representing a type of heat/capsaicin-insensitive but cold-sensitive

1485 A-HTMRs.

1486

1487 **Figure 8. Molecular profile-informed single-unit microneurography recordings revealed**

1488 **novel physiological properties of a population of human C-LTMRs.**

1489 (**A**) Novel physiological properties of hC.LTMR sensory fibers predicted based on gene expres-

1490 sion obtained from the single-soma deep sequencing. (**B**) Spike activity of a hC.LTMR unit in

1491   response to repeated stimulations of the receptive field, superimposed on an expanded time scale.

1492   (**C**) Individual and mean (±SEM) conduction velocities of different HTMR and LTMR types in

1493   response to surface electrical stimulation from upper and lower limbs (the same plot from Fig.

1494   7D) (**D-H**) Responses of a hC.LTMR unit to soft brushing and hair movement (D), cooling (E),

1495   menthol (F), heating (G) and capsaicin (H). Spike activity of that hC.LTMR before and after

1496   capsaicin application, overlaid on an expanded timescale (H). Conduction delay was adjusted

1497   based on the latency of electrically triggered spiking for that recorded afferent. Note, different

1498   scaling.

Figure 1

Figure 2

Figure 3
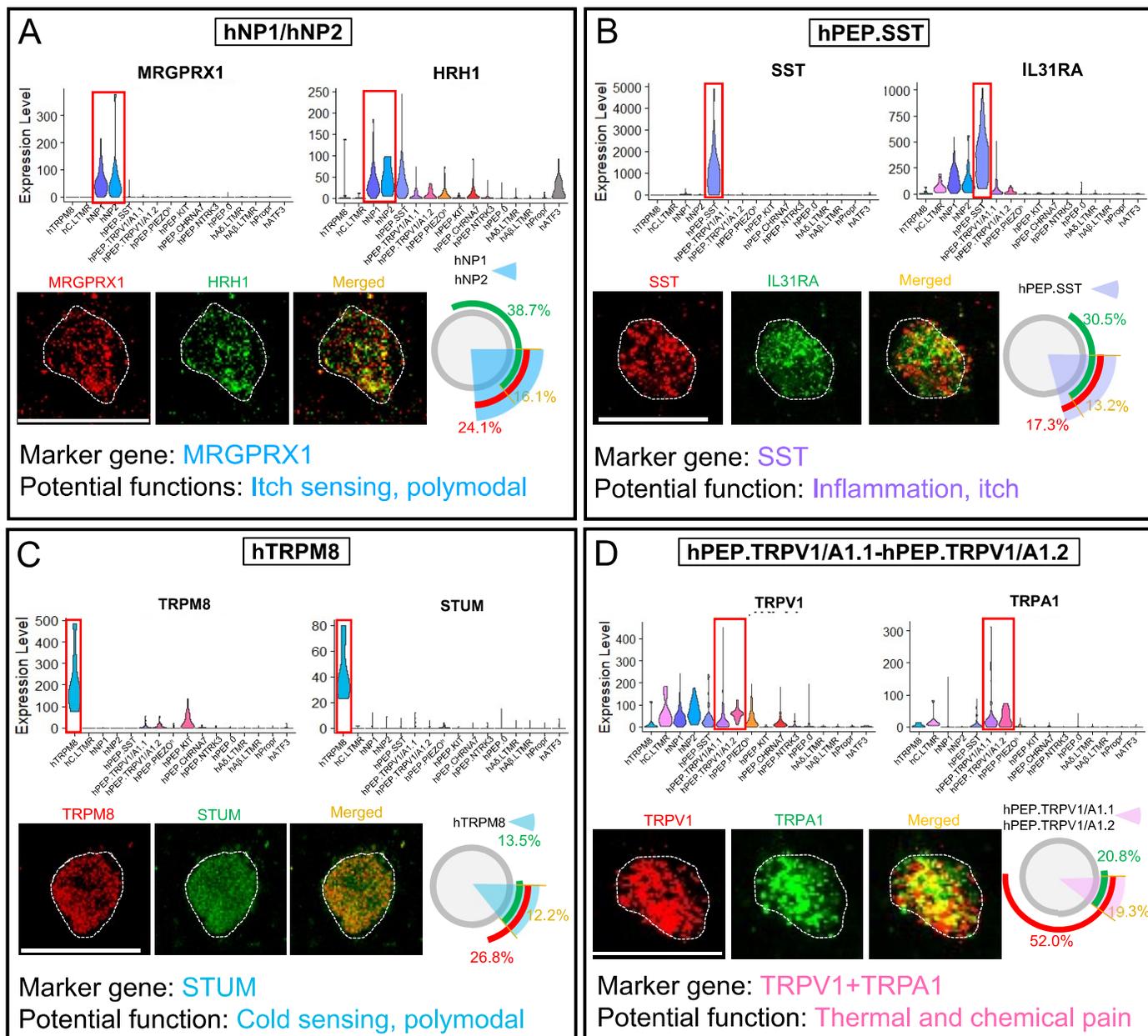
Figure 4

Figure 5

Figure. 6

A

Single soma | Spatial | Single nucleus

GPCRs, Ion channels, Chemokine receptors, Neuropeptides
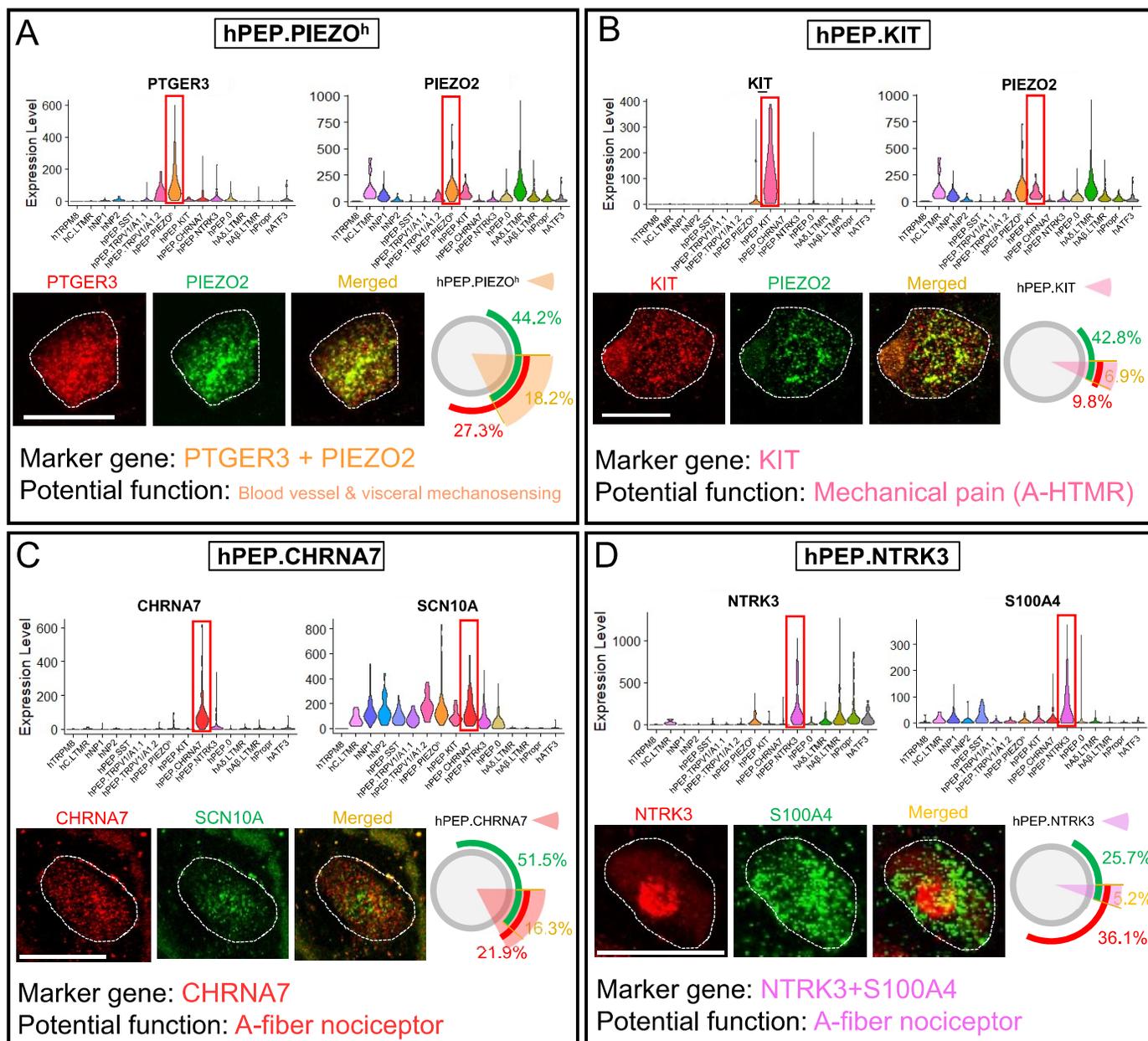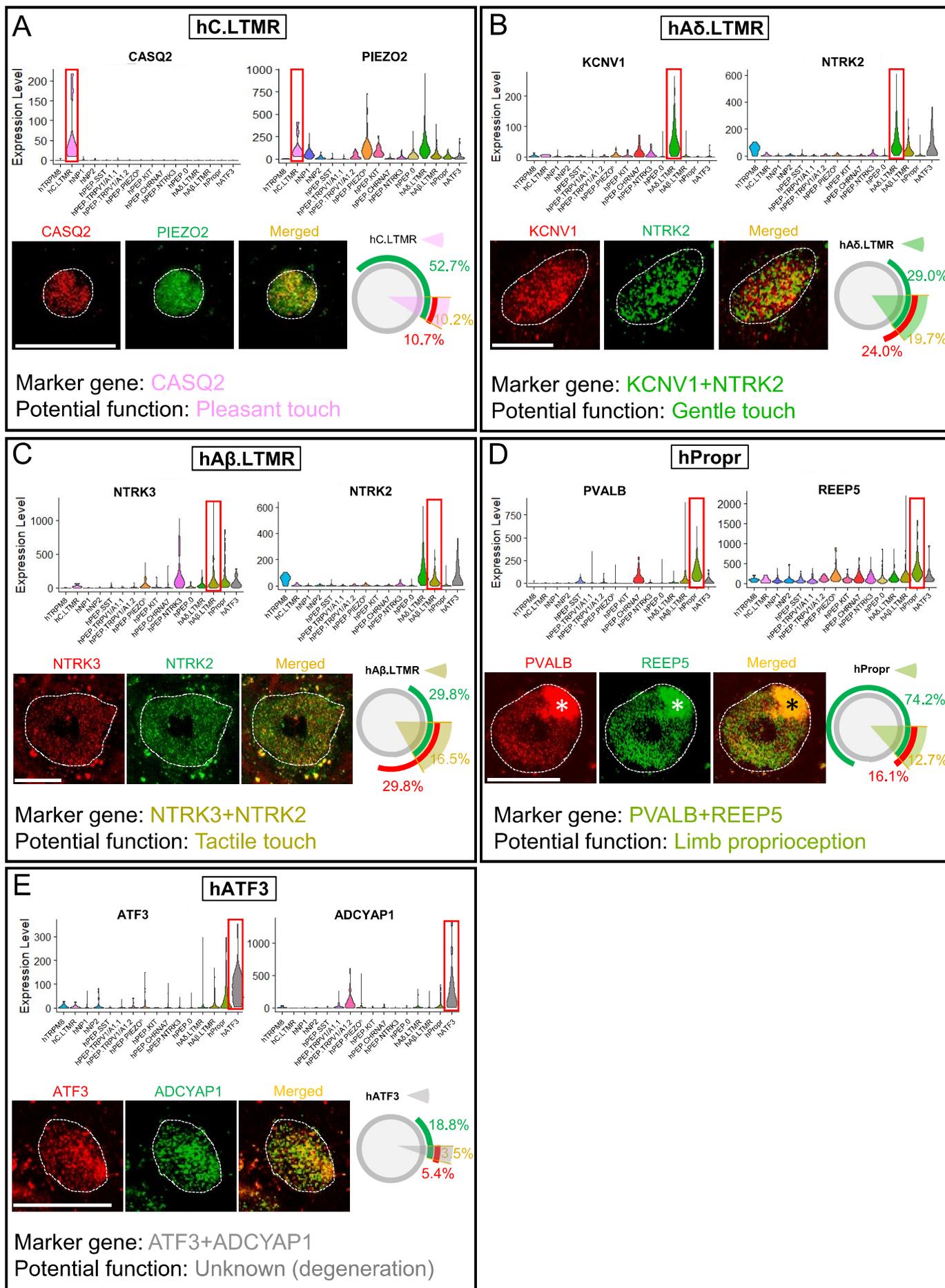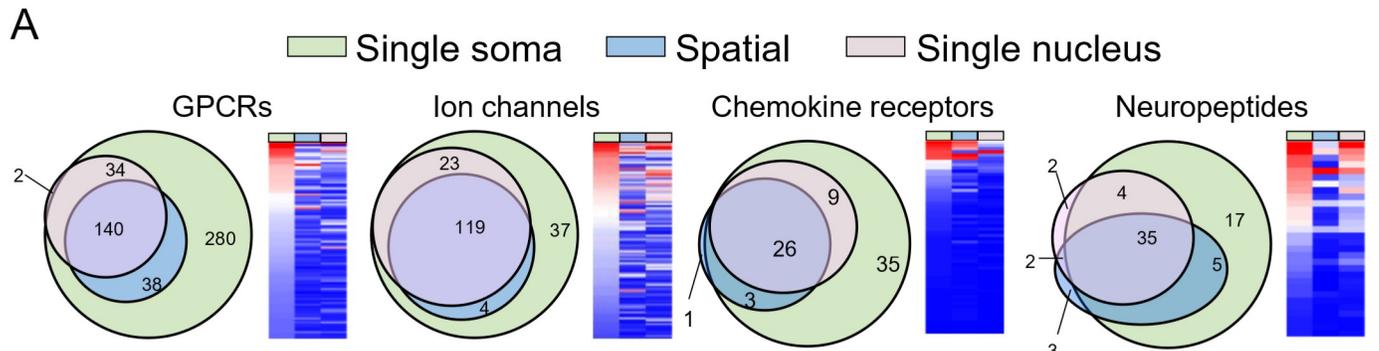
B
Receptors enriched in human itch-sensing DRG neurons
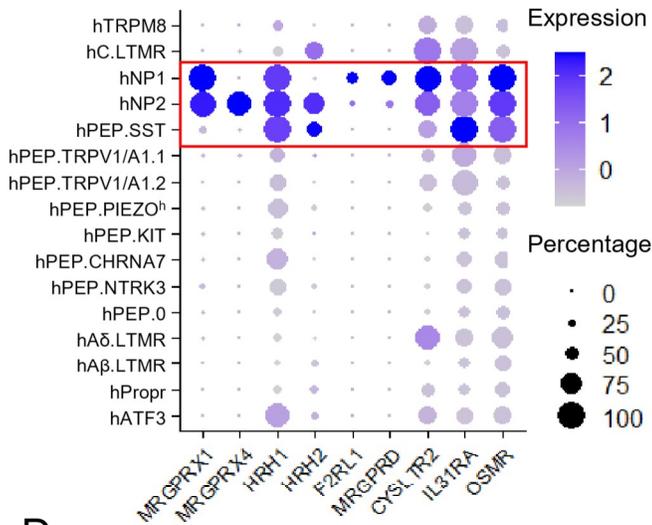
C
Anti-itch drug targets not enriched in human itch-sensing DRG neurons
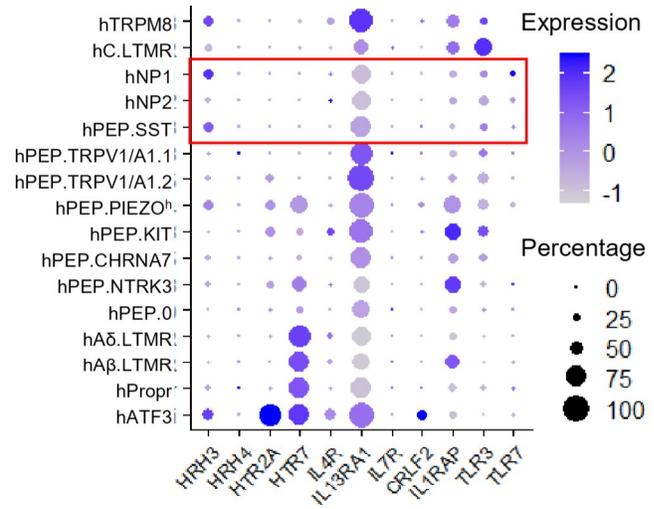
D
Potential novel anti-itch drug targets

