

Somatosensory neuron types identified by high-coverage single-cell RNA-sequencing and functional heterogeneity

Chang-Lin Li^{1*}, Kai-Cheng Li^{1*}, Dan Wu¹, Yan Chen¹, Hao Luo⁴, Jing-Rong Zhao¹, Sa-Shuang Wang⁴, Ming-Ming Sun², Ying-Jin Lu¹, Yan-Qing Zhong¹, Xu-Ye Hu⁵, Rui Hou², Bei-Bei Zhou², Lan Bao^{3,4}, Hua-Sheng Xiao², Xu Zhang^{1,4}

¹Institute of Neuroscience and State Key Laboratory of Neuroscience, CAS Center for Excellence in Brain Science, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 20031, China; ²National Engineering Center for Biochip at Shanghai, Shanghai, China; ³State Key Laboratory of Cell Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences; ⁴School of Life Science and Technology, ShanghaiTec University, Shanghai 200031, China; ⁵Shanghai Clinical Center, Chinese Academy of Sciences/XuHui Central Hospital, Shanghai, China

Sensory neurons are distinguished by distinct signaling networks and receptive characteristics. Thus, sensory neuron types can be defined by linking transcriptome-based neuron typing with the sensory phenotypes. Here we classify somatosensory neurons of the mouse dorsal root ganglion (DRG) by high-coverage single-cell RNA-sequencing (10 950 ± 1 218 genes per neuron) and neuron size-based hierarchical clustering. Moreover, single DRG neurons responding to cutaneous stimuli are recorded using an *in vivo* whole-cell patch clamp technique and classified by neuron-type genetic markers. Small diameter DRG neurons are classified into one type of low-threshold mechanoreceptor and five types of mechanoheat nociceptors (MHNs). Each of the MHN types is further categorized into two subtypes. Large DRG neurons are categorized into four types, including neurexophilin 1-expressing MHNs and mechanical nociceptors (MNs) expressing BAI1-associated protein 2-like 1 (Baiap2l1). Mechanoreceptors expressing trafficking protein particle complex 3-like and Baiap2l1-marked MNs are subdivided into two subtypes each. These results provide a new system for cataloging somatosensory neurons and their transcriptome databases.

Keywords: primary sensory neuron; neuron type; single-cell RNA sequencing; somatosensation
Cell Research (2016) 26:83-102. doi:10.1038/cr.2015.149; published online 22 December 2015

Introduction

Neurons express a higher number of genes than any other cell type, resulting in the high heterogeneity of neuron subtypes. The pseudo-unipolar sensory neurons of the dorsal root ganglion (DRG) transmit peripheral signals to the spinal cord. Small diameter DRG neurons give rise to unmyelinated axons (C-fibers) and thinly myelinated axons (A δ -fibers) that convey the nociceptive, thermal and mechanoreceptive signals generated at peripheral nerve terminals to neurons in lamina I-II of the

spinal cord. Large DRG neurons transmit mechanoreceptive and proprioceptive signals via thickly myelinated afferents (A β -fibers) to spinal lamina III-V [1]. Traditionally, small neurons have been classified into three subsets: isolectin B4 (IB4)-positive (non-peptidergic, NP) subset, peptidergic (PEP, IB4 negative) subset expressing neuropeptide substance P and calcitonin gene-related peptide (CGRP), and tyrosine hydroxylase (TH)-expressing subset, while large neurons express neurofilament 200 (NF200) [2-5].

Cutaneous mechanoreceptors (MRs), nociceptors, thremoreceptors and itch receptors have been physiologically characterized [5, 6]. Many nociceptors respond to multiple stimulus modalities, whereas others have more specialized response properties. Defined nociceptor subtypes have been termed C-fiber (C-) or A δ -fiber (A-) mechanoheat nociceptors (MHNs), mechanical nociceptors (MNs), mechanically insensitive (MI) or sensitive

*These two authors contributed equally to this work.

Correspondence: Xu Zhang

Tel: 86-21-54921726; Fax: 86-21-54921762

E-mail: xu.zhang@ion.ac.cn

Received 13 August 2015; revised 4 November 2015; accepted 1 December 2015; published online 22 December 2015

afferents, and mechanoheat-cold nociceptors [5, 6]. The molecular properties of DRG neurons have been extensively studied, including their expression of various receptors and ion channels such as neuropeptide Y (NPY) receptors [7], MAS-related G-protein-coupled receptors (MRGPRs) [8, 9], voltage-gated Na⁺ channels [10], transient receptor potential (TRP) channels [11, 12], ATP receptors [13], acid-sensing ion channels (ASICs) [14, 15] and tyrosine kinase receptors (TRKs) [16]. Gene expression profiles of DRG tissue have been also analyzed by microarray and RNA-sequencing (RNA-seq) techniques [17, 18]. Although these methods have helped to identify the expression of neuromodulators such as natriuretic peptide B (NPB) [19] and regulators of Na⁺, K⁺-ATPase (NKA), including the γ -subunit of NKA (FXVD domain-containing ion transport regulator 2, *Fxyd2*) [20-22], they cannot provide a global view of the transcriptional profiles of individual neurons.

Recently, multiple efforts have been made to analyze the transcriptional profiles of DRG neurons. The transcriptional profiles of TRPV1 lineage and Na_v1.8 channel-expressing DRG neurons have been examined by RNA-seq [23, 24]. By single-cell PCR, Chiu *et al.* identified six subgroups of DRG neurons [25]. Single-cell RNA-seq enables a better understanding of a cell's transcriptome [26-33]. Usoskin *et al.* performed low-coverage single-cell RNA-seq (3 574 ± 2 010 genes per neuron) and classified the mouse DRG neurons into two PEP types, three NP types, TH type and five NF200-positive types within the traditional classification framework [34]. However, this method resulted in transcriptional variation among DRG neurons due to the limited number of genes detected in each neuron. These studies are useful resources for analyzing the molecular determinants of sensory neuron types, but their interpretation is complicated by the large variation present within a neuron type defined by a given molecular marker. Moreover, the related functional phenotypes of the neurons have not been experimentally analyzed.

Here, by integrating single-cell techniques such as high-coverage RNA-seq (10 950 ± 1 218 genes per neuron), *in vivo* patch clamp recording and single-cell PCR, we classify mouse DRG neurons into 10 types and 14 subordinate subtypes with distinct transcriptional patterns, molecular markers and functional annotations, revealing a new catalog of somatosensory receptors. Our single-cell RNA-seq-based clustering reveal more neuron types and subtypes than previous classifications of DRG neurons, and demonstrate that traditional neuron subset-specific markers in fact label multiple neuron types. Moreover, our study suggests that based on the current understanding of molecular function and signal-

ing networks, transcriptome data can partially predict the functions of neuron types. Further, neuron type-specific functional analyses are needed to confirm and elaborate the precise functions of these neuron types. Thus, neuron types can be defined by integrating their transcriptomic, morphological and functional characteristics.

Results

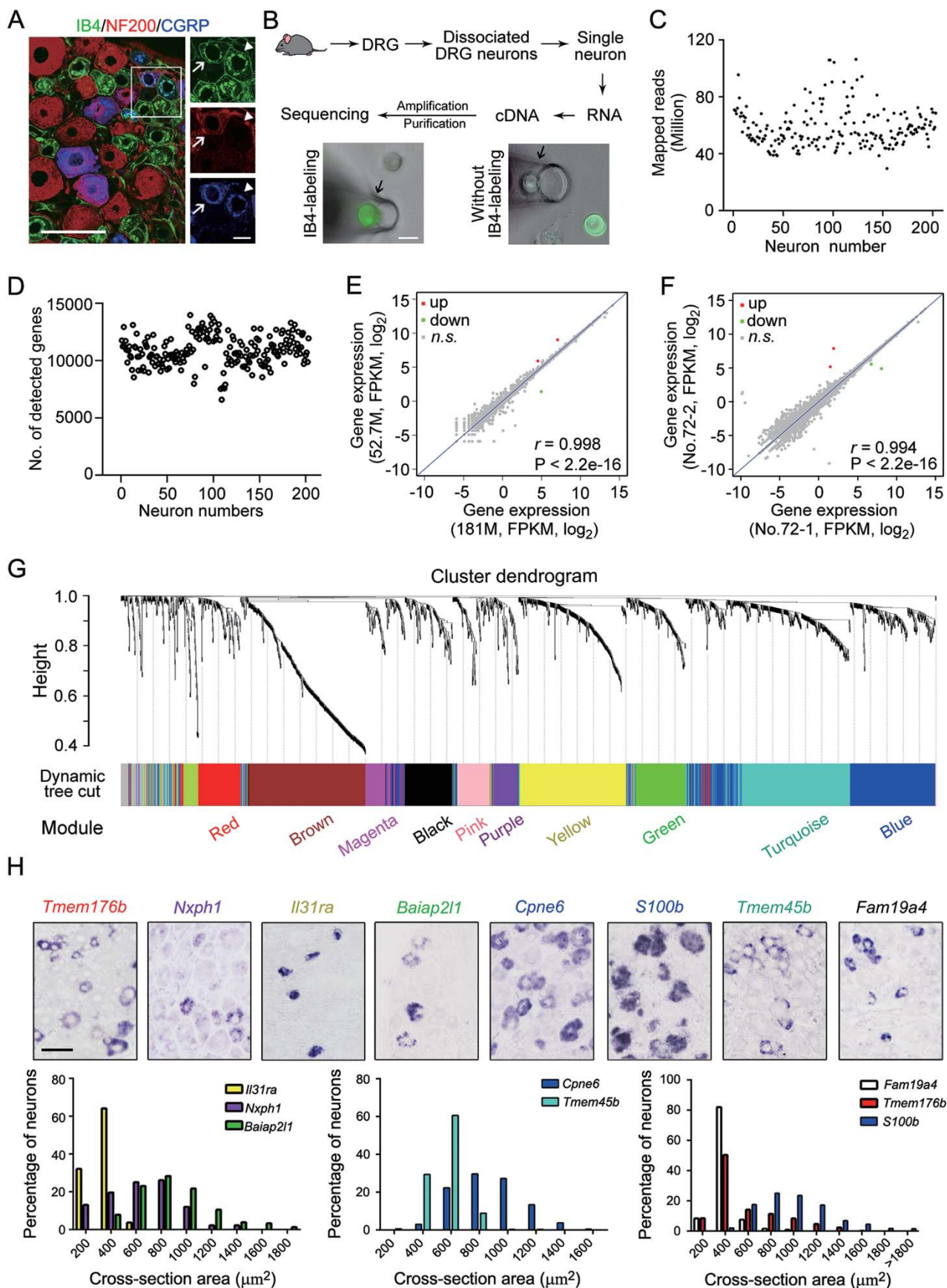
Neuron sampling and quality control of transcriptomic data

RNA-seq analysis showed that the transcriptomes of lumbar 5 (L5) DRG from five adult male mice were homogenous (Supplementary information, Figure S1A and S1B). Immunostaining of CGRP and NF200, combined with IB4 fluorescent labeling showed that ~35% of neurons were positive for IB4, ~41.2% were positive for CGRP and ~46.4% were positive for NF200 (Figure 1A). Approximately 20% of IB4-positive neurons expressed CGRP, whereas ~5.6% expressed NF200. Approximately 28% of CGRP-positive neurons expressed NF200. A small number of neurons were triple labeled.

A DRG-specific approach was designed to increase the efficiency of single-cell RNA-seq used for neuron typing. First, IB4-positive neurons were identified among neurons freshly dissociated from lumbar DRG by IB4 fluorescence labeling (Figure 1B). Under a microscope, neurons without identifiable satellite cells attached to the cell surface were selected and aspirated with glass pipettes. IB4-positive small neurons, IB4-negative small neurons (cross-sectional area < 800 μm^2) and large neurons contributed ~1/3 of the selected neurons. The total number of neurons used was determined by the results of transcriptomic analysis. We profiled neurons until the number of neuron clusters extracted from the data plateaued and the correlation among neurons stabilized. Sequencing libraries were constructed from the cDNA of 203 neurons collected from 19 mice.

To ensure quality of the samples, FPKM (fragments per kilobase of transcript per 10⁶ mapped reads) values for house-keeping genes actin B (*Actb*) or glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*) in each neuron were required to be around 4 675 ± 3 456 (mean ± SD) or 367 ± 268, respectively. Datasets for the corresponding 197 neurons, including 64 IB4-positive neurons, 69 IB4-negative small neurons and 64 large neurons, were processed for transcriptome analysis.

