# Comparative transcriptome profiling of the human and mouse dorsal root ganglia: an RNA-seq–based resource for pain and sensory neuroscience research

Pradipta Ray[a,b], Andrew Torck[a], Lilyana Quigley[a], Andi Wangzhou[a], Matthew Neiman[a], Chandranshu Rao[a], Tiffany Lam[a], Ji-Young Kim[a], Tae Hoon Kim[b], Michael Q. Zhang[b], Gregory Dussor[a], Theodore J. Price[a,*]

## Abstract

Molecular neurobiological insight into human nervous tissues is needed to generate next-generation therapeutics for neurological disorders such as chronic pain. We obtained human dorsal root ganglia (hDRG) samples from organ donors and performed RNA-sequencing (RNA-seq) to study the hDRG transcriptional landscape, systematically comparing it with publicly available data from a variety of human and orthologous mouse tissues, including mouse DRG (mDRG). We characterized the hDRG transcriptional profile in terms of tissue-restricted gene coexpression patterns and putative transcriptional regulators, and formulated an information-theoretic framework to quantify DRG enrichment. Relevant gene families and pathways were also analyzed, including transcription factors, G-protein-coupled receptors, and ion channels. Our analyses reveal an hDRG-enriched protein-coding gene set (~140), some of which have not been described in the context of DRG or pain signaling. Most of these show conserved enrichment in mDRG and were mined for known drug–gene product interactions. Conserved enrichment of the vast majority of transcription factors suggests that the mDRG is a faithful model system for studying hDRG, because of evolutionarily conserved regulatory programs. Comparison of hDRG and tibial nerve transcriptomes suggests trafficking of neuronal mRNA to axons in adult hDRG, and are consistent with studies of axonal transport in rodent sensory neurons. We present our work as an online, searchable repository (https://www.utdallas.edu/bbs/painneurosciencelab/sensoryomics/drgtxome), creating a valuable resource for the community. Our analyses provide insight into DRG biology for guiding development of novel therapeutics and a blueprint for cross-species transcriptomic analyses.

**Keywords:** Dorsal root ganglion, Comparative transcriptomics, Target identification, Human sensory neuroscience

## 1. Introduction

The dorsal root ganglia (DRG) are a primary sensory tissue in vertebrate nervous systems, delivering sensory signals from the body to the central nervous system (CNS) through pseudouni-polar neurons. The DRG is composed of several specialized cell types including proprioceptive, low-threshold and damage-sensing nociceptive sensory neurons (nociceptors), as well as Schwann cells, fibroblasts, and satellite glial cells (SGCs). With

the advent of high-throughput RNA-seq, there has been a concerted effort on generating whole transcriptome snapshots of DRG and other tissues with recent emphasis on single-cell RNA-seq (scRNA-seq) studies. Thus far, efforts have largely been directed toward mouse DRG (mDRG)[59] and rat DRG,[51] with some focusing on single neuron[15,56,112] or specific neuronal subpopulations[31,41,108] highlighting nociceptors. Such studies are informative because they focus on profiling a key neuronal subpopulation in the context of detecting nociceptive signals and generation of pain, enhancing our understanding of the molecular biology and diversity of these neurons. The human DRG (hDRG) in general, and nociceptors in particular, are widely viewed as tissue and cell types with possibilities for discovering biological targets for analgesic drugs, leading to development of therapeutics for both acute and chronic pain[71,98] that do not involve the CNS and avoid resulting complications.[4,20] Studies identifying transcriptional[21,34,120,123] (including scRNA-seq efforts[39]) and proteomic[95] changes in preclinical pain models have been undertaken, but do not account for evolutionary differences in DRG transcriptomes between rodent models and humans. Such preclinical models can potentially suffer from translational issues, as they move toward the clinic.[100] Studies characterizing transcriptome profiles of human nociceptors and DRG (or related tissues and cell types in vivo or in vitro) present opportunities for advancing understanding of basic pain mechanisms and therapeutic targets but a limited number of such in vitro[116] or in

vivo[26,97,104,116] studies have been performed. The SymAtlas[104] project provided the first in vivo, high-throughput (but incomplete) characterization of the hDRG transcriptome using microarrays. Recent in vivo studies[26,97] perform RNA-seq, but have mainly focused on specific gene families or diseases, without systematic comparison with preclinical models.

We performed RNA-seq on female, human organ donor L2 DRG tissues (publicly available at dbGAP phs001158). Comprehensive analyses of the transcriptional landscape with respect to other human tissues identified gene coexpression modules and hDRG-enriched genes in pertinent gene families and signaling pathways. As a starting point for therapeutic target identification, we identified the set of genes with conserved DRG enrichment in human and mouse, and mined drug databases to catalog interactions of known drugs with products of these genes. This effort is the first to present whole transcriptome gene abundances for hDRG and mDRG, contrasted with relevant human and mouse tissues (online at https://www.utdallas.edu/bbs/pain-neurosciencelab/sensoryomics/drgtxome).