Figure 7

A **Physiological properties predicted by deep sequencing**

hPEP.KIT (A-HTMR)

Conduction velocity — A fiber (2-60 m/s)
Mechanical sensitivity ✓
Heat sensitivity ✗
Cold sensitivity ✓

Human *In vivo* single-unit sensory fiber recording

C
- Field-LTMR (n=3)
- SA1-LTMR (n=6)
- SA2-LTMR (n=6)
- A-HTMR cooling- (n=5)
- A-HTMR cooling+ (n=5)
- C-LTMR (n=11)
- C-HTMR (n=11)

D Conduction velocity

F Mechanical Sensitivity — Soft brushing / Coarse brushing

G Heat Sensitivity — Heating 30°C → 50°C

H Cold Sensitivity — Cooling 30°C → 0°C

I

Histological type (Human skin immunostaining)	Molecular type (Single-Soma Deep sequencing)	Physiological type (*In vivo* human recording)
KIT+ CGRP+ NEFH+ PGP9.5+ fibers	hPEP.KIT	A HTMR Cooling+ Heat- Unit

Figure 8

A Physiological properties predicted by deep sequencing

Molecular profiles

CASQ2 PIEZO2 TRPV1 TRPM8

Predicted properties

hC.LTMR ➡

Mechanical sensitivity ✓ Heat sensitivity ✓ Chemical & cold sensitivity — menthol ✗ ❄ ?

B

C C fiber

D Low threshold mechanoreceptor
Soft brushing   Hair movement

Cold & Menthol Sensitivity   **Heat & Capsaicin Sensitivity**

E Cooling   **G** Heating

F Before menthol / After menthol   **H** Before capsaicin / After capsaicin

Supplementary Figure 1

A. Cryo-sections → 70% EtOH (30s) → HistoGene™ staining (10s) → 70% EtOH (30s) → 95% EtOH (30s) → 100% EtOH (30s) → LCM

B.

C.