To obtain single-neuron transcriptomes, we examined the correlation between sequencing depth and the number of genes. We found that 1, 5, 10, 20, 30 and 40 million mapped reads per neuron detected 8 906 ± 79, 9 781 ± 127, 10 158 ± 133, 11 142 ± 223, 11 450 ± 256 and 11



694 ± 288 genes, respectively ($n = 3$, Supplementary information, Figure S2A). The number of mapped reads for each detected gene correlate linearly with the total number of mapped reads (Supplementary information, Figure S2B). Thus, the probability of detecting low-abundance genes depends on the depth of RNA-seq. Therefore, 30 million reads was considered the minimum number of mapped reads to achieve maximal mapping while maintaining efficiency. The average number of mapped reads was 58.2 million (ranging from 29.6 to 106.4 million) for a single neuron (Figure 1C). The number of detectable genes for each neuron ranged from 7 972 to 13 960 ($10\,950 \pm 1\,218$) per neuron and 20 794 in total (Figure 1D; Supplementary information, Figure S2C). To evaluate the effect of sequencing depth on gene number, six neurons were resequenced to a depth of 181 million reads – three times more than the average sequencing depth for all neurons. The libraries obtained from deeper sequencing runs shared 99.8% similarity with the libraries obtained by lower depth sequencing of the same neuron (Figure 1E). Thus, a sequencing depth of 30 million FPKM was sufficient.

Finally, to examine the quality of RNA-seq, cDNA from one neuron was divided into two equal parts and processed for RNA-seq. The transcriptomic datasets of two libraries derived from the same neuron shared 99.4% similarity (Figure 1F), suggesting a high quality of RNA-seq.

Gene modules identified by weighted gene co-expression network analysis

We performed principal component analysis on all 197 single-neuron transcriptomes as previously reported [32]. Genes with the highest loading in the first three principal components were analyzed by unsupervised hierarchical clustering. On the basis of this clustering, 2 043 differentially expressed genes were identified (\geq fivefold change, false discovery rate (FDR) < 5%) from 20 794 genes.

Weighted gene co-expression network analysis (WGC-

NA) is a method used to cluster highly interconnected genes into gene modules [35] that could be relevant for cell typing because highly correlated genes are likely to be detected in cells of the same cluster. The module eigengene (ME) is the first component of a given module and is considered to be the most representative gene expression in the module. The module membership of each gene, also known as eigengene-based connectivity (k_{ME}), is determined by correlating the gene expression profile with the ME and is highly related to the intramodular connectivity [35]. The most highly connected gene (i.e., the hub gene) inside the module tends to have high connectivity and a high k_{ME} . Performing WGCNA on the 2 043 differentially expressed genes identified 12 gene modules (Figure 1G; Supplementary information, Table S1). The yellow module has a strongly connected gene co-expression network, of which the gene encoding NPB (*Nppb*) is a hub gene ($k_{ME} = 0.930$). The MEs include genes for the oncostatin M receptor (*Osmr*), interleukin 31 receptor A (*Il31ra*), 5-hydroxytryptamine (serotonin) receptor 1F (*Htr1f*), neurotensin (*Nts*) and somatostatin (*Sst*, Supplementary information, Table S1). This network may be related to with itch because IL31RA mediates T-helper cell dependent itch [36], and OSMR acts as its heterodimer receptor [37]. The red module contains genes for peptides such as adenylate cyclase activating polypeptide 1 (*Adcyap1*, a hub gene; $k_{ME} = 0.908$), galanin (*Gal*) and tachykinin 1 (*Tac1*). Genes such as transmembrane protein 176A (*Tmem176a*) and 176B (*Tmem176b*), NPY receptor 1 (*Npy1r*), somatostatin receptor 2 (*Sstr2*), zinc-finger CCHC domain-containing 12 (*Zcchc12*) and TRP cation channel, subfamily V, member 1 (*Trpv1*) are the MEs of red module.

In the pink module, the gene for adenosine A2b receptor (*Adora2b*, $k_{ME} = 0.874$) is the hub gene; and genes for cerebellin 1 precursor protein (*Cbln1*), MRGPR B4 (*Mrgprb4*) and A3 (*Mrgpra3*), musculoskeletal, embryonic nuclear protein 1 (*Mustn1*) and protein tyrosine phosphatase, non-receptor type 6 (*Ptpn6*) are closely

Figure 1 Neuron sampling, RNA-sequencing and gene clustering. **(A)** Triple-immunofluorescent staining showing three major neuron subsets labeled by IB4, CGRP and NF200 in mouse lumbar DRG. Scale bar, 50 μ m (left) and 10 μ m (right). **(B)** A schematic of the workflow depicting rapid dissociation and isolation of individual DRG neurons for single-cell RNA-seq profiling. Images show a single neuron positive or negative for IB4 in the tip of a glass pipette. Scale bar, 20 μ m. **(C)** Distribution of mapped reads in 197 neurons. **(D)** Number of detected genes in 197 neurons. **(E)** In samples obtained from individual DRG neurons, the number of genes detected at a given sequencing depth (in this case, 52.7 million mapped reads) was correlates highly with that obtained at a triple depth (181 million mapped reads). **(F)** Transcript expression levels (x and y axes: log2-scale) in two samples obtained by equally dividing cDNA from DRG neuron No. 72 were correlate highly to each other. **(G)** The cluster dendrogram of 2 043 differential genes represents the co-expression modules identified by WGCNA. Modules of highly interconnected groups of genes correspond to branches and are labeled in different colors. **(H)** ISH shows the distribution of the representative module genes including *Tmem176b*, *Nxph1*, *Il31ra*, *Baiap2l1*, *Cpne6*, *S100b*, *Tmem45b* and *Fam19a4* of the red, purple, yellow, green, blue, turquoise and black modules, respectively, in lumbar DRG. Scale bar, 50 μ m.

associated (Supplementary information, Table S1). In the black module, genes for family with sequence similarity 19, member A4 (*Fam19a4*) and TH (*Th*) serve as hub genes ($k_{ME} = 0.920$ and 0.904 , respectively), whereas the genes for zinc-finger protein 521 (*Zfp521*), cholecystokinin A receptor (*Cckar*) and POU domain, class 4, transcription factor 2 (*Pou4f2*) are among the MEs (Supplementary information, Table S1). In the turquoise module, MEs include genes for secretogranin III (*Scg3*, $k_{ME} = 0.934$), cortixin 3 (*Ctxn3*, $k_{ME} = 0.930$), CD24a antigen (*Cd24a*, $k_{ME} = 0.917$), CD55 antigen (*Cd55*, $k_{ME} = 0.916$), synaptoporin (*Synpr*, $k_{ME} = 0.915$), coiled-coil domain-containing protein 68 (*Ccdc68*, $k_{ME} = 0.892$) and MRGPRD (*Mrgprd*, $k_{ME} = 0.889$).

The blue module is characterized by genes encoding the S100 protein, beta polypeptide (*S100b*, $k_{ME} = 0.934$, the hub gene) and contain MEs such as tubulin alpha 4a (*Tuba4a*), ly6/neurotoxin 1 (*Lynx1*), copine VI (*Cpne6*), FXYP7 (*Fxyd7*), potassium voltage-gated channel, shaker-related subfamily member 1 (*Kcna1*) and neurofilament, heavy polypeptide (*Nefh*) encoding NF200 (Supplementary information, Table S1). The hub genes in the green module encode BAI1-associated protein 2-like 1 (*Baiap2l1*, $k_{ME} = 0.922$) and neurotrophic TRK type 2 (*Ntrk2*, $k_{ME} = 0.911$), whereas the MEs include genes for 5-hydroxytryptamine receptor 1D (*Htr1d*), potassium channel subfamily K member 1 (*Kcnk1*), Ca²⁺-dependent activator protein for secretion 2 (*Cadps2*), calcium channel, voltage-dependent, T type, alpha 1H subunit (*Cacna1h*) and N-terminal EF-hand calcium-binding protein 2 (*Necab2*). The green module thus represents large neurons expressing *Ntrk2* [16]. The purple module contain the parvalbumin gene (*Pvalb*), a marker of large proprioceptive neurons [38], and the MEs include neurexophilin 1 (*Nxph1*).

In situ hybridization (ISH) confirmed the expression of *Il31ra* (yellow module), *Tmem45b* (turquoise module) and *Fam19a4* (black module) in small neurons; *Nxph1* (purple module) and *Tmem176b* (red module) in small and some large neurons; *Cpne6* and S100 calcium-binding protein A16 (*S100a16*, blue module) as well as *Pvalb* (purple module) in large neurons and a subset of small neurons, and *Baiap2l1* (green module) in large neurons (Figure 1H; Supplementary information, Figure S3A). Thus, the above gene modules are present in DRG neurons. The genes in brown module, including sclerostin domain containing 1 ($k_{ME} = 0.973$, *Sostdc1*), fatty acid-binding protein 7 ($k_{ME} = 0.944$, *Fabp7*), ATPase, Na⁺/K⁺ transporting, beta 2 polypeptide ($k_{ME} = 0.779$, *Atp1b2*) and TIMP metalloproteinase inhibitor 3 ($k_{ME} = 0.660$, *Timp3*), are found in satellite cells (Supplementary information, Figure S3B). Genes in the magenta module are

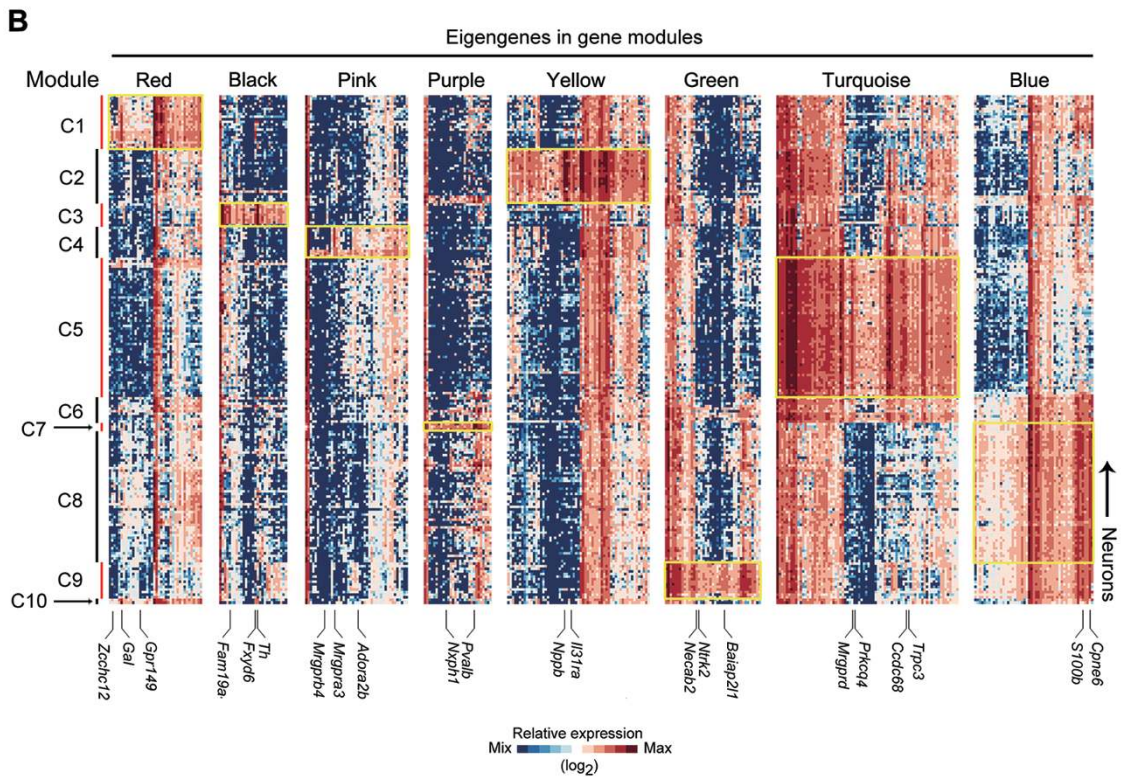
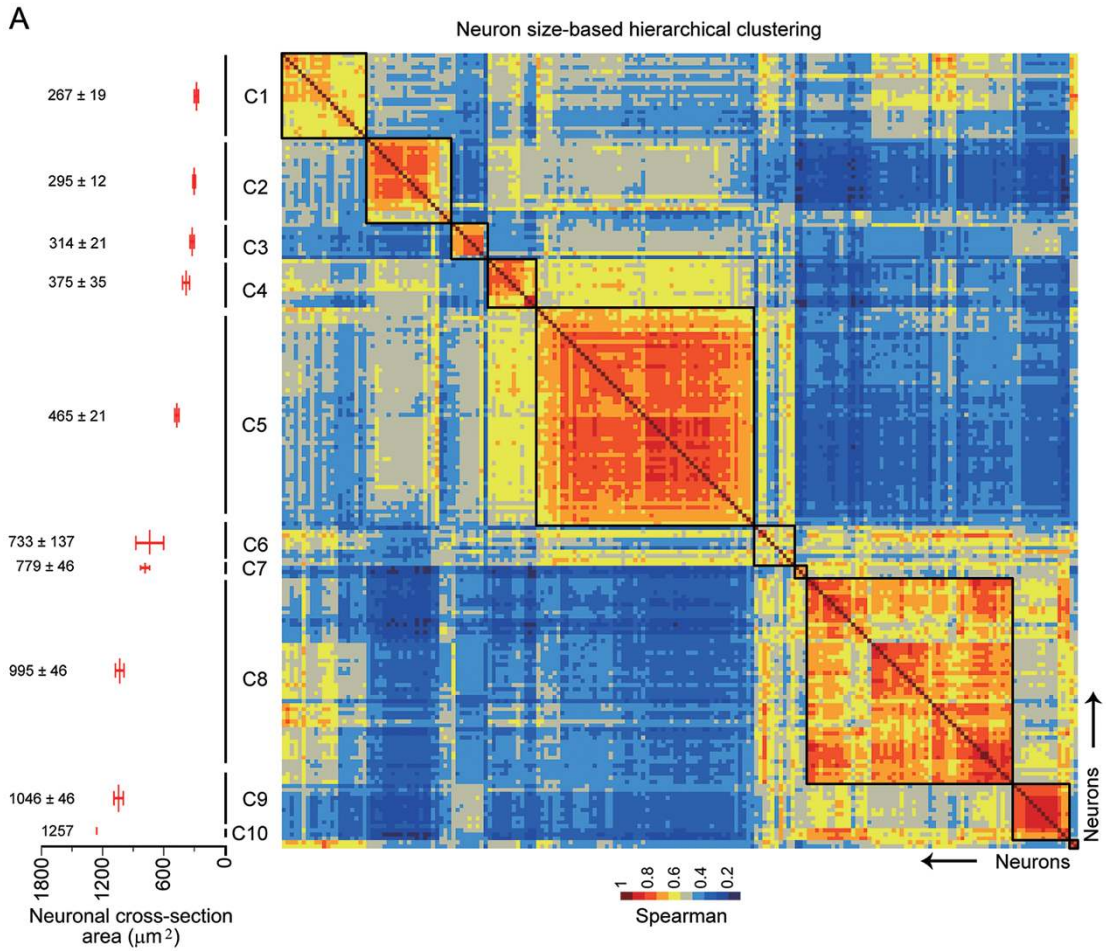
not highly connected. Genes in the green-yellow modules are expressed in blood cells, and genes in the grey module are considered to be non-module genes. In total, WGCNA identifies eight neuronal gene modules and one glial gene module.