## 2. Methods

We performed RNA-seq for hDRG and integratively analyzed the transcriptome by comparing it with publicly available human and mouse RNA-seq data sets (sources shown in **Table 1**, and the workflow for the project shown in Supplementary Fig. 1, available online at http://links.lww.com/PAIN/A559). We also used publicly available human microarray data to contrast whole human DRG and trigeminal ganglia (TG) (GEO PRJNA87249),[104] along with 2 data sets representative of DRG component cell types: cultured primary normal human Schwann cells (NHSCs) from peripheral nerves (GEO GSE14038)[67] and cultured fibroblasts (FIBRO) from the skin (GEO data set GSE21899). We were unable to find a data set for purified SGCs in any database. We further integrated published quantification from mDRG scRNA-seq analyses[112] (male and female lumbar DRG neurons, GEO GSE59739) to putatively identify cell-type–specific expression in DRG-enriched genes. For ease of discourse, human gene names in upper case are used in the text to refer to both human genes and its orthologs, with the species being apparent from context. Mouse gene names and human protein names are capitalized, with human protein names being additionally italicized, when used in the article. At the outset, it is important to clarify that our analyzed tissue panel only queries a finite set of adult, healthy tissue; and the notions of "DRG enriched" and "DRG specific" gene expression is with respect to this set of reference tissues that provide a relevant test-bench for asking questions about mammalian DRG.

### 2.1. Human dorsal root ganglia preparation and RNA-seq

Tissue was sourced from Anabios, Inc (San Diego, CA). For this study, L2 lumbar DRG were removed from 3 female, middle-aged (30-60 years), Caucasian, consented organ donors with no history of chronic pain, before cross-clamp and stored immediately in RNALater (Ambion, Austin, TX). RNA-seq was performed on the whole tissue by Active Motif, Inc (San Diego, CA). An RNEasy Qiagen kit was used for total RNA isolation, and DNA was removed using DNase-I digestion. The extracted total RNA was then used with an Illumina Truseq RNA sample preparation v2 kit to generate polyA+ RNA libraries for sequencing. Over 50 million paired-end 75 bp reads were sequenced from each sample on the Illumina platform, with mean sequencing quality (Phred) scores per base between 29.4 and 34.24 (Supplementary Table 1, available online at http://links.lww.com/PAIN/A559). The RNA-seq data sets generated were contrasted with publicly available RNA-seq data sets in other relevant

human tissues and their orthologous tissues in the house mouse (*Mus musculus*). For mDRG RNA-seq data sets, we used publicly available RNA-seq data from female mice of the C57BL/6 mouse strain, the inbred model strain used to sequence the reference mouse genome.[14] The publicly available mDRG RNA-seq data sets (SRR869619-21) we used for comparison were sourced from individual female mice and homogenized in Trizol, with total RNA treated with 2 rounds of polyA selection (Dynabeads mRNA purification kit; Invitrogen, Carlsbad, CA) followed by fragmentation (Fragmentation Reagents; Applied Biosystems, Foster City, CA). Although not identical to the library preparation for hDRG RNA-seq, this data set was also generated using the Illumina TruSeq kit and yielded 50 bp paired-end reads (of comparable size with the hDRG 75 bp paired-end reads),[29] thus being suitable for comparison. The following subsection details in silico strategies to normalize gene abundance quantifications obtained from these and other data sets we analyzed to minimize the effect of differential library preparation across samples.

### 2.2. Quantification of gene abundances from RNA-seq experimental data

#### 2.2.1. Mapping of sequenced reads and quantification of gene abundances

Sequenced FASTQ files generated by our RNA-seq experiments or downloaded from public databases were mapped to the human and mouse reference transcriptome for the purposes of quantifying relative gene abundance. Reference human (NCBI hg19) and mouse (NCBI mm10) genomes,[22] and reference human (Gencode v14) and mouse (Gencode vM4) transcriptomes[37] were used for mapping sequenced reads. The Tophat/Cufflinks pipeline was used for analyzing the RNA-seq data sets.[111] The Burroughs–Wheeler transform–based RNA-seq mapping tool Tophat[110] (v2.0.13) was used for mapping the sequencing reads with the following command-line parameters: *tophat2 -o <output_path> -p 8 –transcriptome-index <reference_transcriptome> <reference_genome> <leftreads.fastq> <rightreads.fastq>*. The gene abundance quantification tool Cuffdiff[109] (v2.2.1) was used to estimate relative abundances of genes in the reference transcriptome with the following command-line parameters: *cuffdiff -o <output_file> –library-norm-method classic-fpkm -p 8 <reference_transcriptome>*. Relative abundances were calculated for genes with respect to the reference library, based on Gencode coding genes, lincRNAs, pseudogenes, and other noncoding genes with evidence for polyA site or signal.[124] Absolute abundance quantification of genes or transcripts requires calibrating RNA-seq based on exogenous spike-ins with known copy numbers[82] or other methods is rarely performed in practice, and typically require additional amounts of input material, or attachment of unique molecular barcodes to individual transcripts to correct amplification bias in traditional RNA-seq,[102] making them unfeasible or intractable for our samples with limited, varying amounts of tissue sourced from human donors.