D. All DRG neurons

E. Dissected DRG neurons

Supplementary Figure 2

Supplementary Figure 3

Supplementary Figure 4
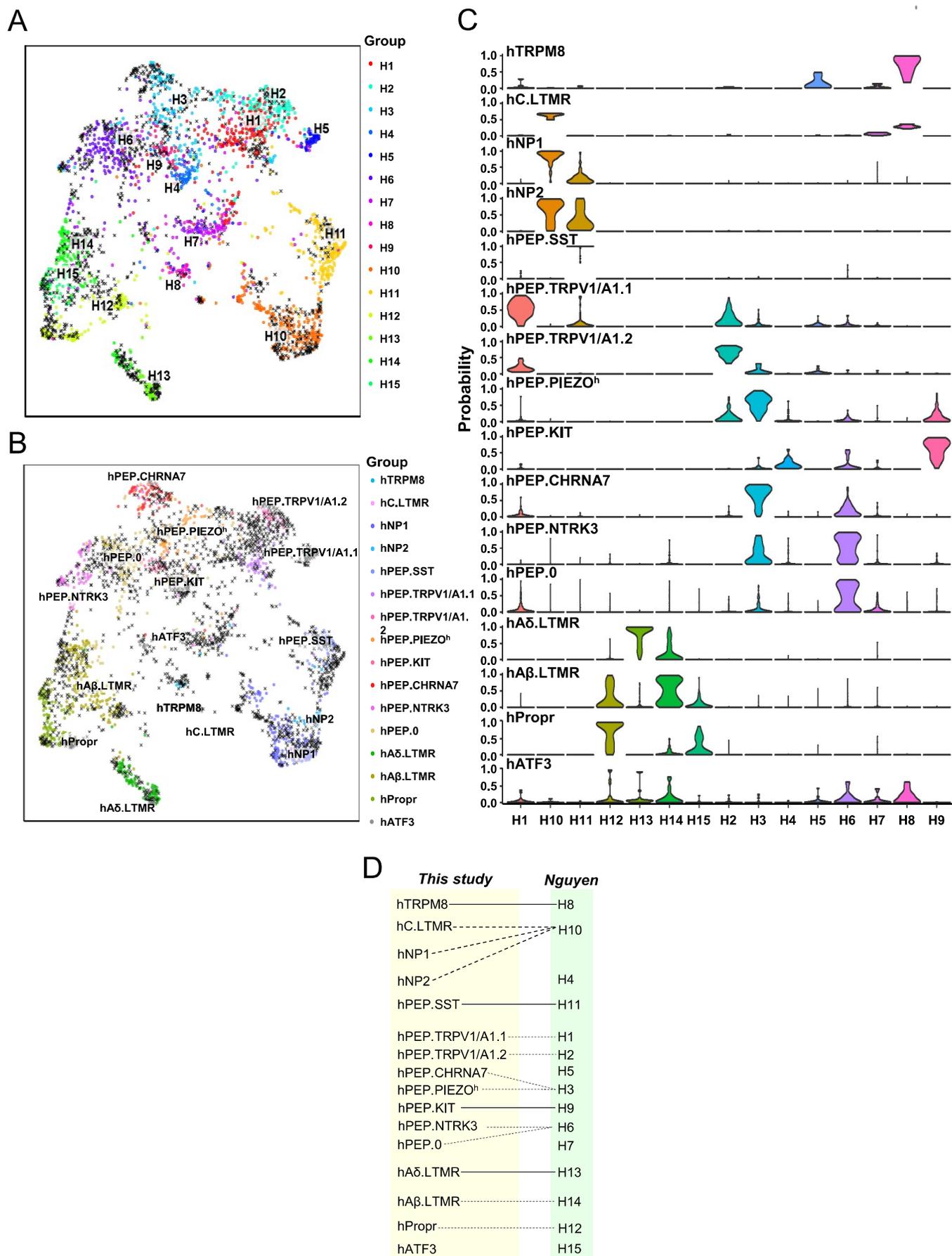
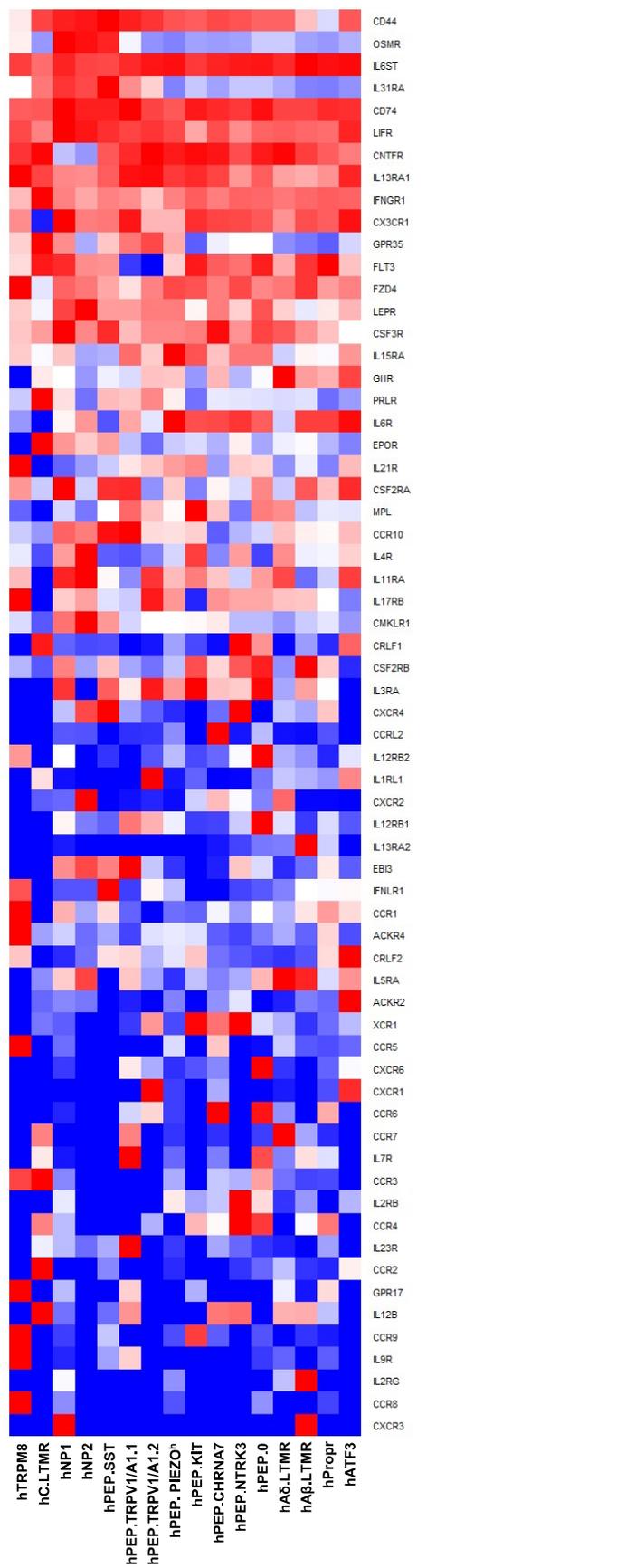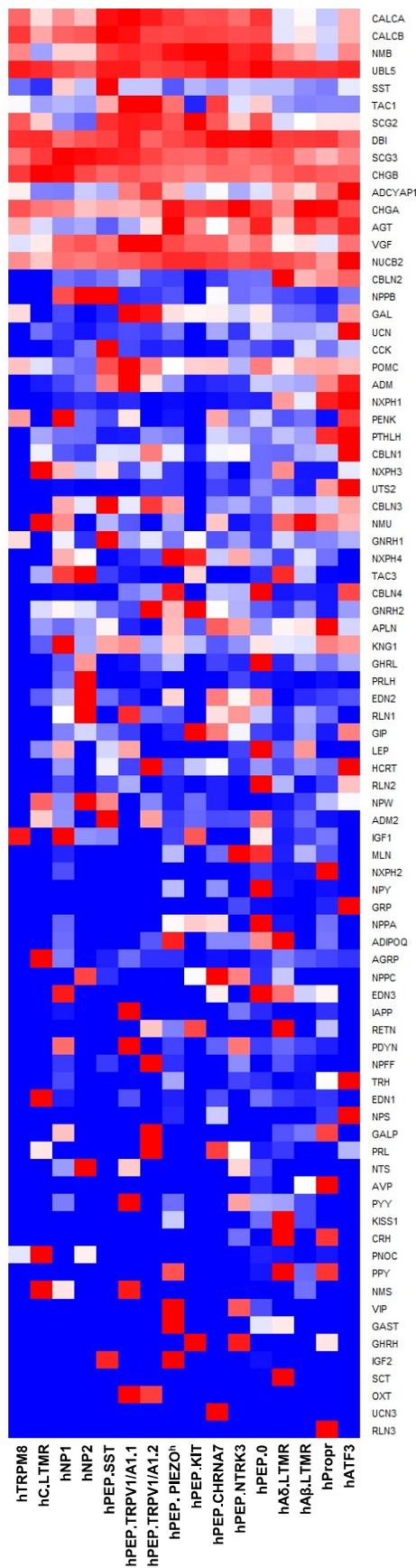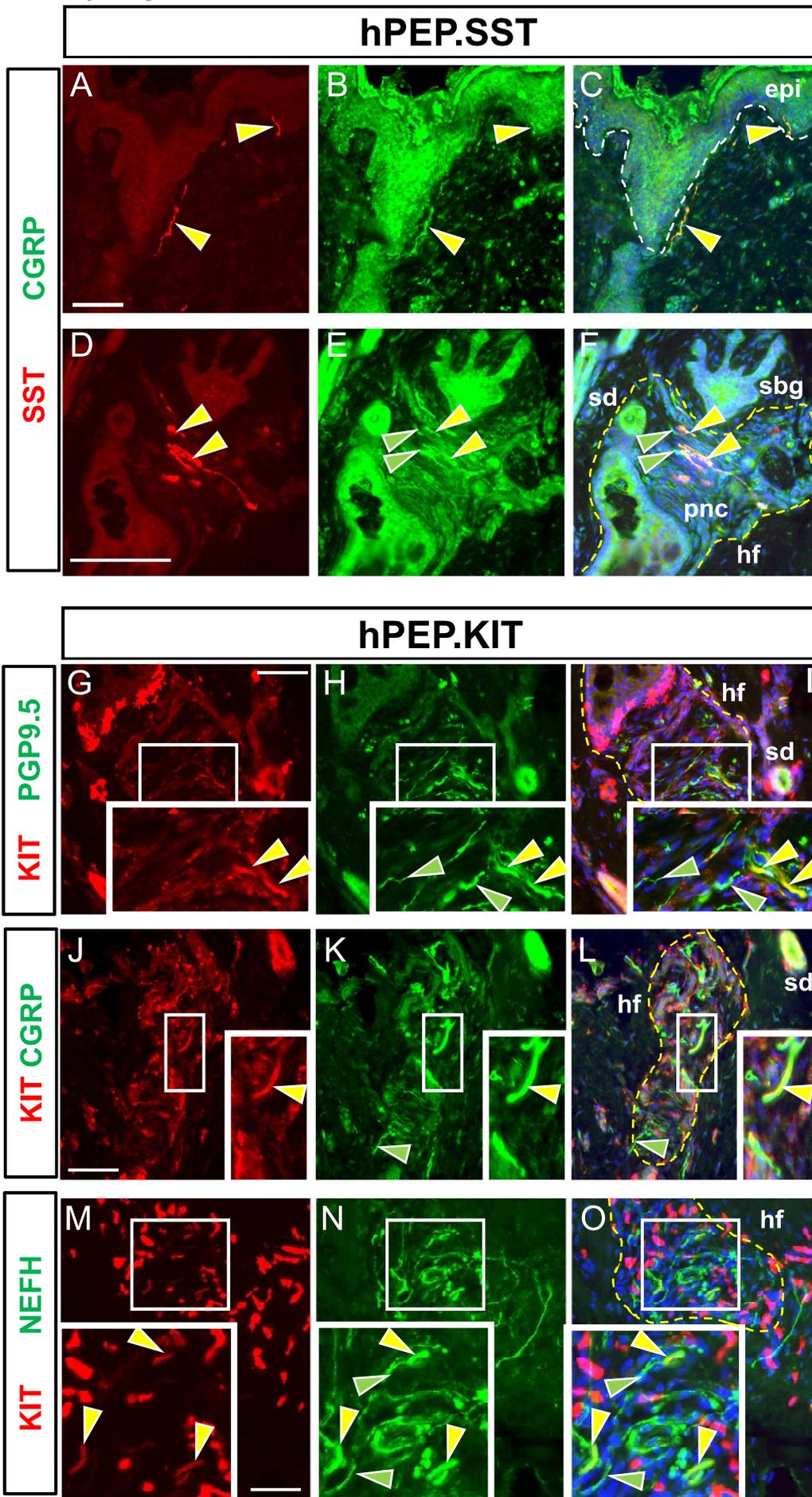Supplementary Figure 5