Ten clusters revealed by neuron size-based hierarchical clustering

To analyze the neuronal clusters, 1 745 differentially expressed genes (≥ 5 -fold change, excluding 185 eigen-genes of satellite cells (brown module, $k_{ME} > 0.6$), all 71 genes of grey module and 42 genes of green-yellow module) from 197 neurons were subjected to principal component analysis. Unsupervised hierarchical clustering analysis was performed on genes with the highest loadings in the first three principal components. This unbiased approach detects eight cell populations (Supplementary information, Figure S4). Each cluster (C) is named in accordance with the average area of its clustered neurons (Supplementary information, Figure S4). *Gal*-containing small neurons are sorted into C1, whereas exceptionally large neurons in the same cell population are separated into C10. C7 neurons are segregated from C8 neurons in the same cell population because they are only weakly correlated. Thus, we here identified 10 clusters in the hierarchical neuronal matrix, including small neurons (averaged area: ~ 267 - $733 \mu\text{m}^2$) in C1-C6 and large neurons (averaged area: ~ 779 - $1257 \mu\text{m}^2$) in C7-C10 (Figure 2A). The clusters C1-C6 and C8-C10 emerge within the first 98 neurons included for analysis, whereas C7 is detectable only when the number of neurons is doubled.

A cell-cell correlation matrix based on 1 745 differentially expressed genes reveal that C9 neurons are highly correlated among themselves (value of Spearman correlation (VSC): 0.780 ± 0.007), followed by C5 (0.726 ± 0.002), C7 (0.705 ± 0.012), C4 (0.700 ± 0.007), C8 (0.681 ± 0.002), C2 (0.671 ± 0.006) and C3 (0.659 ± 0.020) neurons (Figure 2A). Small neurons in C1 and C6 are moderately correlated (VSC: 0.618 ± 0.003 and 0.574 ± 0.011 , respectively). The correlation among neurons in a cluster suggests the intra-cluster homogeneity of neurons, making them more likely to comprise a specific subtype. The correlation matrix also shows the correlation between neurons within different clusters (Figure 2A). C4 neurons are closely correlated with C5 neurons (VSC: 0.578 ± 0.001), whereas large neurons in C8 and C9 are weakly correlated with C1 and C3, respectively (Figure 2A). These inter-cluster correlation patterns could represent the molecular basis of responses by multiple neuron clusters to a specific stimulus.

We next correlated the MEs ($k_{ME} > 0.700$) identified



by WGCNA with neuron clusters. Generally, the MEs of red module ($k_{ME} > 0.600$) are found in C1 and C10, yellow module in C2, black module in C3, pink module ($k_{ME} > 0.500$, *Mrgpra3* ($k_{ME} = 0.566$)) in C4, turquoise module in C5, magenta module in C7 ($k_{ME} > 0.600$), blue module in C8 and green module in C9 (Figure 2B). Analysis of the eigengenes of each cluster reveals *Gal*, *Adcyap1* and *Npy1r* for C1; *Nppb*, *Il31ra*, *Sst* and *Htr1f* for C2; *Th*, *Fam19a4*, *Zfp521* and *Cckar* for C3; and *Adora2b*, *Mrgprb4* and *Mrgpra3* for C4 (Figure 2B; Supplementary information, Table S1). *Mrgprd* and the gene encoding lysophosphatidic acid receptor 3 (*Lpar3*) are the eigengenes for C5; *Nxph1* for C7; and *S100b*, trafficking protein particle complex 3-like (*Trappc3l*), Cingulin-like 1 (*Cgnl1*), *Cpne6* and *Nefh* for C8. The eigengenes of C9 included *Ntrk2* and genes encoding *Baiap211*, *Necab2* and *Cadps2*. Thus, differential expression of the eigengenes is important for the formation of neuron clusters.

Fourteen neuron subclusters and the “hybrid state”

According to the cluster tree, C1, C2, C4, C5, C6, C8 and C9 could be further classified into two subclusters each, forming a total of 14 subclusters (Supplementary information, Figure S4) that are also indicated by the correlation matrix (Figure 2A). Neuron subclusters can be generated by additional gene networks in the same gene module. The neuron correlation matrix shows that small neurons in C4 can be further divided into two subclusters, C4-1 and C4-2 (Figure 3A). The differential genes shown in the corresponding module of WGCNA reveals that C4-1 neurons express *Mrgpra3*, whereas C4-2 neurons express both *Mrgpra3* and *Mrgprb4* (Figure 3A). Both single-cell real-time PCR of the sequenced neurons in C4 and double-fluorescent ISH proved the existence of the C4 subclusters (Figure 3B-D). *Mrgprb4* is present in ~22.1% of *Mrgpra3*-expressing neurons ($n = 104$, Figure 3C). The gene co-expression network of *Mrgprb4* is separate from that of *Mrgpra3*, although they are found in the same module (Figure 3E). Thus, this is a canonical case of subclustering based on an additional gene network.

Neuron subclusters can also be formed by the addition of functional genes from other gene modules. Some neurons within a cluster expressed the representative genes of other clusters, a situation referred to “hybrid

state.” In this manner, C2 is divided into subclusters C2-1 and C2-2 (Figure 3F). In the corresponding modules of WGCNA, C2-2 neurons expressed *Nefh*, *S100b*, *Cpne6* and *Pvalb* (Figure 2B and 3F). The level of *Nppb* and *Il31ra* in C2-2 is lower than that in C2-1, and *Pvalb* and *Nefh* were expressed strongly in C2-2 but are not detectable in C2-1 as shown by single-cell real-time PCR (Figure 3G). To confirm the existence of C2 subclusters, IB4-negative small neurons (area < 314 μm^2), selected randomly from dissociated DRG neurons, were analyzed by single-cell PCR (Figure 3H). We found that *Pvalb* is expressed in ~10% of *Nppb*-containing neurons ($n = 30$). Double fluorescent ISH proved that *Il31ra* is co-expressed with *S100b* and *Cpne6* in small DRG neurons (Figure 3I). Thus, C2-1 and C2-2 represent two distinct subclusters, and C2-2 represents a hybrid state. The hybrid state is found to be a property of some clusters or subclusters of small neurons that are genetically correlated with large neurons, suggesting that they might be small neurons with A δ -fibers (Figure 3J).

The small neurons in C1 are correlated with the large neurons in C8. C8-2 neurons are correlated moderately with C1-1 neurons and weakly with C1-2 neurons (Supplementary information, Figure S5A). Analysis of differentially expressed genes shows that C1-1 neurons express high levels of claudin 9 (*Cldn9*) and acid-sensitive ion channel 3 (*Asic3*), whereas C1-2 neurons expressed high levels of *Zcche12* and *Sstr2* (Supplementary information, Figure S5B and S5C). C8-1 neurons express high levels of *Ntrk3*, *Ret*, *Htr1d* and *Fam19a2*, while C8-2 express high levels of *Ntrk1*, *Calca* and *Htr3a* (Supplementary information, Figure S5C).

C5-1 represents a subcluster of *Mrgprd*-expressing small neurons that also express *Gal*, *Adcyap1* and *Tmem176b* (Supplementary information, Figure S5D and S5E), suggesting a molecular correlation with C1. The neuron matrix also indicates that C6 neurons are correlated with both C5 neurons and large neurons (Figure 2A). Statistical analysis shows that C6 neurons expressed high levels of *Nefh* and *S100b* and a low level of *Mrgprd* (Supplementary information, Figure S5E and S5F). Double fluorescent labeling shows that *Mrgprd* is expressed in *S100b*- or NF200-containing DRG neurons (Supplementary information, Figure S5G and S5H). These results indicate the “hybrid state” of this neuron cluster.

Figure 2 Clustering of DRG neurons. **(A)** A heatmap of size-based hierarchical clustering shows the color-coded correlation matrix of 197 neurons based on the mRNA profiles of 1 745 differentially expressed genes (> 5 fold change). Ten clusters of strongly correlated neurons are marked with black frames and shown with the given names and their averaged area of neurons to the left. **(B)** The WGCNA heatmap shows the expression profiles of module genes. The eigengenes of neuronal clusters are indicated by yellow frames.