#### 2.2.2. Design decisions for RNA-seq analyses

While mapping RNA-seq data sets using Tophat, the number of mismatches allowed per alignment segment was set to the default (2), with all our analyzed data sets, including the 3 hDRG samples we sequenced, having acceptable mapping rates for postmortem samples (>70%, Supplementary Table 1, available online at http://links.lww.com/PAIN/A559).[93] For Tophat, all data sets were assumed to be generated from

**Details of RNA-seq data sets used in analyses.**

| Tissue | Human data id | Mouse data id |
|---|---|---|
| DRG | SRA: SRR3734492 (F), SRR3734493 (F), and SRR3734494 (F) | SRA: SRR869619 (C57BL/6, F), SRR869620 (C57BL/6, F), and SRR869621 (C57BL/6, F) |
| Spinal cord | SRA: SRR1330498 (F) and SRR1369219 (F) | SRA: SRR789190 (C57BL/6, F) and SRR789191 (C57BL/6, F) |
| Frontal cortex | SRA: SRR1413562 (F) and SRR1432123 (F) | SRA: SRR1646331 (C57BL/6, M) and SRR1646335 (C57BL/6, M) |
| Hippocampus | SRA: SRR1348339 (F) and SRR1348424 (F) | SRA: SRR1646332 (C57BL/6, M) and SRR1646336 (C57BL/6, M) |
| Nucleus accumbens | SRA: SRR598196 (F) and SRR598253 (F) | SRA: SRR647495 (C57BL/6, M), SRR647501 (C57BL/6, M), and SRR647498 (C57BL/6, M) |
| Caudate nucleus (h)/striatum (m) | SRA: SRR598428 (F) and SRR1332856 (F) | SRA: SRR1519958 (CBAxC57BL/6, F), SRR1519959 (CBAxC57BL/6, F), and SRR1519960 (CBAxC57BL/6, F) |
| Skeletal muscle | ENA: ERR579149 (M), ERR579152 (F), and ERR579141 (F) | SRA: SRR545833 (C57BL/6, M) and SRR545834 (C57BL/6, M) |
| Heart | ENCODE: ENCFF001RTS (F) and ENCFF770NYA (F) | ENCODE: ENCFF001IVS (C57BL/6, pooled M + F) and ENCFF001IVU (C57BL/6, pooled M + F) |
| Lung | SRA: SRR820596 (F) and SRR821302 (F) | SRA: SRR453156 (C57BL/6, pooled M + F), SRR453157 (C57BL/6, pooled M + F), SRR453158 (C57BL/6, pooled M + F), and SRR453159 (C57BL/6, pooled M + F) |
| Whole blood | SRA: SRR614503 (F) and SRR607042 (F) | SRA: SRR1774203 (C57BL/6, M), SRR1774204 (C57BL/6, M), SRR1774205 (C57BL/6, M), and SRR1774206 (C57BL/6, M) |
| Small intestine | SRA: SRR1957202 (F) | ENCODE: ENCFF996KVO (C57BL/6, pooled M + F), ENCFF001JLL (C57BL/6, pooled M + F), and ENCFF001JLQ (C57BL/6, pooled M + F) |
| Liver | ENCODE: ENCFF187OKV (M) | ENCODE: ENCFF001JAT (C57BL/6, pooled M + F), ENCFF001JAW (C57BL/6, pooled M + F), ENCFF001RTN (C57BL/6, pooled M + F), and ENCFF001RTD (C57BL/6, pooled M + F) |
| Tibial nerve | SRA: SRR1312622 (F) and SRR1312802 (F) | |
| Olfactory epithelium | | SRA: SRR1646337 (C57BL/6, M) and SRR1646333 (C57BL/6, M) |

Nucleotide databases and corresponding IDs of analyzed data sets. The hDRG samples were sequenced as part of the study and deposited in dbGAP/SRAdb, whereas the rest of the data sets was obtained from the dbGAP, SRAdb, GEO, ENCODE, and ENA (ArrayExpress) databases. Sex of the samples and strains for the mouse data sets is also noted.
DRG, dorsal root ganglia; hDRG, human DRG.

strand-agnostic libraries for the purposes of mapping. Although some analyzed libraries were actually strand specific, such an approach allows us to avoid the inherently different nature of strand specific and nonspecific libraries, and minimizes the chances that downstream differential expression analysis will yield artifacts. Because the reference transcriptomes for human and mouse are well annotated by the GENCODE project, de novo assembly was not performed—instead, all concordantly mapped read pairs (fragments) mapping to each known transcript were used by Cuffdiff to quantify relative abundance. Finally, we performed a second round of in silico library selection limiting analysis to only genes with known or predicted polyA+ transcripts (detailed in Section 2.3 under transcripts per million [TPM] calculations), to minimize the effect of different library selection techniques (rRNA depletion or polyA+ selection) in the different analyzed samples.