Supplementary Figure 6
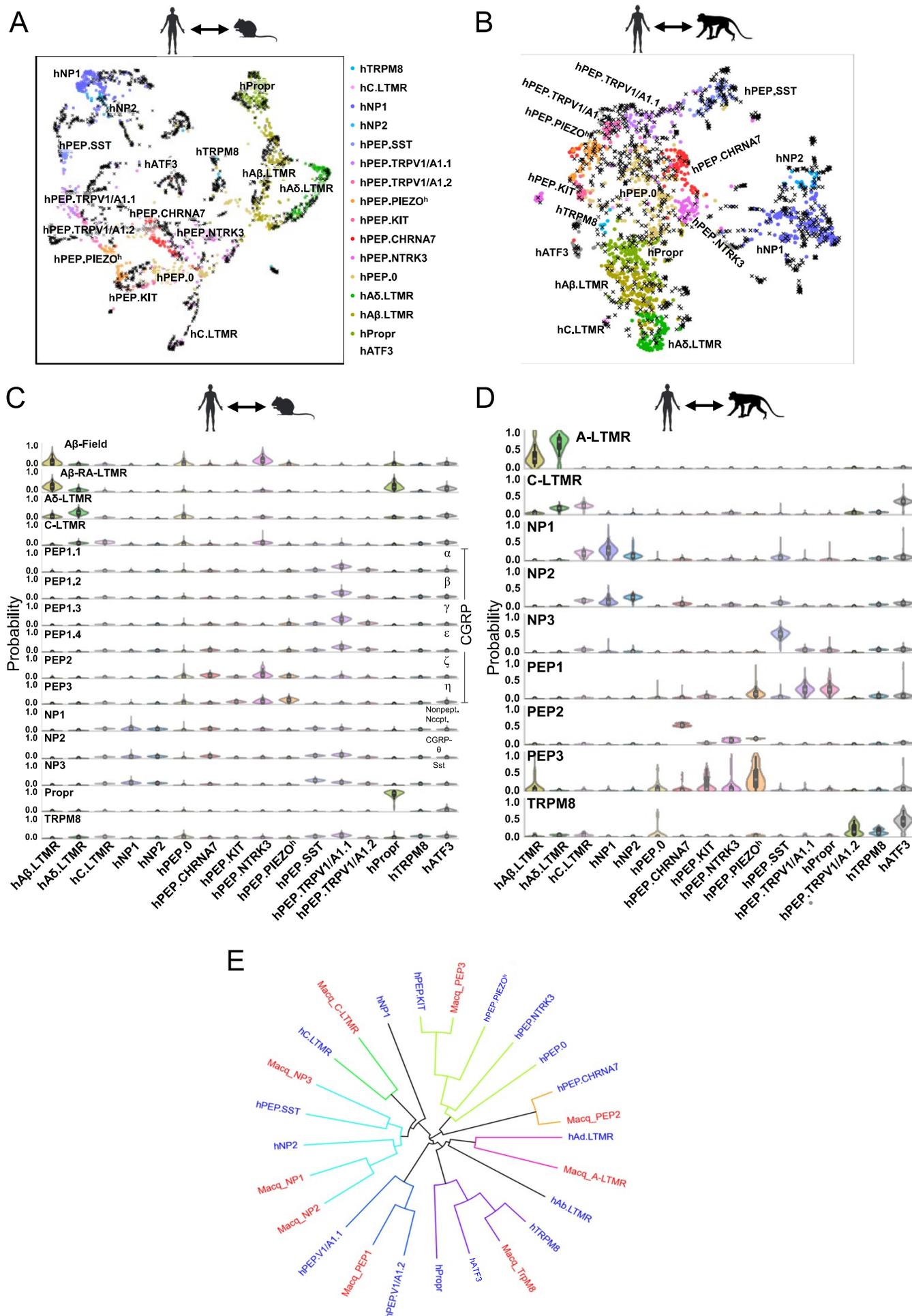
Supplementary Figure 7