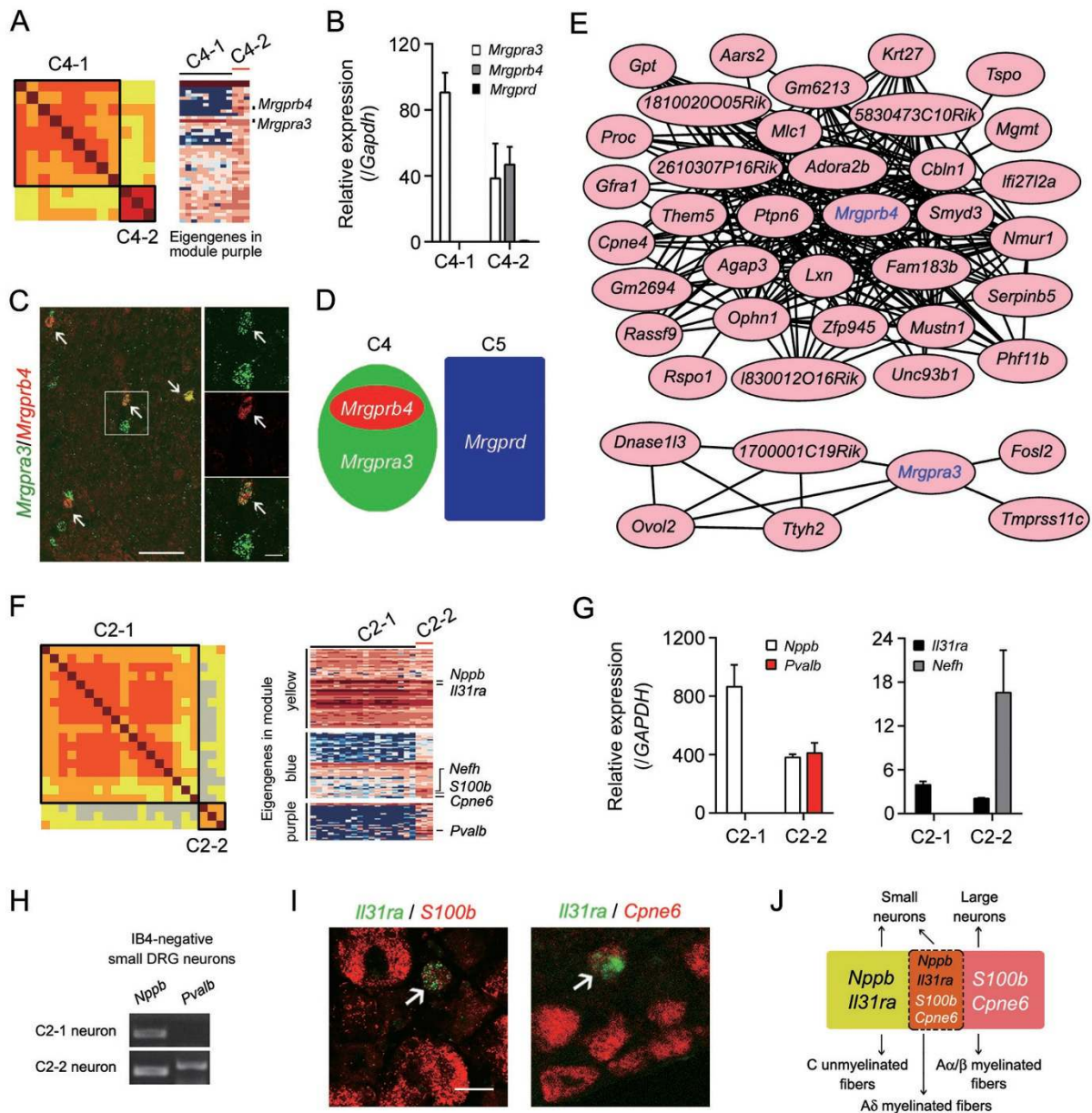


Figure 3 DRG neuron subclusters and the “hybrid states” of subclusters. **(A)** An enlargement of C4 on the heatmap of the correlation matrix and WGCNA eigengenes (Figure 2A and 2B) reveal two distinct subclusters, C4-1 and C4-2, outlined by black frames. **(B)** Single-cell real-time PCR confirming the differential expression of *Mrgpra3*, *Mrgprb4* and *Mrgprd* in C4-1 and C4-2 neurons ($n = 3$). **(C)** Double fluorescent ISH showing the co-expression of *Mrgprb4* in a subpopulation of *Mrgpra3*-positive small DRG neurons (arrows). Scale bars, 100 μm (left) and 20 μm (right). **(D)** Differential distribution of *Mrgpra3*, *Mrgprb4* and *Mrgprd* in C4 and C5. **(E)** Gene co-expression network identified by WGCNA shows that the *Mrgpra3*-containing network is separate from the *Mrgprb4* network in the pink module. **(F)** An enlargement of C2-1 and C2-2 in the heatmap of the correlation matrix and WGCNA (Figure 2A and 2B). **(G)** Single-cell PCR confirming the differential expression of representative genes in C2-1 and C2-2 neurons ($n = 3$). **(H)** Single-cell PCR showing expression of *Pvalb* in an *Nppb*-positive, IB4-negative neuron. **(I)** Double fluorescent ISH showing co-expression of *Il131ra* with *S100b* or *Cpne6* in a small DRG neuron (arrow). Scale bar, 20 μm . **(J)** The correlation among *Nppb*, *Il131ra*, *S100b* and *Cpne6* expression in clusters of small and large neurons and their predicted relationships with different types of afferent fibers.

The correlation of C9-2 with C8 is stronger than that of C9-1 with C8 (Supplementary information, Figure

S5I). C9-2 neurons express higher levels of *S100b*, *Asic3* and *Cgn1l* but lower levels of *Ntrk2*, *Baiap21l* and

Cadps2 compared with C9-1 (Supplementary information, Figure S5J). In summary, clusters and subclusters of DRG neurons are revealed by analyzing single-cell transcriptome data and the neuronal correlation matrix.

Differential genes of neuron clusters

To identify differentially expressed genes within neuron clusters and subclusters, we compared neuron clusters using the Fisher test. C1 neurons contain high levels of *Tac1* and *Calca* and represented a population of the traditionally classified PEP subset of DRG neurons, whereas C2 expressed a high level of *Nppb*. Comparisons between C1 and C2 reveal 152 differentially expressed genes (≥ 10 -fold change, FDR < 0.05), 83 of which are highly expressed in C1 (e.g., *Gal* and *Adcyap1*) and 69 in C2 (e.g., *Nppb*, *Il31ra*, *Nts* and *Htr1f*; Supplementary information, Figure S6A). Moreover, the representative genes that are differentially expressed between subclusters within C1 and C2 are shown in Supplementary information, Figure S6A. For example, *Cldn9* and *Asic3* are expressed mainly in C1-1, whereas *Zcchc12* and *Sstr2* are expressed in C1-2 (Supplementary information, Figure S6A). The expression pattern of C2 also clearly revealed two subclusters.

C3 neurons are *Th* positive, whereas C4 neurons expressed *Mrgpra3* and C5 neurons contained *Mrgprd*. Traditionally, these neurons are classified as NP small neurons. Comparisons among C3, C4 and C5 neurons revealed representative differentially expressed genes (≥ 10 -fold change, FDR < 0.05 , Supplementary information, Figure S6B), such as GDNF family receptor alpha 2 (*Gfra2*), *Fam19a4*, *Zfp521* and *Th* in C3; *Gfra1*, *Mrgpra3*, *Mrgprc11* (also known as *Mrgprx1*), the gene for R-spondin homolog (*Rspo1*) and *Mrgprb4* in C4; and *Gfra1*, *Gfra2*, *Mrgprd*, *Lpar3*, protein kinase C theta (*Prkcg*) and *Mrgprb5* in C5.

Large neurons in C7, C8 and C9 contained high levels of *Nefh*. Comparisons among C7, C8 and C9 produce representative differentially expressed genes (≥ 10 -fold change, FDR < 0.05 , Supplementary information, Figure S7), including *Nxph1* and genes for phospholamban (*Pln*) and wingless-type MMTV integration site family member 7A (*Wnt7a*) in C7; and *Ntrk2*, *Baiap211*, *Cadps2*, *Necab2* and the gene encoding Ly6/Plaur domain-containing 1 (*Lypd1*) in C9. C10 neurons express not only *Gal* and *Adcyap1*, the representative genes for C1, but also *Th* for C3, *Rspo1* for C4 and *Cpne6* for large neurons.

Markers of neuron clusters

To be considered a marker gene, the prerequisites are selectivity, a relatively high level of expression, and utility in multiple experimental approaches. *Gal*, *Nppb*, *Th*

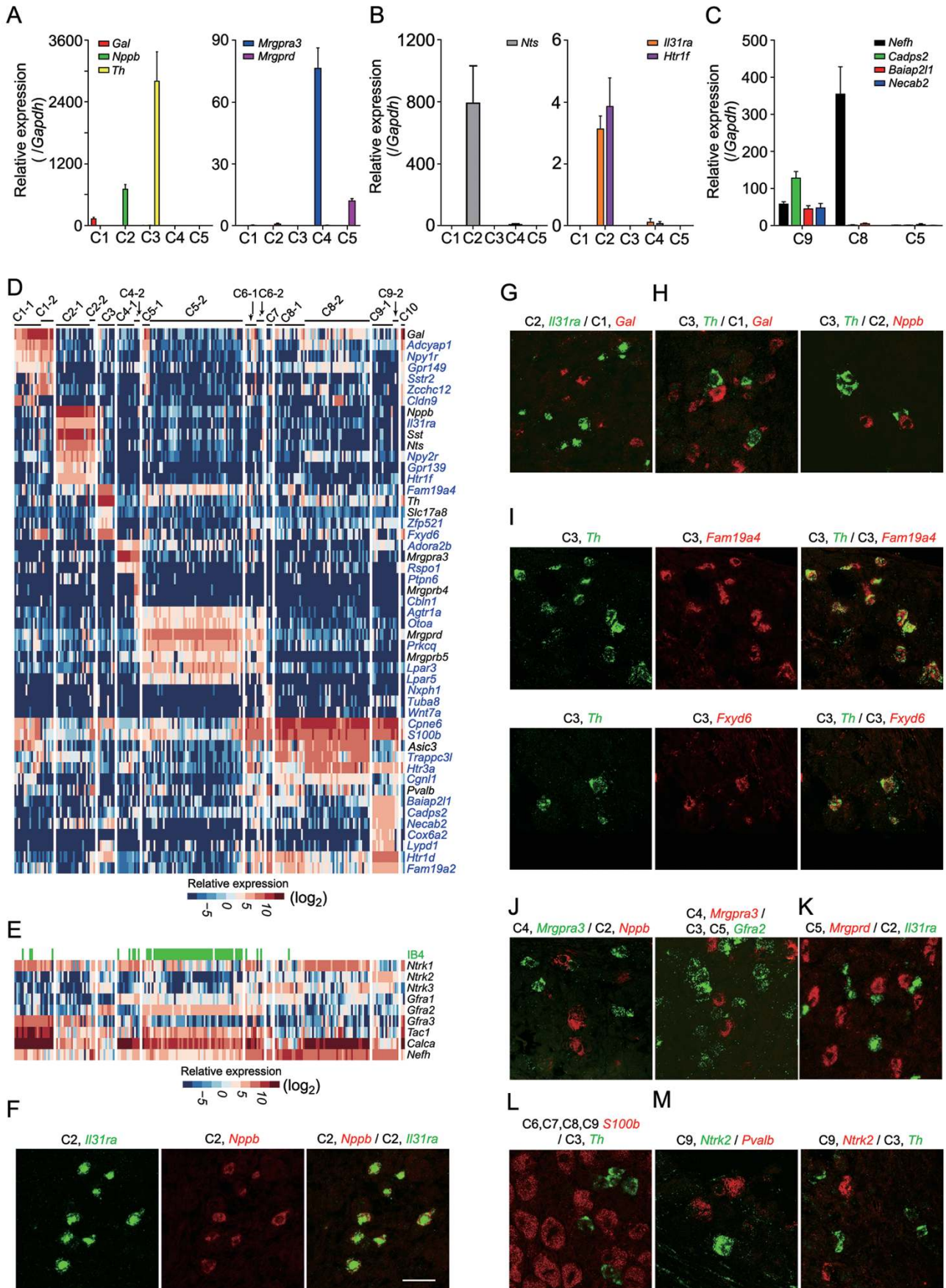
and *Mrgpra3* are selected as markers for C1, C2, C3 and C4, respectively. Single-cell real-time PCR show that *Gal*, *Nppb*, *Th* and *Mrgpra3* are selectively expressed in C1, C2, C3 and C4 neurons, respectively (Figure 4A). Each cluster could have several markers, e.g., *Nts*, *Il31ra* and *Htr1f* for C2 (Figure 4B) and *Cadps2*, *Baiap211* and *Necab2* for C9 (Figure 4C). The subclusters C1-1 and C1-2 are marked by *Gal* together with *Cldn9* and *Asic3*, and *Gal* combined with *Zcchc12* and *Sstr2*, respectively (Figure 4D). C2-2 is separated from C2-1 by the expression of *Pvalb* and *S100b* (Figures 3F-3H and 4D). Both *Mrgpra3* and *Mrgprb4* are combined markers for C4-2, whereas *Mrgpra3* alone is a marker for C4-1 (Figures 3A and 4D).