### 2.3. Sample statistics

#### 2.3.1. Sample statistics for characterizing gene abundances in individual tissues for each species

The relative abundance of a transcript was calculated based on the number of fragments (paired-end reads) that map to it, by calculating fragments per kilobase per million mapped fragments (FPKM) as follows:

$$\text{FPKM}_{\text{transcript } i} = \frac{\text{Mapped fragment count}_{\text{transcript } i}}{\frac{\text{No of mapped fragments in expt}}{10^6} \times \frac{\text{Effective transcript length in base pairs}}{10^3}}.$$

Fragments per kilobase per million mapped fragments for genes were calculated by summing the FPKMs for individual transcripts corresponding to the gene. Cuffdiff was configured to use all replicates to estimate a mean FPKM for each gene. Our analysis ultimately required us to compare across tissues and species for many different RNA-seq experiments, performed across multiple laboratories using different library preparations and sequencing depth. Hence, we decided to transform the FPKMs to another commonly used measure for quantifying relative abundance: TPM.[83] To calculate TPMs, FPKMs are renormalized with respect to the sum of FPKMs of the reference library of transcripts to generate a relative abundance score scaled to parts per million.

$$TPM_{transcript\ i} = 10^6 \times \frac{FPKM_{transcript\ i}}{\sum_j FPKM_{transcript\ j}}.$$

We only chose to quantify relative abundance with respect to transcripts that were polyA+ because most data sets in our analysis used a polyA selection step. LincRNAs, pseudogenes, and other noncoding genes with evidence of polyA signal or site (based on APASdb[124]) were retained in the reference library, whereas remaining noncoding genes (including ribosomal rRNA genes) were not evaluated. Abundances for genes on chromosome Y were also not evaluated to minimize detection of male sex-linked gene expression patterns. By only quantifying the relative abundance of genes with respect to the set of common genes that are minimally targeted by most RNA-seq protocols, we effectively performed a second round of library selection in silico. Finally, to scale the distribution of TPMs in across tissue and species comparably, the upper quartile (75th percentile) TPM was calculated for each sample and values scaled with respect to it to calculate upper quartile–normalized TPMs (uqTPMs), based on previous approaches in the literature.[30] A complete list of evaluated genes, with relative abundances in TPM and uqTPM across all analyzed tissues, is available at https://www.utdallas.edu/bbs/painneurosciencelab/sensoryomics/drgtxome. Empirical density functions for uqTPMs for different tissues in both human and mouse are presented in Supplementary Figure 2 (available online at http://links.lww.com/PAIN/A559). We used upper quartile normalization to scale TPM measures such that the bimodal distribution of relative abundances was comparable across both tissues and species.

### 2.3.2. Sample statistics for characterizing tissue specificity of genes in individual species

A tissue-specific transcriptomic signature inevitably depends on a set of tissue-restricted genes, potentially contributing to tissue-specific functionality or phenotype. The information-theoretic measure of Shannon entropy[101] has been historically used in various domains as a framework for assessing diversity in a population,[38] including the identification of tissue-restricted genes in transcriptomics.[121] For each gene, the uqTPM levels were first normalized to 1, generating a probability distribution $\vec{X}$ with $x_i$ corresponding to the gene's normalized abundance and $t_i$ corresponding to its uqTPM in the ith tissue. Based on this distribution, Shannon entropy[101] $H(\vec{X})$ was calculated to identify tissue-restricted genes that are expressed in the DRG:

$$x_i = \frac{t_i}{\sum_i [t_i]}; H(\vec{X}) = -\sum_i [x_i \times \log_2(x_i)] \text{ where } \log_2(0) = 0.$$

Because the value of Shannon entropy can vary between 0 and $\log_2(n)$ where $n$ = information length (number of tissues in our case), we use Shannon's normalized measure of entropy which

ranges between 0 (highest tissue specificity) and 1 (lowest tissue specificity): $\bar{H}(\vec{X}) = \frac{1}{\log_2(n)} H(\vec{X})$.

We defined 2 additional formulations derived from this framework to characterize tissue-restricted genes of interest. We profiled nervous system–enriched genes among the set of tissue-restricted genes we characterized in this work because these are potential targets or excluded targets for drugs which do or do not cross the blood–brain barrier (BBB). We calculated a neural proportion score by summing the proportion $x_i$ calculated for Shannon entropy across all neural tissues in our tissue panel. We also calculated a score for DRG enrichment that would take into account the magnitude of relative abundance values across different tissues. For each gene, we also calculated a DRG-enrichment score to quantify enrichment in the DRG with respect to other analyzed tissues, ranging from 0 (undetected in the DRG or uniformly expressed across tissues) to 1 (only expressed in the DRG), defined as $x_{DRG} \times (1 - \bar{H}(\vec{X}))$. Empirical density functions for the neural proportion and DRG-enrichment scores for both human and mouse are presented in Supplementary Figure 3 (available online at http://links.lww.com/PAIN/A559). The degree of neural or DRG enrichment for a gene can thus be estimated by contrasting the enrichment score with the overall distribution of the score across all genes.