Supplementary Figure 8

Supplementary Figure 9

A. Human

B. Mouse (Sharma)

C. Macaque (Kupari, SmartSeq2)
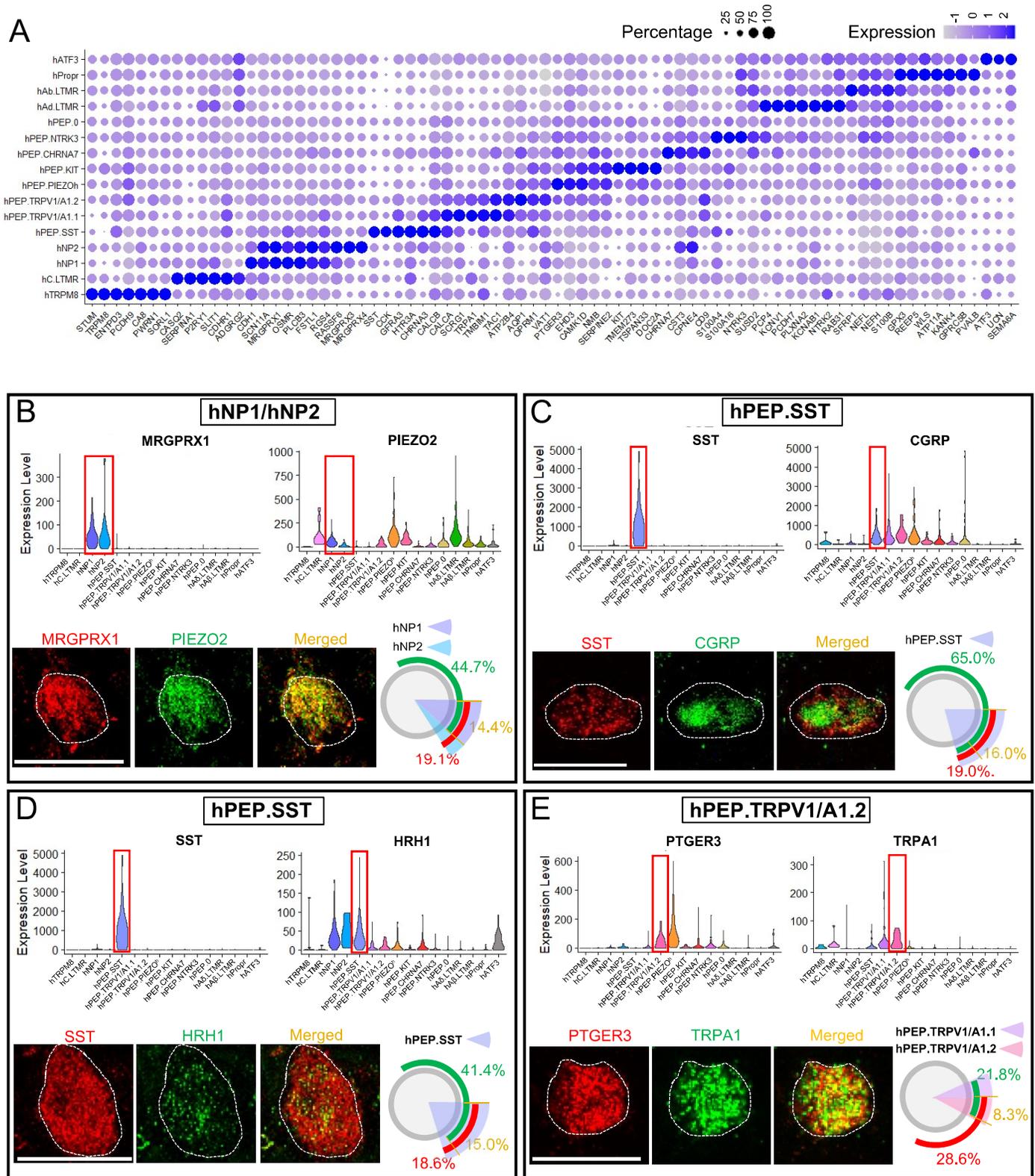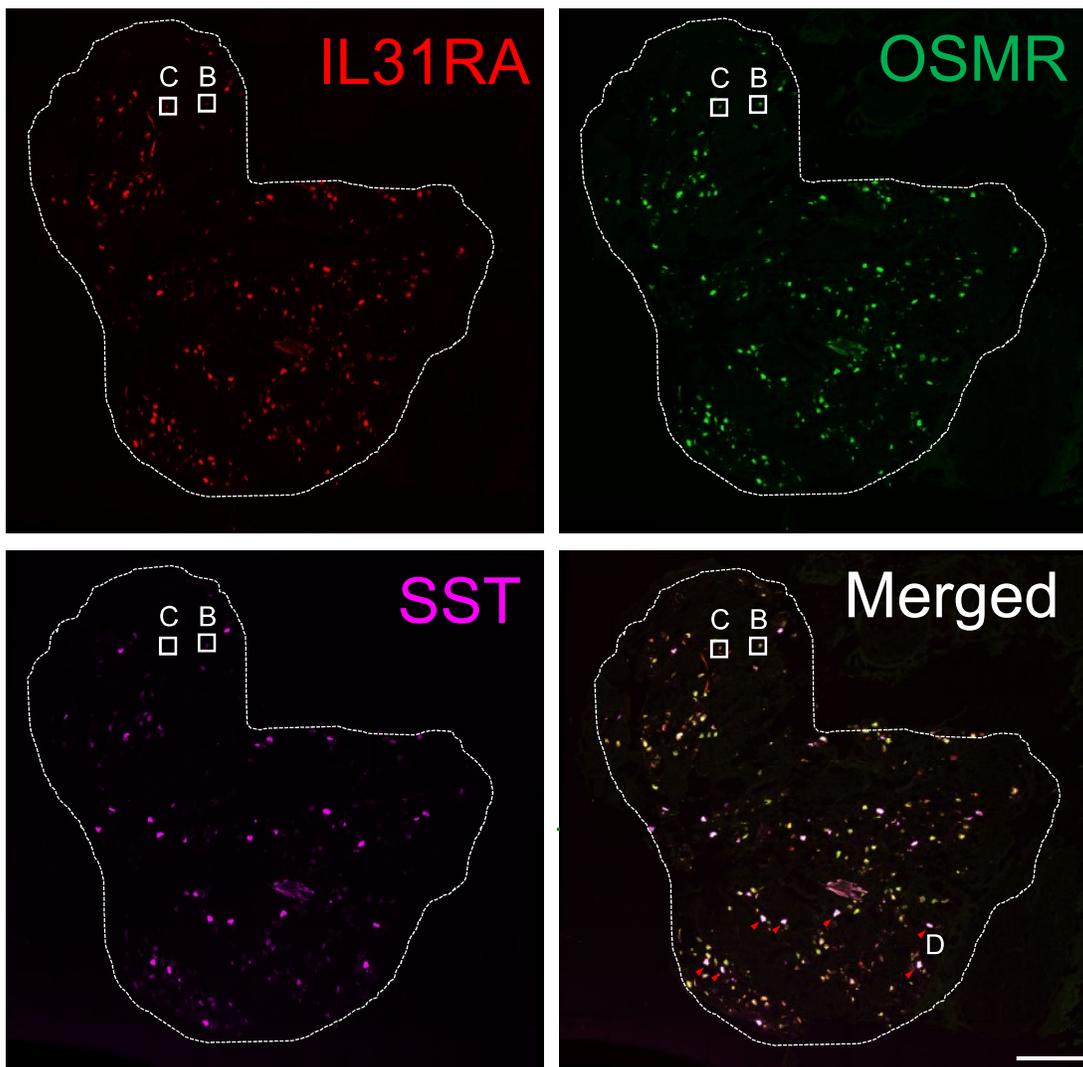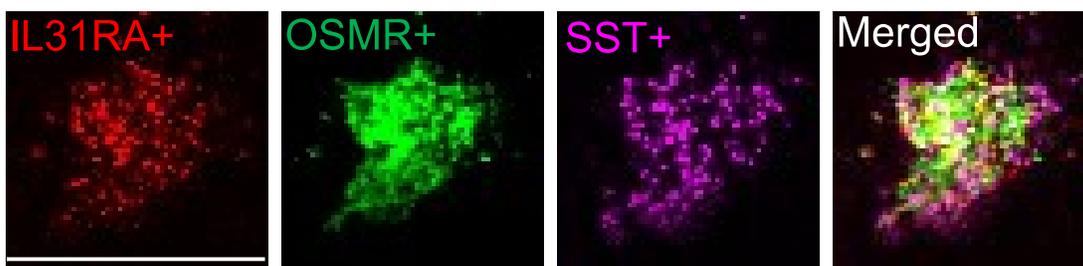