Mrgprd is expressed highly in C5 and moderately in C6 (Figure 4A and 4D). In *Mrgprd*-marked C5 neurons, the presence of *Gal*, *Sstr2*, *Zcchc12*, *Asic3* and *Gfra3* distinguishes C5-1 from C5-2 neurons (Figure 4D and 4E). C6 neurons are distinguished from C5 by their expression of *Cpne6* and *S100b* (Figure 4D). C6-1 contain *Gal*, whereas C6-2 express *Prkcg* and *Lpar3* at high levels (Figure 4D). Notably, abundant *Gal* expression is also found in C5-1 and C6-1 neurons. C5-1 differ from *Gal*-marked C1 by the expression of *Mrgprd*, whereas C6-1 is differentiated by *Mrgprd* and high levels of *Cpne6* and *S100b* expression.

For large DRG neurons, *Nxph1*, *Trappc3l* and *Cadps2* are markers for C7, C8 and C9, respectively (Figure 4D). C8-1 express high levels of *Ntrk3* and *Htr1d*, whereas C8-2 contain high levels of *Ntrk1* and *Htr3a* (Figure 4D and 4E). C9-2 is distinguished from C9-1 by higher levels of *Asic3* and *Cgn11* (Figure 4D). The large neurons of C10 express *Gal*, *Th*, *Rspo1* and *Kcnk2* (Figures 4D and 5A).

Traditional markers of neuronal subsets, such as *Tac1* and *Calca* for the PEP subset, IB4 for the NP subset and *Nehf* for large neurons, are observed in multiple clusters (Figure 4E). IB4 labels C4 and C5 neurons as well as some neurons in C1 and C6. Some IB4-positive neurons in C4 and C5 also express *Tac1* and *Calca*. Moreover, many recently suggested marker genes, such as *Cntnap2* (contactin-associated protein-like 2) [34], are also found in multiple clusters (Supplementary information, Figure S8A).

The growth factor receptors *Ntrk* and *Gfra* are also not restricted to a single neuron cluster. However, differential combinations of *Ntrk* and *Gfra* are consistent with the clustering of DRG neurons (Figure 4E). *Ntrk1* and *Gfra3* are present in C1, C10 and C2 neurons; *Ntrk1* and *Gfra1* in C4 neurons; *Gfra2* and low levels of *Gfra1* expression in C5 neurons; *Gfra1* and high levels of *Ntrk3* expression in C8-1 neurons; *Ntrk1* and low levels of *Ntrk3* expres-



sion in C8-2 neurons, and *Ntrk2* and low levels of *Ntrk3* expression in C9 neurons.

Double fluorescent ISH was used to examine the distribution of the markers. *Il31ra* is co-expressed with *Nppb* but not *Gal* in small neurons (Figure 4F and 4G). Almost all *Th*-positive neurons express *Fam19a4* and *Fxyd6*, but not *Gal* and *Nppb* (Figure 4H and 4I). *Mrgpra3* is not co-expressed with *Nppb* or *Gfra2* (Figure 4J). *Mrgprd* is not co-expressed with *Il31ra* in DRG neurons (Figure 4K). *S100b* is not expressed in *Th*-positive neurons (Figure 4L), and *Pvalb* and *Th* are absent from *Ntrk2*-positive large neurons (Figure 4M). Thus, markers identified in this study can be used to identify neuron types in DRG tissue.

Functional phenotypes of neuron clusters

We next asked whether neuron clusters represent neurons of different sensory modalities. A heterogeneity analysis of signaling molecules in neuron clusters as well as previous reports implies the functionalities of these neuron clusters (Figure 5A; Supplementary information, Data S1). TRP channels are involved in temperature sensing [12]. *Trpv1* is expressed strongly in C1-1 and moderately in C1-2 and C2-1 (Figure 5A; Supplementary information, Figure S8B). *Trpv2* is expressed in almost all clusters but at a low level in C8, whereas *Trpv4* is present in some C3 neurons. *Trpm8* is expressed in some neurons in C1-2, C6-1 and C8-2, whereas *Trpa1* is expressed in C1-2, C4-2, C5-2, C6-2 and C10. *Ano1* is present in C2-2 and C7. Thus, several neuron clusters are likely to be differentially sensitive to temperature. The cross-cluster expression of genes encoding mechanoreceptive molecules such as *Trpcs*, *Kcnks*, *Asic3* and *Piezo2* suggests that mechanical stimuli are sensed by neurons in multiple clusters (Figure 5A). The predicted sensitivity of clusters to chemicals depends on the differential expression of chemical receptors (Figure 5A). The differential expression of pruritogen receptors and NPB encoded by *Nppb* suggests that C2, C4 and C5 selectively transduce itch signals generated by pruritogens (Figure

5A). For example, C4 contains mRNAs encoding the histamine H1 receptor (*Hrh1*), *Mrgprc11* and *Mrgpra3*, and therefore could be specific for histamine-, BAM8-22- and chloroquine-induced itch.

To evaluate the functional phenotypes of neuron clusters, we carried out *in vivo* whole-cell patch clamp recording of single neurons selected at random in the L5 DRG of mice upon application of a noxious mechanical stimulus (pinch), heat (42-52 °C), cold (ice water) or innocuous mechanical stimulus (brush and pressure) to the plantar skin of the hindpaw. The appropriate peripheral stimulus induced an inward Na⁺ current and generated an action potential in the responsive neuron (110 of 129 recorded neurons). mRNA diffusing into the internal solution of the recording pipette was collected and processed for single-cell quantitative real-time PCR to evaluate the expression levels of neuron-type markers, including *Gal*, *Zcchc12*, *Cldn9*, *Nppb*, *Nts*, *Sst*, *Th*, *Mrgpra3*, *Mrgprb4*, *Mrgprd*, *Nefh*, *Pvalb*, *Nxph1* and *Baiap211* (Figure 5B). This analysis was successfully carried out for 69 responsive neurons (Figure 5C and 5D).

Cluster sorting showed that C1, C2, C4, C5, C6 and C7 are involved in sensing heat and pinch (Figure 5D) but have different preferred stimuli. Most recorded C1-2 neurons responded to heat and pinch. Over half of the neurons in C6 and C7 were activated by pressure and pinch, whereas most neurons in C4 responded to pinch. Furthermore, these results suggest the sensory selectivity of neuron clusters. C2 neurons responded to heat and pinch but not to pressure. All recorded C3 neurons were stimulated by brush but not heat, and these neurons were also responsive to pressure and pinch. C9 neurons primarily responded to pressure and pinch. Neurons responding to cold are scattered in C2, C4-1 C5 and C6, but none of these neurons were found to respond only to cold. These results suggest that the transcriptional heterogeneity among neuronal clusters contributes to the functional specificity and probability of neurons to respond to cutaneous stimuli. The sensory thresholds of neurons in the same cluster could be different due to intra-cluster

Figure 4 Evaluation of neuron clusters markers. **(A)** Single-cell real-time PCR confirms the expression of marker genes in C1 (*n* = 6), C2 (*n* = 6), C3 (*n* = 6), C4 (*n* = 3) and C5 (*n* = 6) neurons. **(B)** Single-cell PCR reveals the expression of marker genes in C1 (*n* = 3), C2 (*n* = 6), C3 (*n* = 3), C4 (*n* = 3) and C5 (*n* = 6) neurons. **(C)** Single-cell PCR reveals the expression of marker genes in C5 (*n* = 3), C8 (*n* = 8) and C9 (*n* = 4) neurons. **(D)** A heatmap shows the expression patterns of cluster markers. New markers are indicated in blue. **(E)** A heatmap shows the expression patterns of *Calca*, *Tac1* and *Nefh*, and the neurotrophin receptors *Ntrk* and *Gfra* in the neuron clusters. IB4 labeling is marked with a green square. **(F, G)** Double fluorescent ISH shows the cluster-specific expression of markers in DRG neurons. *Il31ra* is co-expressed with *Nppb* but not *Gal* in small neurons. Scale bar, 50 μm. **(H, I)** *Th* is co-expressed with both *Fam19a4* and *Fxyd6* but not *Gal* and *Nppb* in small neurons. **(J)** *Mrgpra3* is not co-expressed with *Nppb* and *Gfra2* in small neurons. **(K)** *Mrgprd* is not co-expressed with *Il31ra* in small neurons. **(L)** *S100b* and *Th* are not co-expressed in DRG neurons. **(M)** *Pvalb*-positive large neurons and *Th*-positive small neurons do not contain *Ntrk2*.

diversity in gene networks and their activity states.

Types and subtypes of somatosensory neurons

We defined neuron types by integrating transcriptomic, morphological and functional characteristics. In the hierarchy of pseudo-unipolar DRG neurons, small neurons are sorted into 6 types and 10 subordinate subtypes, whereas large neurons are grouped into 4 types and 4 subordinate subtypes (Figure 6A and 6B). The neuron types are highly related to neuron size (Figure 6B), suggesting a principle of somatosensory organization. Neuron types C6, C7 and C8 and all subtypes therein have not been previously described (Figure 6B). These subtypes are distinguished by the expression of either additional genes or the representative genes of large neurons. Each neuron type and subtype has a distinct pattern of gene expression and several type-selective markers, including many newly defined markers (Figures 4D and 6B). The traditional markers, such as *Tac1*, *Calca*, *Nehf*, *IB4*, as well as several newly proposed markers, such as plexin C1 (PLXNC1) and lactate dehydrogenase B (LDHB) [34], can not uniquely identify specific types of DRG neurons (Figure 6C; Supplementary information, Figure S8A), although *Th* has been demonstrated to be a marker of one type of MR.

Previous and current experimental data suggest that small neurons in C1, C2, C4, C5 and C6 are generally sensitive to noxious mechanical stimulus and heat, suggesting that they act as MHNs (Figures 5A, 6D and 6E). In addition, they can be further specified into sensory phenotypes. C1, C4, C5 and C6 represent MHNs that are moderately sensitive to innocuous mechanical stimuli such as pressure. C2 may contain MHNs insensitive to innocuous mechanical stimuli (MI) but sensitive to itch-provoking chemicals (itch-sensitive, IS). C4 and C6 preferentially respond to noxious mechanical stimuli. C4 and C5 are additionally IS neurons acting as MHNs (IS). All MHNs appear to be polymodal nociceptors because they are also chemoreceptive. C3 neurons are the C-fiber low-threshold MR (C-LTMR), which are also sensitive to noxious mechanical stimuli but not to heat. Large DRG neurons are categorized into four types, including C9 as MNs. C7 contains MHNs that are highly sensitive to both noxious and innocuous mechanical stimuli (mechanically sensitive, MS). C8 neurons are typical MRs because the thermoceptive channel *Trps* is mostly absent in this type.

Moreover, C1, C2, C4, C5, C6, C8 and C9 are further classified into 14 subtypes (Figure 6B). C1-2 neurons are highly sensitive to both heat and noxious mechanical stimuli, whereas C1-1 could be more mechanoreceptive (Figures 5A and 6D). C4-2 neurons are sensitive to massage-like stroking, which is mediated by the MRG-

PR B4-signaling network, whereas C4-1 neurons are low-threshold mechanoreceptive. C2-2 neurons may be A δ -fiber MHN (IS). Cold-sensitive neurons are scattered among C2-1, C4-1 and C5, which can also act like MHNs. This classification suggests a new framework for classifying somatosensory neuron types (Figure 6E).