### 2.3.3. Sample statistics for characterizing conservation of gene expression patterns in human and mouse

Based on the Homologene database,[28] we identified orthology mappings between human and mouse genes, using the Homologene gene family IDs. Every individual human to mouse gene mapping was categorized as being part of a one-to-one, one-to-many, many-to-one, or many-to-many mapping between gene sets, based on the relative timing estimates for duplication and speciation events. In addition, genes with no orthologs in the other species were categorized as one-to-none, many-to-none, none-to-one, or none-to-many mapping. We additionally curated the Homologene database to correct orthology relationships that were inconsistent with the literature, most notably the relationships in the MRGPR family. Ensembl Compara[114] gene trees were used to guide this process, and the final set of homology mappings used is presented beside abundance data for each gene in the RNA-seq tables on our web site. Correlational measures (Pearson correlation coefficient [PCC[85]]) previously used in the literature[96] were calculated between the gene expression (uqTPM) vectors of human genes and their mouse orthologs across corresponding tissues, for those human and mouse genes that have a one-to-one orthology mapping and show either human or mouse DRG enrichment. We analyzed the empirical distribution of the DRG-enrichment scores to identify hDRG-enriched and mDRG-enriched genes (DRG-enrichment score >0.29 in at least 1 species, corresponding to inclusion of all genes where TPM is highest in DRG across tissues) for all human and mouse genes, and restricted this gene set to genes with high correlation (determined from empirical distribution of the PCC) between human and mouse tissue expression panels.

### 2.4. Similarity analysis of gene and tissue expression profiles to identify gene expression signatures

For microarray comparisons, only microarray data sets from variations of the Affymetrix HG-U133A were used. For each chip, the tool Oligo was used for microarray probe normalization,[10] followed by quantile normalization across chips,[9] available on our web site. Averaging across replicates, a clear unimodal

distribution for probe intensity was identified (Supplementary Fig. 4, available online at http://links.lww.com/PAIN/A559) for classifying probesets as expressed or unexpressed. Because culturing cells can alter gene expression profiles,[126] we ignore probesets that were not expressed in hDRG or TG and characterized which of the hDRG- or TG-expressed probesets were expressed in NHSCs or fibroblasts. Gene abundance profiles of RNA-seq data sets were not contrasted with microarray data sets because of the different nature of the inherent bias of the 2 technologies and the limited dynamic range of microarrays.[86]

For RNA-seq, we analyzed only genes with polyA+ transcripts, and by performing upper quartile normalization, we aimed to minimize the effect of technical variation in our data. Within human data sets, genes were analyzed for coexpression patterns across all analyzed tissues, to identify sets of genes with similar expression patterns and cluster them into "modules."[24] We assigned each gene a digital "co-expression pattern," based on clustering the gene's TPM values into 2 top-level clusters. The most frequent gene coexpression patterns were identified, along with transcription factors (TFs) and other relevant genes in each module. The genes in the corresponding modules were then hierarchically clustered[24] and the resulting dendrogram optimally leaf ordered on the genes[6] to depict the gene expression patterns across the set of analyzed human tissues. This allowed us to visualize relevant coexpression patterns and putatively identify regulatory circuitry underlying such coexpression. In addition, this allows for simplification of downstream analyses such as coexpression module detection, identifying regulatory networks or gene set enrichment analysis.[12]

We also performed hierarchical clustering of all human and mouse tissue transcriptome profiles, restricted to the set of genes to those with one-to-one orthology between human and mouse based on our analysis.[28] For the hDRG and mDRG, when all 3 individual replicates were individually used instead of a single tissue profile to identify the sample-to-sample variation relative to observed tissue-to-tissue variation. The set of genes used for this analysis was limited to genes with tissue-restricted expression to identify a clear tissue-specific gene signature.[90]

### 2.5. Functional enrichment analysis

We identified human and mouse genes with relevant functional gene ontology descriptions, based on the Gene Ontology (GO) database.[16] Wherever possible, the set of human genes were augmented using the HGNC database.[32] The gene sets used can be downloaded from our web site. Functional enrichment analysis was then performed using Enrichr,[12] allowing for multiple testing correction, with the Reactome database[18] being used for pathway analysis. To identify all known drug–gene product interactions for genes present in the conserved DRG-enriched gene set, the Drug–Gene Interaction Database (DGIdb)[115] was used. Protein–protein interactions were mined based on validated interactions of genes with relevant GO terms in the StringDB database.[27] The KEGG database and visualization tool[46] was used to analyze the neurotrophin signaling pathway, used with permission from Kanehisa Laboratories (available on our web site).

### 2.6. In silico detection of mRNA axonal localization

We contrasted relative abundances of mRNAs in the hDRG from our own RNA-seq experiments to human tibial nerve (hTN) RNA-seq experiments from the GTex consortium.[17] Changes in gene expression between the hDRG and hTN can be brought about by

region-specific gene expression in nonneuronal cells, subcellular localization in sensory neurons, variable composition of cell types between hDRG and hTN samples, differential expression between sensory (present in hDRG and hTN) and motor neurons (present only in hTN), or a combination of all these. Thus, the analysis was restricted to characterizing genes known to be present in the axonal compartment of sensory neurons in adult rats.[35] Uniquely identifiable human and mouse orthologs of the rat genes in our curated Homologene database were used in the analysis. The hDRG:hTN fold change was calculated based on TPM values for all genes, and is presented in Sheet 1 of the human RNA-seq data, available on our web site.

## 3. Results

We introduce an open-data, searchable web site for our analyses allowing users to visualize gene expression profiles (https://www.utdallas.edu/bbs/painneurosciencelab/sensoryomics/drgtxome) across orthologous human and mouse genes. For reproducibility, code for performing calculation of the data underlying the figures: uqTPMs, and sample statistics for characterizing tissue specificity and conservation of expression are available on our web site.