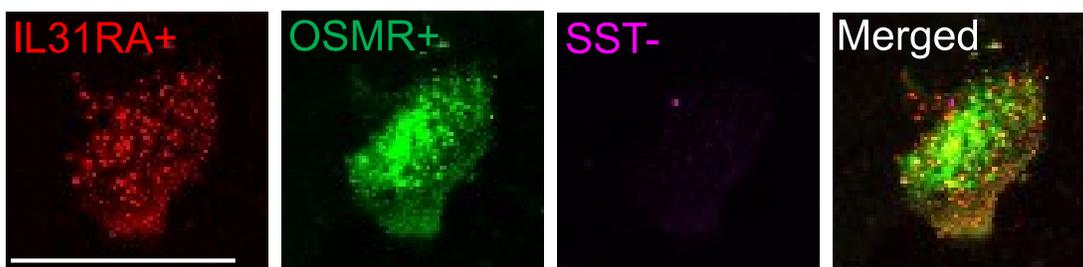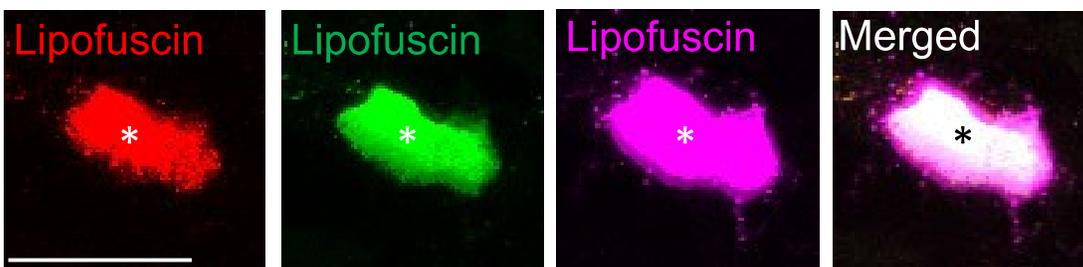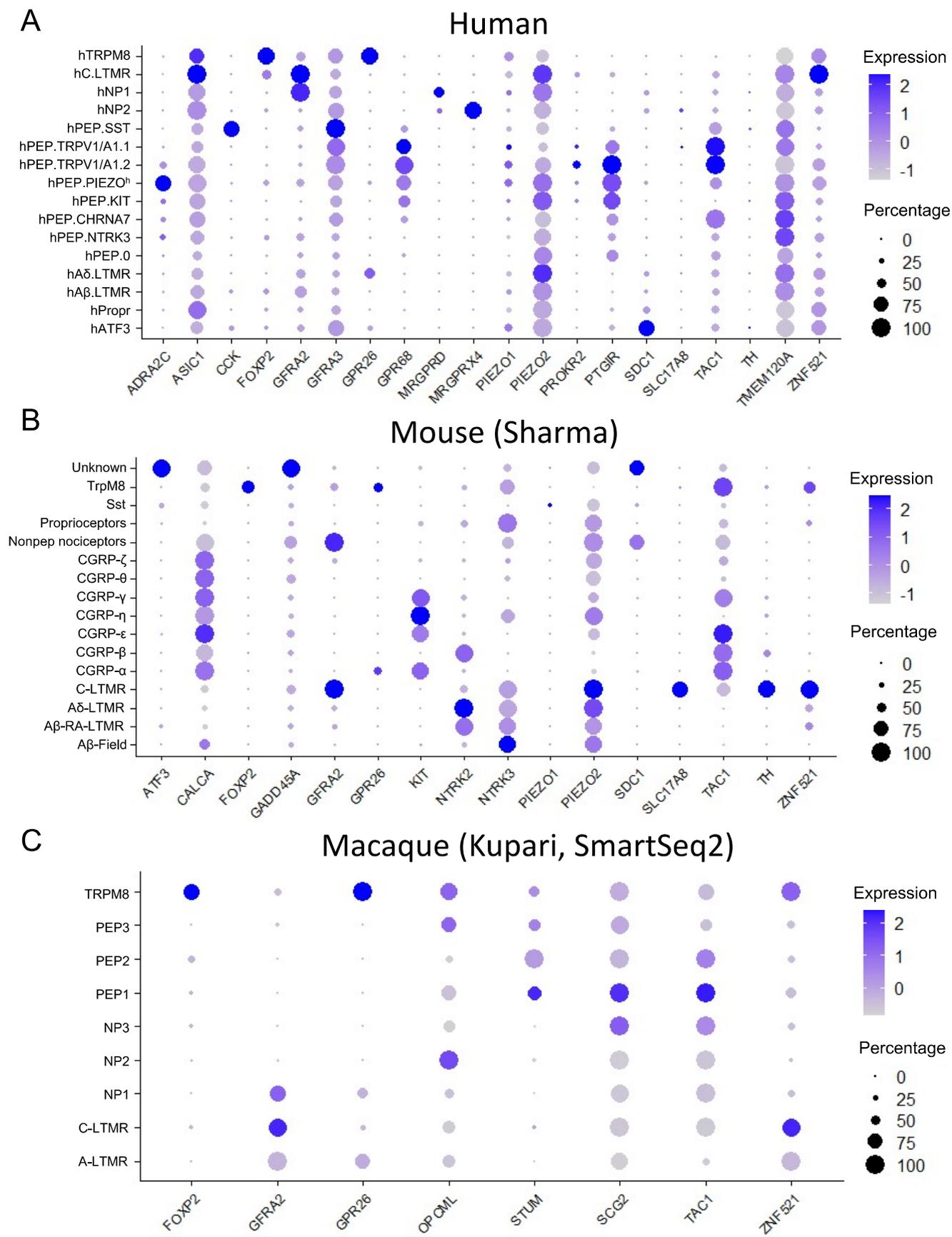
Supplementary Figure 10
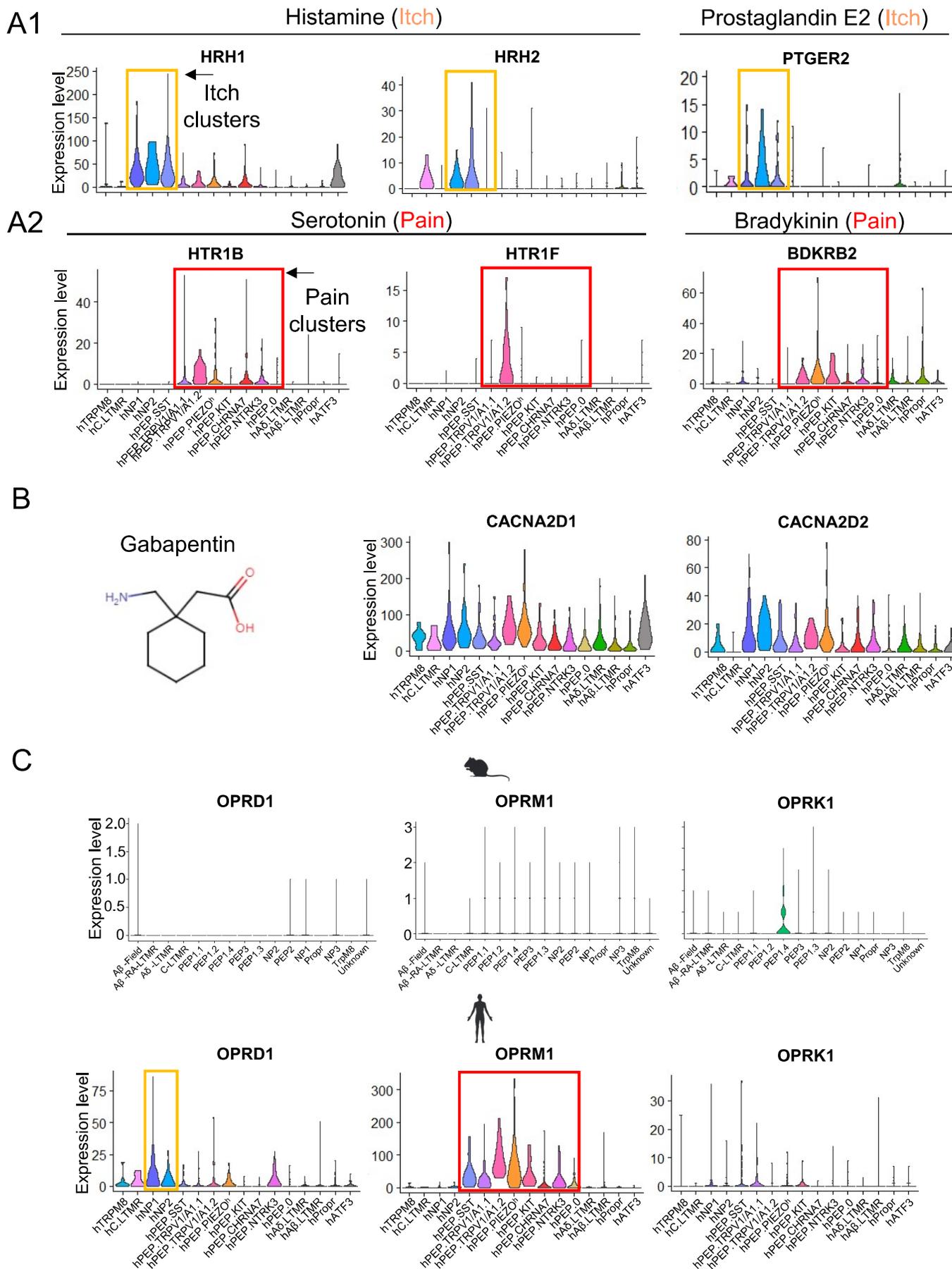
Supplementary Figure 11
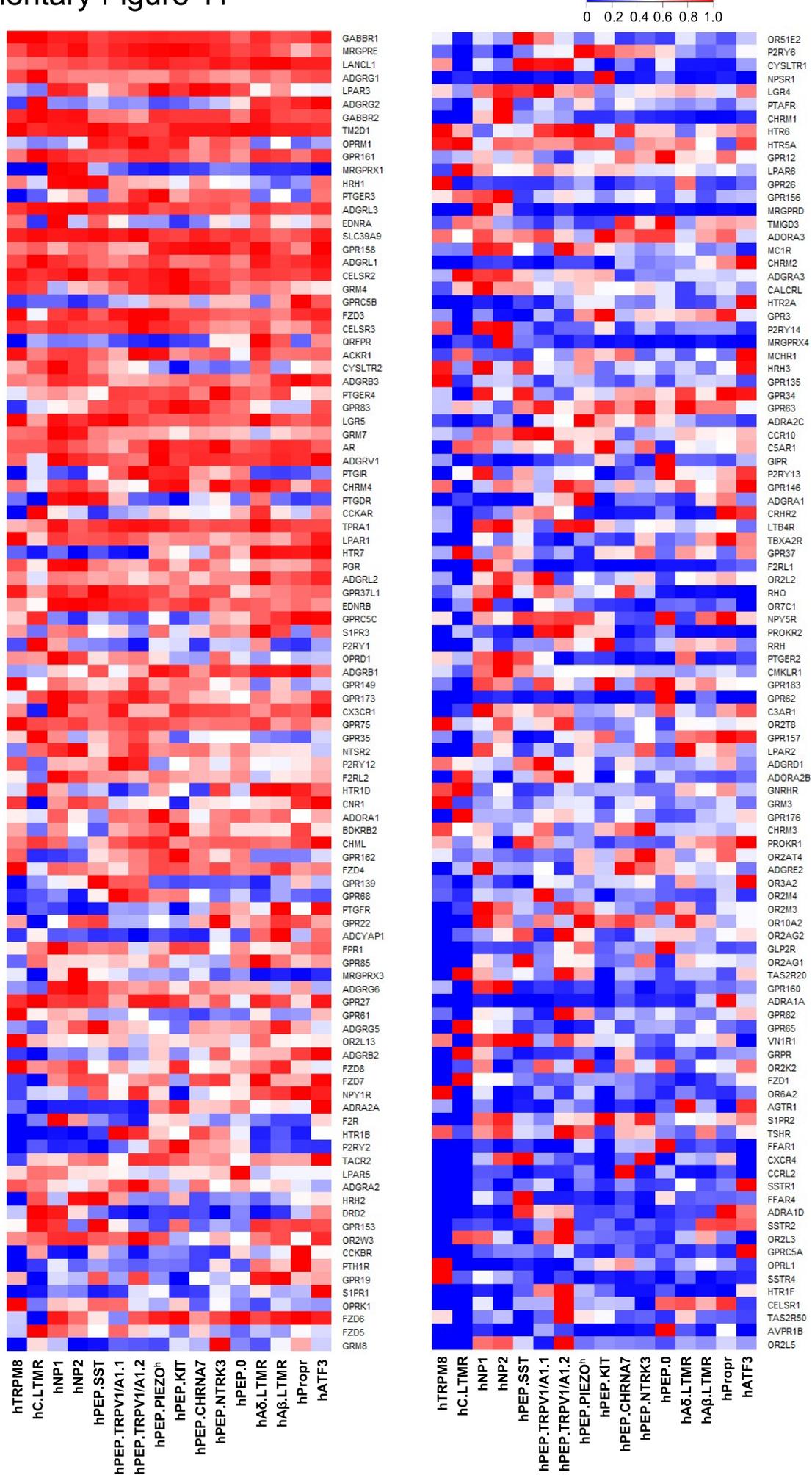