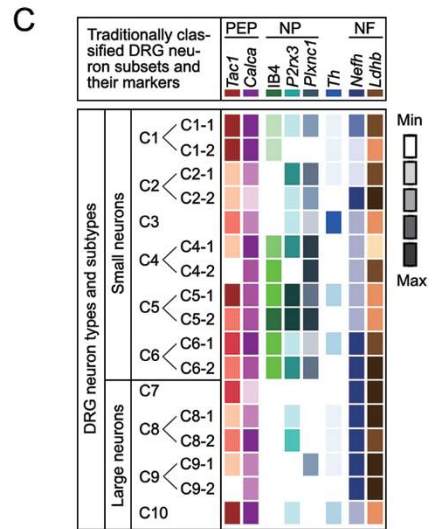
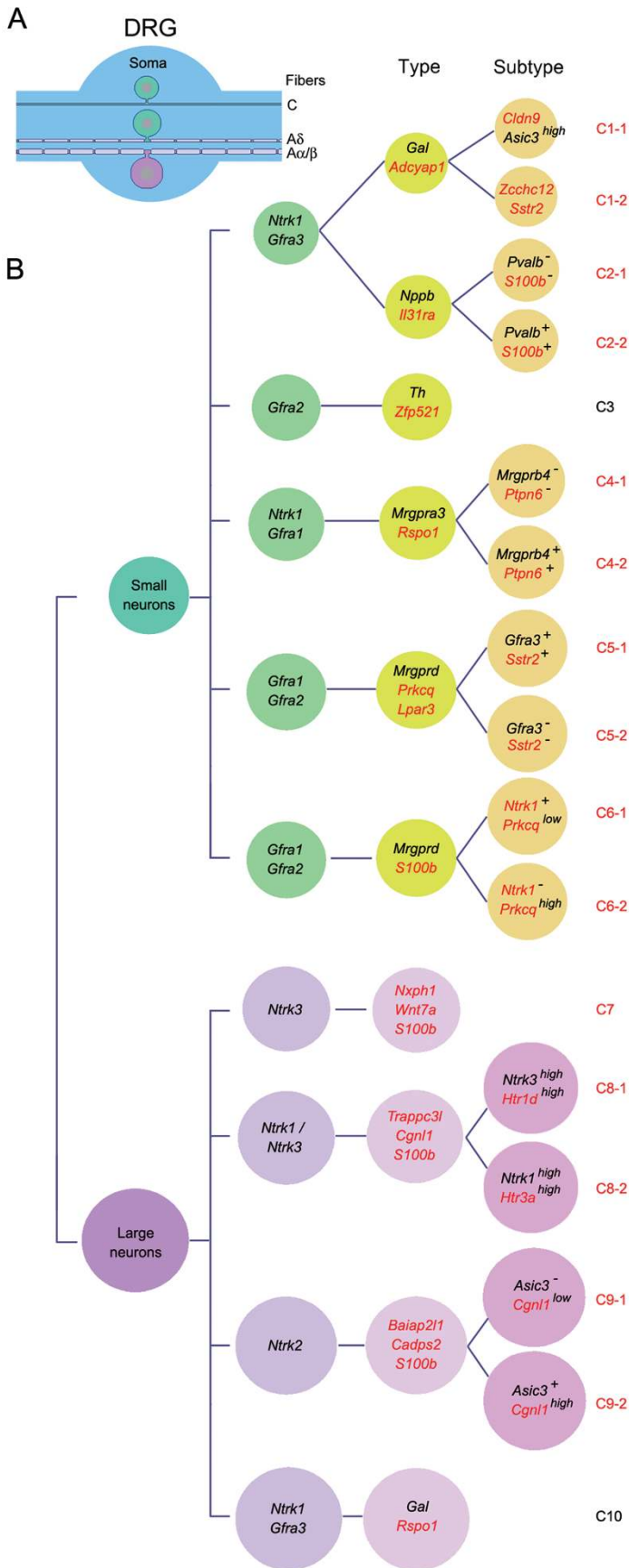
Discussion

Tissue-specific approach incorporating single-cell RNA-seq for neuron-typing

The present study presents a DRG-specific approach to neuron sampling for neuron typing that incorporates single-cell RNA-seq with other methods. First, neuron sizes are recorded because the differences between small and large neurons correlate with the projections of afferent fibers in the spinal cord and their functions. The correlation between neuron size and neuron type was proven by transcriptome analysis. Second, the principal neuron subsets classified according to *IB4* labeling and their proportions are incorporated into this approach, enabling the maximal coverage of neuron types. One of the technical concerns is the number of sequenced neurons required to cover all neuron types in the DRG. The transcriptome similarity of lumbar DRG in mice and the traditionally classified neuron subsets aided our neuron typing through single-cell RNA-seq in a number of neurons. The required number of neurons was determined by subsequent clustering analysis; most neuron clusters and gene modules of DRG neurons appeared when ~100 neurons were sequenced. The clusters and, in particular, the subclusters became more apparent when more neurons were analyzed. Therefore, the current knowledge regarding neuronal characteristics is important for the design of tissue-specific approaches for neuron typing.

Neuron-typing with high-coverage single-cell RNA-seq

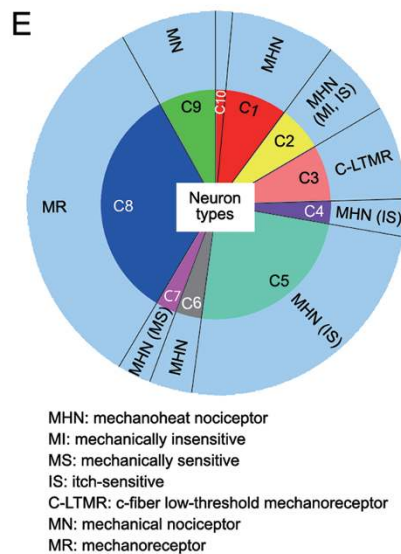
To characterize the transcriptome, a high-coverage of at least 30 million mapped reads per neuron appears to be adequate to detect the majority of expressed genes. Moreover, the number of detected genes and gene copies in an individual neuron should be confirmed to ensure the data quality of the transcriptome analysis. The complete transcriptome of a single neuron is required to define neuron clusters by integrating WGCNA and neuron hierarchical clustering. A recent study has suggested that low-coverage single-cell RNA-seq (i.e., 0.27 million reads for ~4 600 detected genes) is sufficient to distinguish cell types, including blood cells, dermal cells and neurons [30]. Similarly, low-coverage single-cell RNA-seq with a large number of neurons has been applied for cell typing in the mouse DRG [34]. However, this approach resulted in



D

Types and subtypes	Functional phenotypes					
	Nociceptors		Mechano-receptor		Pruriceptors	
	Thermal	Mechanical	Pressure	Brush		
	Cold	Heat				
C1	-	M	M	M	L	-
C1-1	-	-	M	M	M	-
C1-2	-	H	H	M	L	-
C2	L	M	M	-	L	H
C2-1	L	M	M	-	L	H
C2-2	-	M	M	-	-	H
C3	-	-	M	M	H	-
C4	L	M	H	M	L	H
C4-1	L	M	H	M	M	H
C4-2	-	M	H	M	L	H
C5	L	M	M	M	L	H
C6	L	M	H	M	M	-
C7	-	M	H	H	L	-
C9	-	L	H	M	-	-

Probability: H (high), M (medium), L (low), - (undetected)



large variations in transcriptional data and only detected $3\,574 \pm 2\,010$ genes per neuron (Supplementary information, Figure S9A and S9B). A large number of genes, including neuron type-specific markers, were not detected by the low-coverage single-cell RNA-seq. For example, some marker genes of C2 neurons, including *Sst*, *Nppb* and *Il31ra*, were not detected in all neurons of the corresponding NP3 population reported by Usoskin *et al* [34], and the representative genes *Bmp8a* and *Gpr139*, albeit expressed at a low level in C2 neurons, were absent from the gene list of NP3 (Supplementary information, Figure S9C). Therefore, the low-coverage RNA-seq is not sufficient to identify all representative genes for DRG neuron types, particularly those expressed at low levels, and cannot capture the entire transcriptome of a neuron. In contrast, the transcriptome of individual neurons can be comparatively well characterized by high-coverage single-cell RNA-seq in combination with the methods used to evaluate the quality of datasets. In addition, our data can be used to profile non-coding RNAs and to compare alternative splicing and RNA editing across different types of DRG neurons. Thus, the present study, for the first time, provides the classification of neuron types by comprehensive transcriptome analysis.

DRG neuron types under physiological condition

Neuron types have often been defined by their morphological and/or functional characteristics, with or without the use of molecular markers. Recently, single-cell RNA-seq has been used to identify neuron types in an unbiased manner. We propose that transcriptomic, morphological and functional characteristics are the three major components of a defined neuron type. A specific type of neuron has certain morphological properties, a unique transcriptome and transcriptome-derived signaling networks and markers, and a functional phenotype. Our single-cell RNAseq-based clustering analysis reveals more subtypes of DRG neurons than previously expected. We suggest a new catalog of DRG neuron types, including 6 types and 10 subtypes of small neurons, and 4 types and 4 subtypes of large neurons with single or combined markers. In addition to the well-known neuron types (C1, C3, C4, C5, C9 and C10), our present data indicate that *Nppb*-marked C2 is an independent neuron

type, consistent with other recent proposals [25, 34]. Neuron types C6, C7 and C8, and all corresponding subtypes have not been previously reported.

Our study suggests a new architecture of the classification of DRG neurons. The “NP subset” and “PEP subset” are no longer suitable for defining neuron types because transcriptome analysis clearly shows that IB4-positive (NP) neurons express neuropeptides at various levels, and expression of neuropeptides such as substance P and CGRP is not limited to C1 (PEP) neurons. The transcriptional profiles of TRPV1-lineage DRG neurons and Na_v1.8 channel-expressing nociceptors [23, 24] could be shared by several neuron types. Interestingly, some C1 neurons express high levels of both *Trpa1* and *Trpv1*, whereas C4, C5 and C6 express relatively lower levels of *Trpa1*. Previous ISH results using digoxigenin (DIG)- or fluorescein-labeled probes showed the co-existence of *Trpa1* and *Trpv1* in ~3.6% DRG neurons [39], and ISH using sensitive radioisotope-labeled probes showed *Trpa1* expression in ~32.4% DRG neurons [40]. Thus, to fully characterize these neuron types, it is worth to integrate multiple sources of information including molecular characteristics, physiological functions, and sensory circuits.

The neuron types described here represent the physiological, but not the pathological, states of DRG neurons. The gene expression patterns of DRG neurons are markedly changed in animal models of chronic pain induced by peripheral nerve injury or inflammation [17]. For example, the number of *Gal*-expressing DRG neurons is increased after peripheral nerve injury, whereas the expression of *Tac1* is reduced. Therefore, neuron types and their markers should be studied independently under pathological conditions in various animal models of disease. Moreover, in addition to the regulation of gene expression, the mechanisms of chronic pain could also relate to the translational and post-translational regulation of proteins.

Functional annotations of neuron types and related gene-networks

The present study shows that transcriptome data can partially predict functions of neuron types. The transcriptomic heterogeneity among DRG neuron types correlates

Figure 6 Neuron types and correlated somatosensory receptors. **(A)** The morphological characteristics of DRG neurons. **(B)** Classification of the types and subtypes of DRG neurons, their markers and the type hierarchy of DRG neurons. New types, subtypes and markers are indicated in red. **(C)** Traditionally classified subsets (upper block) are not suitable for the classification of neuron types in DRG (lower block). **(D)** The diagram shows the functional probability of the neuron types as somatosensory receptors. **(E)** A schematic showing the proposed framework of DRG neuron types and their proportions based on both transcriptomic analysis and ISH, as well as functional annotations suggested by electrophysiological analysis results and published data. C8 and C10 are predicted to be mechanoreceptors or/and proprioceptors.

with the somatosensory diversity. Identification of six types of MHN suggests that the sensing of mechanoheat stimuli is a basic function for many nociceptors. These MHNs could be further differentiated by the specific signaling network to sense other stimuli. C1-2 neurons act as major MHNs, consistent with previous reports that Gal is involved in mechanical and thermal nociception [41]. The mechanoreception of C1-1 may be attributed to the expression of TRPC1, KCNK4 and ASIC3, which mediates both cutaneous and muscle nociception induced by mechanical and acid stimuli [14]. C6 could be MHNs expressing *Trpv2*, *Trpm8*, *Trpm3*, *Trpc1* and *Asic3*, and respond to a broad range of temperature and mechanical stimuli. C7 as an MHN (MS) expresses not only heat sensor ANO1 [42] and touch-related *Trpc1* but also the genes typically expressed in large neurons, such as *Cpne6* and *Pvalb*.

C2, particularly C2-1, are MHNs (MI, IS) that express itch-related NPB [43], IL31RA [36] and proteinase-activated receptor 2 (PAR2) encoded by *F2r11* [44]. The thermal sensitivity of C2 correlates with the presence of *Trpv1* and *Trpv2* in the *Nppb* gene network and the role of NBP in thermal hyperalgesia [19, 45]. Both C4 and C5 are MHNs (IS). MRGPR A3 in C4, MRGPR D in C5 and C6, and MRGPR C11 in C2 and C4 are major itch receptors (Figure 5A). The *Mrgprb4* gene network in C4-2 could be responsible for detecting gentle touch and massage-like stroking [46, 47]. TRPV2, TRPC3, TRPC6 and Piezo2 in C5 may correlate to C4-2's moderate sensitivity to heat and mechanical stimuli and its contribution to nociceptive behavior [47, 48].