### 3.1. Comparison of human dorsal root ganglia whole transcriptome profile with other transcriptome profiles

From a transcriptional point of view, the hDRG is a highly heterogeneous conglomerate of cells with multiple types of sensory neurons, Schwann cells, SGCs, fibroblasts, and resident macrophages.[76] In vivo studies of heterogeneous tissues such as the hDRG or hTN pose a special challenge because mRNA species may be expressed specifically in individual cell types or more generically. We estimated the degree of overlap between the whole DRG transcriptomic profiles and some of its non-neuronal constituent cell types. We looked at assembled microarray data sets for human DRG, TG, and cultured NHSCs and skin fibroblasts (FIBRO) (**Fig. 1A**). We identified 12,422 probesets that show expression in hDRG. Of these, 11,407 (94.6%) are shared with the human TG, reiterating the overall similarity of the 2 tissues. Interestingly, we found that approximately 70% (8666/12,422) of the DRG-expressed probesets show coexpression in either the NHSC or fibroblast data set or both. The remaining 30% of the DRG-expressed probesets were putatively due to expression restricted to sensory neurons or SGCs and included nociceptor-enriched channels such as SCN10A and SCN11A,[112] and nociceptor-enriched TFs such as PRDM12[13] which are well characterized in mammalian DRG. Our analysis indicates that cDNA libraries constructed from whole hDRG have a strong nonneuronal component.

To look at conservation of gene expression across tissues in humans and mice in a transcriptome-wide fashion, we performed hierarchical clustering of the RNA-seq transcriptomes (**Fig. 1B**). We find that for nonneural tissues, homologous human and mouse tissues cluster in pairs. The hDRG replicates, mDRG replicates, and hTN sample also cluster together, showing a distinct peripheral nervous system (PNS) gene expression profile, and low sample-to-sample variation in our hDRG and mDRG samples. Our findings agree with meta-analyses that suggest distinct signatures for human sensory tissues and brain tissues,[106] and that interstudy distances among homologous tissue transcriptomes are typically less than intrastudy distances between different tissues.[105] However, inside the CNS tissue cluster, we find that the human samples and mouse samples do
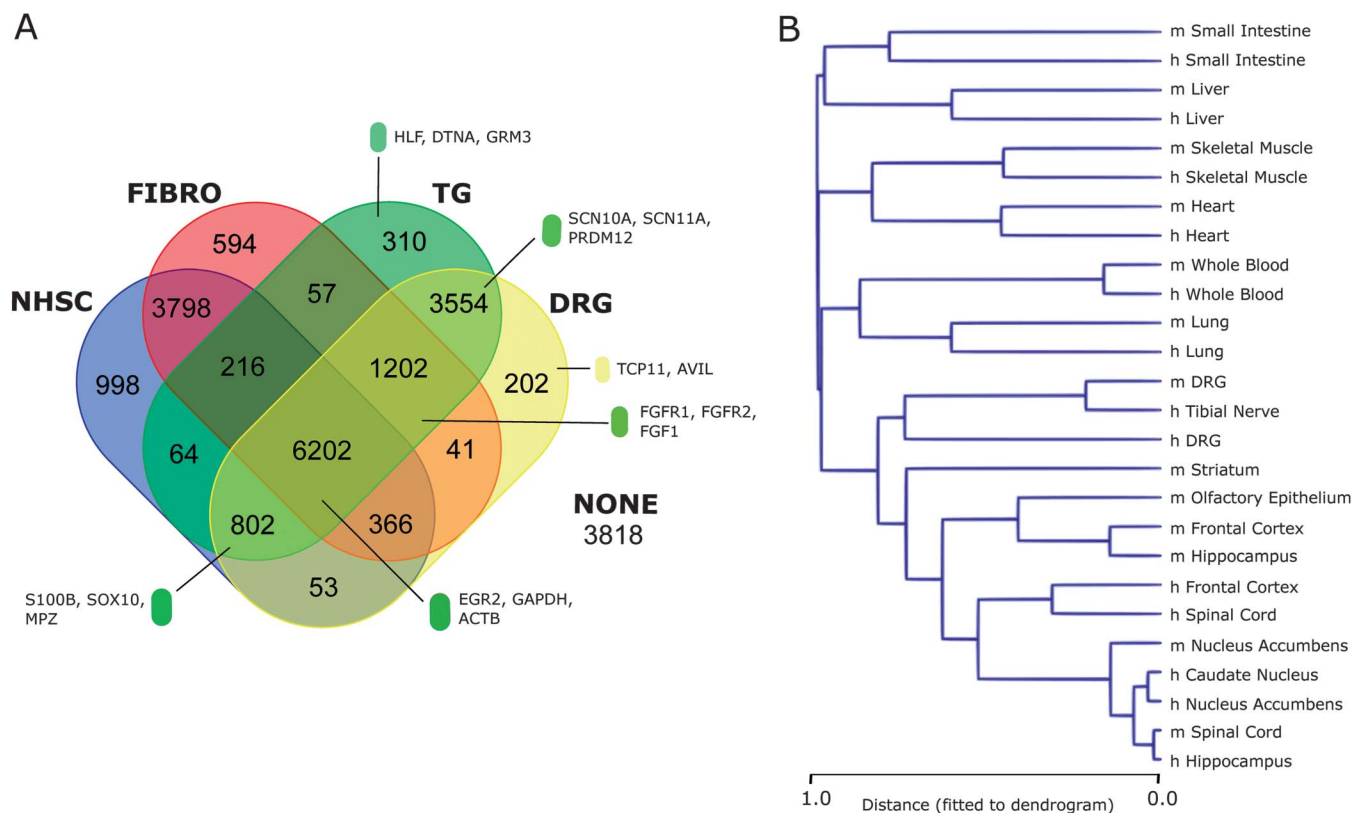
**Figure 1.** Cell-type–specific and cell compartment–specific gene expression in the hDRG. (A) Microarray analysis of human DRG, TG, cultured Schwann cells (NHSC), and fibroblasts (FIBRO) reveals that over 3500 probes are expressed in vivo in DRG and TG but not detected in Schwann cells or fibroblasts, suggesting regulatory programs driving gene expression specifically in sensory neurons and SGCs. (B) Clustering of human and mouse tissue transcriptomes based on a Pearson correlation coefficient (PCC)-based distance measure (1 − PCC) and average linkage for tissue-restricted genes. DRG, dorsal root ganglia; hDRG, human DRG; NHSC, normal human Schwann cell; SGC, satellite glial cell; TG, trigeminal ganglia.