Supplementary Figure 12
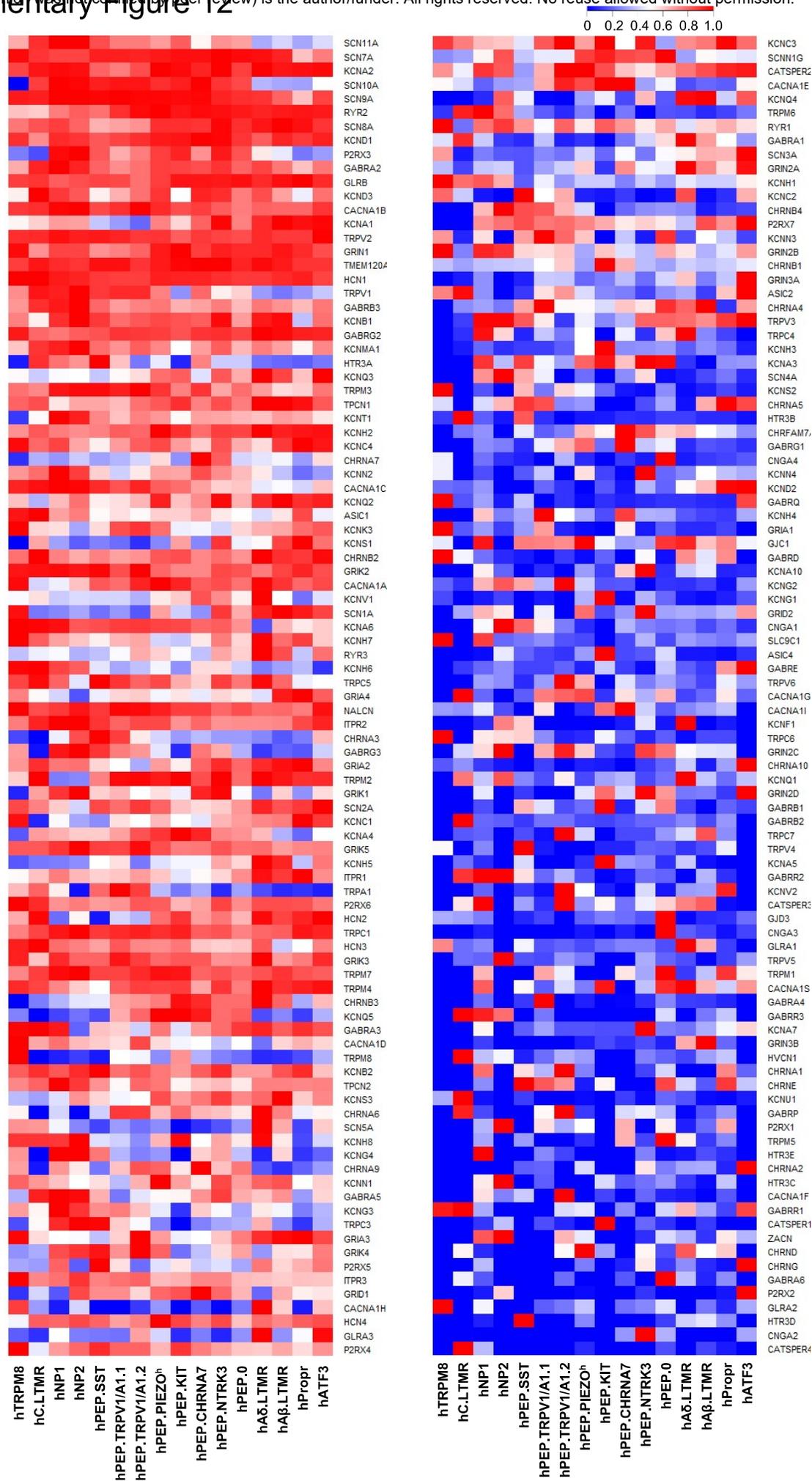
Supplementary Figure 13

Supplementary Figure 14

Supplementary Figure 15

Supplementary Figure 16

A

hPEP.KIT

B

Cooling

C

Heating

Supplementary Figure 17