Our analysis supports the notion that the TH-positive DRG neuron functions as C-LTMR [2, 49]. C3 expresses *Th*, *Fam19a4*, *Cpne6* and *Piezo2* [50]. Our finding indicates that C9 neurons are A β -MNs, consistent with the early idea regarding A β -fiber nociceptive DRG neurons [51]. C8 neurons differentially express *Trpc1*, *Kcnk4*, *Asic3* and *Piezo2*, suggesting their roles in mechanoreception. They were not recorded by our *in vivo* electrophysiological approach, suggesting that either their receptive fields are not in the plantar skin or they possess other modalities that were not tested, such as other touch modalities and limb proprioception. C3 and C8 are two of the most unique populations of neurons that may serve as mechanoreceptors. It might be possible to identify novel mechanoreceptive molecules in these neuron types.

Taken together, our findings indicate that most types of nociceptors respond to multiple stimulus modalities and others have more specialized response properties. Sensing of heat and noxious mechanical stimuli may be a principal function for many types and subtypes of nociceptive small DRG neurons that could also be clas-

sified by other somatosensory properties such as itch. Two types of nociceptors more specialized for noxious mechanical stimulus exist in large DRG neurons. Most MHNs are also polymodal nociceptors because they are also sensitive to various chemical stimuli. Thus, our study provides new insights into the biological function of somatosensory neurons, particularly for the nociceptors, and suggests that neuron types are defined by their transcriptomic, morphological and functional characteristics. The transcriptome data of DRG neuron types can be a resource for studying sensory mechanisms and pain therapy.

Materials and Methods

Preparation of single DRG neurons

Lumbar DRGs dissected from adult mice (C57BL/6, male, 8-10 weeks) were digested with 1 mg/ml collagenase type 1A, 0.4 mg/ml trypsin type I and 0.1 mg/ml DNase I (Sigma) in DMEM at 37 °C for 25-30 min, triturated and plated on coverslips. Neurons were treated with fluorescein-labeled IB4 (1:1 000; Vector Lab) for 10 min at room temperature. Cell size was recorded according to the scaleplate of an Olympus microscope (IX72). Single DRG neurons with or without IB4 labeling were randomly selected and aspirated into a glass pipette for single-cell RNA extraction. To minimize the changes in gene expression and meet the quality requirement for cDNA used to construct sequencing libraries, all neurons were collected within 4 h of neuron dissociation.

One technical concern was the number of sequenced neurons required to cover all neuron types in DRG. This was determined by subsequent clustering analysis. In fact, most neuron clusters and gene modules for DRG neurons appeared after ~100 neurons were analyzed; the clusters and subclusters became more apparent when the number of neurons analyzed was increased.

RNA isolation and library construction

The selected neurons were processed for single-cell RNA extraction and reverse transcription within 1 h and were subjected to cDNA amplification and purification. Single DRG neurons were gently transferred into lysis buffer, and reverse transcription was performed using a SMARTer Ultra Low RNA Kit (Clontech) directly in the cell lysate. Single-cell cDNA was amplified using an Advantage 2 PCR Kit (Clontech) according to the manufacturer's protocol. After cDNA amplification, 50 μ l of the sample was subjected to cDNA purification, whereas the remaining sample was kept for single-cell PCR experiments. To construct the sequencing library, we used over 5 ng of the purified cDNA with fragments over 600-bp length from each neuron. Purified cDNA from a single cell was sheared into 150-350-bp fragments by Covaris S2. A TrueSeq DNA library preparation kit (Illumina) was used to prepare the sequencing library according to the manufacturer's protocol.

Sequence alignment and analysis

cDNA library was sequenced using the Illumina sequencing platform (HiSeq 2000). The size of the library was ~300 bp, and both ends of the library were sequenced to a length of 100 bp. The

raw reads were cleaned by removing adaptor sequences, short sequences (length < 35 bp), low-quality bases (quality < 20) and ambiguous sequences (i.e., reads with more than two unknown bases 'N').

Genome mapping is a key step in the analysis. We used tophat v2.0.9 to map the cleaned RNA-seq reads to the mouse mm10 genome with two mismatches, two gaps and one multihit allowed. We utilized the tophat2 splicing algorithm to map reads covering splice junctions, thereby improving the utilization of reads. After genome mapping, cufflinks v2.1.0 was used to quantify gene expression. The gene expression value was normalized by FPKM and adjusted by a geometric algorithm.

For the saturation plot, BAM files were generated from genome alignment and genome annotation (gtf/gff). The gradient reads, such as 1, 3, 5, 10, 20 and 30 million reads, were processed to establish the relationship between mapped reads and mm10 genes. Next, relation between the reads and number of genes in each read gradient was determined. The result is shown as the saturation curve.

WGCNA was performed using custom R scripts to identify gene co-expression modules in the transcriptome values of differentially expressed genes. In brief, a correlation matrix among all pairs of genes was created across the measured samples. Next, an adjacency matrix was constructed by raising the co-expression measure (power = 4), which was a soft-threshold of the correlation matrix. The topological overlap measure (TOM) was calculated based on the adjacency matrix. Then, the genes were hierarchically clustered using 1-TOM as the distance measure, and the modules were determined for the resulting dendrogram.

The ME is the first principal component in a given module. The module membership (k_{ME}) of each gene in a module refers to the correlation of the gene expression level between the gene and the ME. The genes with the highest value of module membership were considered to be intramodular hub genes. The gene connection network was visualized using the bioinformatics software Cytoscape (v3.1.1) for complex networks.

Custom R scripts were used to perform hierarchical clustering. To identify the differentially expressed genes in each neuron cluster, the transcriptomic data of all qualified neurons were analyzed using unsupervised hierarchical clustering and Fisher's exact test.

Single-cell PCR

Sequences of the primers used in PCR are provided in Supplementary information, Table S3. For cDNA amplification from a single DRG neuron, 50 μ l of the sample was subjected to cDNA purification, and the remaining sample was kept for single-cell PCR experiments. Quantitative PCR was performed using Premix Ex Taq (Takara) on a 7500 Fast Real-time PCR system (Applied Biosystems) according to the manufacturer's protocol. The data were analyzed using 7500 System SDS Software 2.3 (Applied Biosystems) by the standard curve method. The mRNA level was normalized to GAPDH mRNA. The data are presented as the mean \pm SEM. The PCR products were analyzed on ethidium bromide-stained 1.5% agarose gels.

For single-cell PCR, individual cells were placed gently into a reaction tube containing reverse transcription reagents and processed according to the manufacturer's protocol (Superscript III, Invitrogen). Subsequently, cDNA was used in separate PCR reactions. The reaction mixture (20 μ l) was amplified using Ex Taq (Takara) for 45 cycles. *Gapdh* was used as a positive control.

A negative control was obtained from clean pipettes containing bath solution. The PCR products were analyzed on ethidium bromide-stained 1.5% agarose gels.

In situ hybridization

Lumbar 4 and 5 DRG were dissected from adult male mice and processed for ISH. Probes for *Gal*, *Nppb*, *Il31ra*, *Th*, *Mrgpa3* and *Mrgprd* mRNA, among other genes, are listed in Supplementary information, Table S4. For ISH, cRNA riboprobes were labeled with DIG. Fresh sections of DRG were fixed with 4% paraformaldehyde in DEPC-PBS for 20 min, then acetylated and prehybridized in hybridization buffer for 3 h at 67 $^{\circ}$ C. The sections were then incubated with the hybridization buffer containing 1 μ g/ml of the antisense probe for 16 h at 67 $^{\circ}$ C. After hybridization, the sections were incubated in alkaline phosphatase-conjugated sheep anti-DIG antibodies (1:2 000; Roche Molecular Biochemicals). The slides were developed in a solution of 1 μ l/ml NBT and 3.5 μ l/ml BCIP in alkaline phosphatase buffer.

For double fluorescent ISH, one probe was labeled with fluorescein (FITC), whereas the other was labeled with DIG. Both probes were added to the hybridization buffer. After washing and blocking, the sections were incubated with anti-FITC-HRP (1:4 000) overnight at 4 $^{\circ}$ C. After washing with TNT buffer, the sections were incubated with TSA-Plus (DNP, 1:100) for 10 min. Then, the sections were washed with TNT and incubated with anti-DIG-AP (1:1 000) and anti-DNP Alexa488 (1:500) in 1% blocking reagent overnight at 4 $^{\circ}$ C. The sections were washed with TNT and TS8.0, and then were developed with HNPP/FR (1:1:100) in TS8.0.

Immunohistochemistry

Adult male mice were fixed with 4% paraformaldehyde and 1% picric acid. To label IB4-positive neurons, cryostat sections of DRG were first incubated with 5 μ g/ml IB4 (Vector Laboratories). For triple immunofluorescence, the sections were stained with goat antibodies against GSL (1:1 000; Vector Lab), mouse antibodies against NF200 (1:2 000; Sigma) and rabbit antibodies against CGRP (1:2 000; Dia Sorin), followed by FITC-, Cy3- and Cy5-conjugated donkey secondary antibodies against goat, mouse and rabbit (1:500; Jackson ImmunoResearch). The tissues were examined under a confocal microscope.

In vivo whole-cell patch-clamp recording and neuron classification

Mice were anesthetized with pentobarbital sodium (Nembutal, 50 mg/kg i.p. for initial dose and 20 mg/kg i.p. for supplemental dose). After laminectomy at lumbar levels 2-6, L5 DRG were exposed and isolated from the surrounding tissues. Oxygenated artificial cerebrospinal fluid (ACSF) was dripped periodically onto the surface of the ganglia to prevent drying and hypoxia. Two lumbar vertebrae (lumbar 4 and 6) were clamped. The epineurium covering the DRG was removed under a dissection microscope. The skin was sewn to a ring to maintain a pool of warm oxygenated ACSF. The DRG was continuously superfused at a rate of 3 ml/min by oxygenated ACSF that was preheated by an in-line heater to achieve the desired temperature of 37 $^{\circ}$ C. The ACSF contained (in mM) 130 NaCl, 3.5 KCl, 24 NaHCO₃, 1.25 NaH₂PO₄, 1.2 MgCl₂, 1.2 CaCl₂ and 10 dextrose. The solution was bubbled with 95% O₂ and 5% CO₂, and had a pH of 7.4 and an osmolarity of 290-310 Osm. The vasoconstricting drug (Arg8)-vasopressin (10 μ M in oxygenated ACSF, Sigma Aldrich) was applied via a syringe. The va-

soconstrictor stopped or substantially reduced the blood flow to the capillary bed in the DRG, thereby preventing bleeding that would otherwise occur after the application of collagenase. Collagenase P (1 mg/ml, Roche Diagnostics) was applied to a local region of the DRG via a pipette located close to the surface. The bath perfusion was turned off for ~5 min during the application. Repetitive collagenase application was performed until the neuronal cell bodies were loosened from the surrounding connective tissue and cells (4-6 applications for patch clamp recording).

Next, we performed *in vivo* whole-cell patch clamp recording of randomly selected DRG neurons. The patch pipette (resistance 5-8 M Ω) was filled with an internal solution containing (in mM) 135 K-gluconate, 0.5 CaCl₂, 2 MgCl₂, 5 KCl, 5 EGTA, 5 HEPES and 5 D-glucose. Somatic stimuli were applied to the plantar skin of the hindpaw. A heat stimulus (30-52 °C; 20 s/stimulus) was applied via a thermal stimulator at an inter-stimulus interval of 90 s. Cold stimulus was applied with ice water (20 s per stimulus). An innocuous brush on the skin was performed using a cotton swab. Innocuous pressure was induced by picking up a fold of skin with a flattened alligator clip to produce a consistent strength (~6.5 mN/mm²). Noxious mild pinching at the hindpaw was performed using calibrated serrated forceps (~6 N/mm²). Mechanical stimuli (brush, pressure and pinch) were applied (2 s per stimulus) in ascending order with an inter-stimulus interval of 30 s. Noxious stimulation was applied sparingly to avoid neuronal sensitization.

Whole-cell recordings were performed using an Axopatch 200B amplifier and were digitized with a Digidata 1440A interface and controlled by pCLAMP10.1 software. The whole-cell configuration was obtained in voltage-clamp mode. In brief, we filled a glass electrode with intracellular solution and advanced it quickly to the target depth under moderate positive pressure. When the tip of the electrode was close to a neuron, the tip resistance was increased. We removed the positive pressure and applied negative pressure to aid in seal formation. We achieved the whole-cell configuration using brief bursts of negative pressure to rupture the cell membrane. For detection of responses induced by stimuli, the DRG neuron was subjected to whole-cell recording in voltage-clamp mode at -60 mV. The peripheral stimulus caused a large inward Na⁺ current and a small outward K⁺ current, and generated an action potential in responsive neurons.