not cluster according to brain regions, and we attribute this to well-characterized evolutionary divergence of the CNS transcriptome in humans with respect to other primates and rodents[66,80] rather than batch or laboratory effects, as the subclusters are not determined by source laboratory of the samples. Preclinical rodent models have been shown to mimic human response in several domains.[106] Our analysis identifies a distinct, evolutionarily conserved tissue-specific signature for healthy human and mouse PNS tissues, opening the door to the follow-up question of how faithfully rodent pain models correlate with human pain signatures at the molecular level.

### 3.2. Coexpression patterns and regulatory programs shaping the human dorsal root ganglia

Based on our binary coexpression pattern for each coding or noncoding human gene, we found that 22,140 genes were expressed in 1 or more tissues, with 12,462 genes being generically expressed across the reference tissue panel based on their entropy (Supplementary Figs. 5A and 5B, available online at http://links.lww.com/PAIN/A559). Of the remaining 9498 genes, we found a power law–like distribution for frequencies of possible coexpression patterns across the tissues we queried (Supplementary Fig. 5C, available online at http://links.lww.com/PAIN/A559), suggesting that a few dominant regulatory paradigms primarily shape the hDRG transcriptome. Of the top 25 most frequent patterns (5542 of the 9498 tissue-restricted genes in our analysis), 15 correspond to enrichment or de-enrichment in subsets of nonneural tissues, 6 to enrichment in subsets of CNS tissue (Supplementary Fig. 5D,