After recording, the internal solution was collected in a PCR tube containing 2 μ l polymerase buffer, 0.5 μ l oligo dT primer, 0.5 μ l Quantiscript RT enzyme mix (Qiagen) and incubated for 1 h at 42 °C. The cDNA product was treated with 4 μ l ligase buffer and 1 μ l ligase mix for 30 min at 24 °C, and then amplified with 14.5 μ l REPLI-g sc reaction buffer and 0.5 μ l REPLI-g SensiPhi DNA polymerase for 2 h at 30 °C according to the manufacturer's instructions (Qiagen). After amplification, the reaction solution was diluted to 300 μ l. Then, we performed quantitative real-time PCR to evaluate the gene expression level of markers, including *Gal*, *Zcchc12*, *Cldn9*, *Nppb*, *Sst*, *Nts*, *Th*, *Mrgpra3*, *Mrgprb4*, *Mrgprd*, *Nxph1*, *Baiap211*, *Pvalb* and *Nefh*. Recorded neurons were sorted into corresponding clusters when the expression level of the marker genes for one cluster was at least fivefold higher than that of other clusters. Recorded neurons were sorted into C6 when several marker genes were detected simultaneously.

Data availability

The gene data sets are deposited in the public resource Gene Expression Omnibus with accession number of GSE63576.

Acknowledgments

We thank Drs Li Yang and Luo-Nan Chen for comments on data analysis and manuscript. This work was supported by the National Natural Science Foundation of China (31130066 and 81300961) and the Strategic Priority Research Program (B) of Chinese Academy of Sciences (XDB01020300).

Author Contributions

XZ, H-S X, LB, C-L L and K-C L conceived and designed the study. Y-J L and Y-Q Z collected neurons. B-B Z constructed the library. C-L L, M-M S and RH analyzed RNA-seq data. DW, C-L L, HL, J-R Z, Y-J L and X-Y H performed ISH. C-L L, DW, S-S W and X-Y H performed single-cell real-time PCR. YC performed single cell PCR. DW performed IHC. K-C L recorded neurons. C-L L treated recorded neurons for Q-PCR. XZ, C-L L and LB wrote the manuscript.

Competing Financial Interests

The authors declare no conflict of interest.

References

- Maxwell DJ, Réthelyi M. Ultrastructure and synaptic connections of cutaneous afferent fibres in the spinal cord. *Trends Neurosci* 1987; **10**:117-123.
- Li L, Rutlin M, Abaira VE, *et al.* The functional organization of cutaneous low-threshold mechanosensory neurons. *Cell* 2011; **147**:1615-1627.
- Ju G, Hökfelt T, Brodin E, *et al.* Primary sensory neurons of the rat showing calcitonin gene-related peptide immunoreactivity and their relation to substance P-, somatostatin-, galanin-, vasoactive intestinal polypeptide- and cholecystokinin-immunoreactive ganglion cells. *Cell Tissue Res* 1987; **247**:417-431.
- Wang H, Rivero-Melian C, Robertson B, Grant G. Transganglionic transport and binding of the isolectin B4 from *Griffonia simplicifolia* I in rat primary sensory neurons. *Neuroscience* 1994; **62**:539-551.
- Meyer RA, Ringkamp M, Campbell JN, Raja SN. Peripheral mechanisms of cutaneous nociception. In: McMahon SB, Koltzenburg M. eds. *Wall and Melzack's Textbook of Pain*. Elsevier 2006:3-34.
- Willis WD, Coggeshall RE. Sensory receptors and peripheral nerves. In: Willis WD, Coggeshall RE. eds. *Sensory Mechanisms of the Spinal Cord*. New York, Boston, USA; Dordrecht, London, UK; Moscow, Russia: Kluwer Academic/Plenum Publishers 2004:19-87.
- Zhang X, Shi T, Holmberg K, *et al.* Expression and regulation of the neuropeptide Y Y2 receptor in sensory and autonomic ganglia. *Proc Natl Acad Sci USA* 1997; **94**:729-734.
- Dong X, Han S, Zylka MJ, Simon MI, Anderson DJ. A diverse family of GPCRs expressed in specific subsets of nociceptive sensory neurons. *Cell* 2001; **106**:619-632.
- Liu Y, Yang FC, Okuda T, *et al.* Mechanisms of compartmentalized expression of Mrg class G-protein-coupled sensory receptors. *J Neurosci* 2008; **28**:125-132.
- Baker MD, Wood JN. Involvement of Na⁺ channels in pain pathways. *Trends Pharmacol Sci* 2001; **22**:27-31.

- 11 Caterina MJ, Schumacher MA, Tominaga M, Rosen TA, Levine JD, Julius D. The capsaicin receptor: a heat-activated ion channel in the pain pathway. *Nature* 1997; **389**:816-824.
- 12 Dhaka A, Viswanath V, Patapoutian A. Trp ion channels and temperature sensation. *Annu Rev Neurosci* 2006; **29**:135-161.
- 13 Kobayashi K, Yamanaka H, Noguchi K. Expression of ATP receptors in the rat dorsal root ganglion and spinal cord. *Anat Sci Int* 2013; **88**:10-16.
- 14 Li WG, Xu TL. ASIC3 channels in multimodal sensory perception. *ACS Chem Neurosci* 2011; **2**:26-37.
- 15 Geffeney SL, Goodman MB. How we feel: ion channel partnerships that detect mechanical inputs and give rise to touch and pain perception. *Neuron* 2012; **74**:609-619.
- 16 Kashiba H, Uchida Y, Senba E. Distribution and colocalization of NGF and GDNF family ligand receptor mRNAs in dorsal root and nodose ganglion neurons of adult rats. *Brain Res Mol Brain Res* 2003; **110**:52-62.
- 17 Xiao HS, Huang QH, Zhang FX, *et al.* Identification of gene expression profile of dorsal root ganglion in the rat peripheral axotomy model of neuropathic pain. *Proc Natl Acad Sci USA* 2002; **99**:8360-8365.
- 18 Mantioti S, Lehmann R, Flegel C, *et al.* Comprehensive RNA-Seq expression analysis of sensory ganglia with a focus on ion channels and GPCRs in Trigeminal ganglia. *PLoS One* 2013; **8**:e79523.
- 19 Zhang FX, Liu XJ, Gong LQ, *et al.* Inhibition of inflammatory pain by activating B-type natriuretic peptide signal pathway in nociceptive sensory neurons. *J Neurosci* 2010; **30**:10927-10938.
- 20 Li KC, Zhang FX, Li CL, *et al.* Follistatin-like 1 suppresses sensory afferent transmission by activating Na⁺,K⁺-ATPase. *Neuron* 2011; **69**:974-987.
- 21 Wang F, Cai B, Li KC, *et al.* FXYD2, a gamma subunit of Na⁺, K⁺-ATPase, maintains persistent mechanical allodynia induced by inflammation. *Cell Res* 2015; **25**:318-334.
- 22 Venteo S, Bourane S, Mechaly I, *et al.* Regulation of the Na,K-ATPase gamma-subunit FXYD2 by Runx1 and Ret signaling in normal and injured non-peptidergic nociceptive sensory neurons. *PLoS One* 2012; **7**:e29852.
- 23 Goswami SC, Mishra SK, Maric D, *et al.* Molecular signatures of mouse TRPV1-lineage neurons revealed by RNA-Seq transcriptome analysis. *J Pain* 2014; **15**:1338-1359.
- 24 Thakur M, Crow M, Richards N, *et al.* Defining the nociceptor transcriptome. *Front Mol Neurosci* 2014; **7**:87.
- 25 Chiu IM, Barrett LB, Williams EK, *et al.* Transcriptional profiling at whole population and single cell levels reveals somatosensory neuron molecular diversity. *Elife* 2014; **3**.
- 26 Ramskold D, Luo S, Wang YC, *et al.* Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol* 2012; **30**:777-782.
- 27 Wills QF, Livak KJ, Tipping AJ, *et al.* Single-cell gene expression analysis reveals genetic associations masked in whole-tissue experiments. *Nat Biotechnol* 2013; **31**:748-752.
- 28 Jaitin DA, Kenigsberg E, Keren-Shaul H, *et al.* Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* 2014; **343**:776-779.
- 29 Patel AP, Tirosh I, Trombetta JJ, *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 2014; **344**:1396-1401.
- 30 Pollen AA, Nowakowski TJ, Shuga J, *et al.* Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat Biotechnol* 2014; **32**:1053-1058.
- 31 Shalek AK, Satija R, Shuga J, *et al.* Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* 2014; **510**:363-369.
- 32 Treutlein B, Brownfield DG, Wu AR, *et al.* Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* 2014; **509**:371-375.
- 33 Wu AR, Neff NF, Kalisky T, *et al.* Quantitative assessment of single-cell RNA-sequencing methods. *Nat Methods* 2014; **11**:41-46.
- 34 Usoskin D, Furlan A, Islam S, *et al.* Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. *Nat Neurosci* 2015; **18**:145-153.
- 35 Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008; **9**:559.
- 36 Cevikbas F, Wang X, Akiyama T, *et al.* A sensory neuron-expressed IL-31 receptor mediates T helper cell-dependent itch: involvement of TRPV1 and TRPA1. *J Allergy Clin Immunol* 2014; **133**:448-460.
- 37 Diveu C, Lak-Hal AH, Froger J, *et al.* Predominant expression of the long isoform of GP130-like (GPL) receptor is required for interleukin-31 signaling. *Eur Cytokine Netw* 2004; **15**:291-302.
- 38 Celio MR. Calbindin D-28k and parvalbumin in the rat nervous system. *Neuroscience* 1990; **35**:375-475.
- 39 Story GM, Peier AM, Reeve AJ, *et al.* ANKTM1, a TRP-like channel expressed in nociceptive neurons, is activated by cold temperatures. *Cell* 2003; **112**:819-829.
- 40 Obata K, Katsura H, Mizushima T, *et al.* TRPA1 induced in sensory neurons contributes to cold hyperalgesia after inflammation and nerve injury. *J Clin Invest* 2005; **115**:2393-2401.
- 41 Kerr BJ, Cafferty WB, Gupta YK, *et al.* Galanin knockout mice reveal nociceptive deficits following peripheral nerve injury. *Eur J Neurosci* 2000; **12**:793-802.
- 42 Cho H, Yang YD, Lee J, *et al.* The calcium-activated chloride channel anoctamin 1 acts as a heat sensor in nociceptive neurons. *Nat Neurosci* 2012; **15**:1015-1021.
- 43 Mishra SK, Hoon MA. The cells and circuitry for itch responses in mice. *Science* 2013; **340**:968-971.
- 44 Han L, Dong X. Itch mechanisms and circuits. *Annu Rev Biophys* 2014; **43**:331-355.
- 45 Liu XY, Wan L, Huo FQ, *et al.* B-type natriuretic peptide is neither itch-specific nor functions upstream of the GRP-GR-PR signaling pathway. *Mol Pain* 2014; **10**:4.
- 46 Liu Q, Vrontou S, Rice FL, Zylka MJ, Dong X, Anderson DJ. Molecular genetic visualization of a rare subset of unmyelinated sensory neurons that may detect gentle touch. *Nat Neurosci* 2007; **10**:946-948.
- 47 Vrontou S, Wong AM, Rau KK, Koerber HR, Anderson DJ. Genetic identification of C fibres that detect massage-like stroking of hairy skin *in vivo*. *Nature* 2013; **493**:669-673.
- 48 Rau KK, McIlwrath SL, Wang H, *et al.* Mrgprd enhances excitability in specific populations of cutaneous murine polymodal nociceptors. *J Neurosci* 2009; **29**:8612-8619.
- 49 Delfini MC, Mantilleri A, Gaillard S, *et al.* TAF4A, a chemo-

- kine-like protein, modulates injury-induced mechanical and chemical pain hypersensitivity in mice. *Cell Rep* 2013; **5**:378-388.
- 50 Ranade SS, Woo SH, Dubin AE, *et al.* Piezo2 is the major transducer of mechanical forces for touch sensation in mice. *Nature* 2014; **516**:121-125.
- 51 Djouhri L, Lawson SN. Abeta-fiber nociceptive primary afferent neurons: a review of incidence and properties in relation to other afferent A-fiber neurons in mammals. *Brain Res Brain Res Rev* 2004; **46**:131-145.

(**Supplementary information** is linked to the online version of the paper on the *Cell Research* website.)