available online at http://links.lww.com/PAIN/A559), and the remaining 4 affecting hDRG expression. The most common coexpression pattern of these was the set expressed pan-neuronally (**Fig. 2A**), the basis of a neuronal gene signature.[25] Previous analyses with hDRG microarray data sets were limited to analyzing a set of predefined probesets, and were unable to identify a broad pan-neuronal transcriptome signature including the hDRG.[94,104] In our analysis, pan-neuronal genes include splicing factor RBFOX3, ion channels (SCN1A and KCNK1), tyrosine receptor kinase NTRK2, cell adhesion molecule (CAM) gene DOCK3, and the early neural lineage TF SOX1.[87] The second relevant gene coexpression module consists of genes which are downregulated in the hDRG and spinal cord but expressed in the rest of the CNS (**Fig. 2B**), potentially driven by TFs such as HOXA3, HOXA4, and HOXB3 involved in rostral–caudal patterning[77] and are common to these 2 adult human tissues but not more rostral parts of the nervous system. The TF MYT1L is downregulated in the hDRG and spinal cord with respect to the rest of the CNS. A third prominent coexpression pattern identifies genes enriched primarily in the CNS but undetectable in the hDRG (**Fig. 2C**). These include CNS-expressed genes such as SLC32A1[79] and GLRA2,[125] known to be suppressed in the DRG, the metabotropic glutamate receptor GRM5 (linked to postsynaptic plasticity in the spinal dorsal horn[40]), RNA-binding protein PABPC1L2A, serotonin 2C receptor HTR2C, and TFs OLIG2 (present in astrocyte and oligodendrocyte cell types, and subpopulations of motor neuron[61]) and POU3F4. The fourth prominent coexpression pattern was primarily enriched in the hDRG (**Fig. 2D**), with the gene set explored in more detail in Section 3.3.
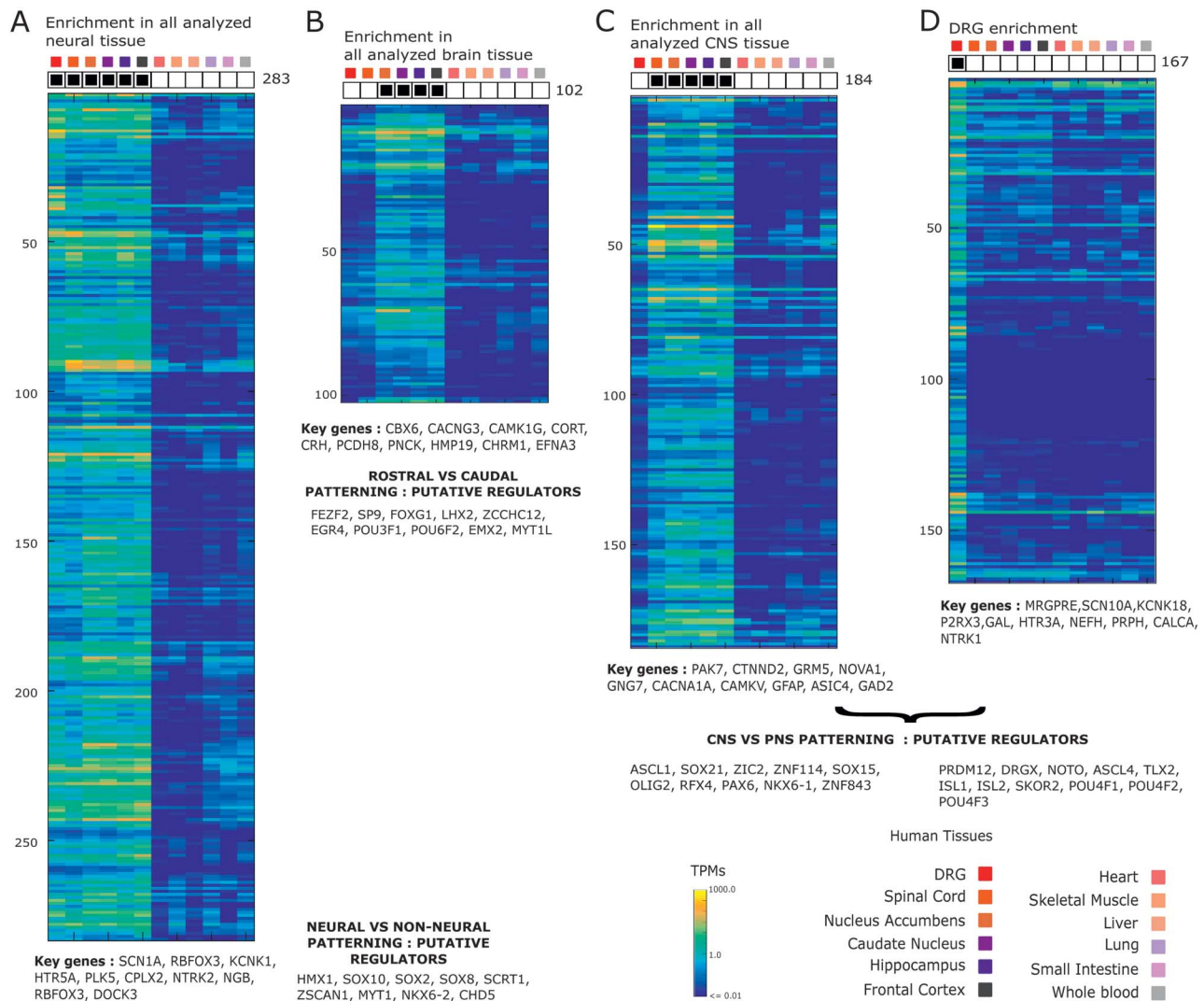
**Figure 2.** Identification of common tissue-restricted human gene coexpression patterns in our analyzed tissue panel. Heatmaps with hallmark genes and potential transcriptional regulators, which are differentially expressed in (A) neural vs nonneural tissue, (B) brain tissues vs remaining tissue, (C) CNS vs non-CNS tissue, and (D) DRG vs remaining tissues. CNS, central nervous system; DRG, dorsal root ganglia; PNS, peripheral nervous system; TPM, transcripts per million.

To identify a more comprehensive picture of the regulatory forces driving gene expression in the hDRG, we catalogued all hDRG-enriched and neurally enriched TFs (heatmaps for human and orthologous mouse gene expression in **Fig. 3A**). Multiple evolutionary developmental studies have been performed,[13,58,73] suggesting an important role of TLX3, RUNX1, DRGX, POU4F1, and PRDM12 in mouse, human, and *Xenopus* sensory tissue development. We find that most of the studied regulatory TFs have conserved DRG-enriched gene expression in adult humans and mouse, suggesting that regulatory interactions are conserved through evolutionary history. The most well studied among these TFs is possibly PRDM12, which is essential for mammalian DRG development and pain sensation.[13] The ancient origins of somatic and visceral neurons are well documented in bilaterians (arthropods and vertebrates),[78] and among the transcriptional determinants of such neurons are DRG-enriched TFs DRGX, POU4F1, and POU4F2 (both of the BRN family). Based on mDRG scRNA-seq data,[112] several of the DRG-enriched TFs (HOXD1, PRDM12, POU4F2, POU4F3, and ISL2) are restricted to subpopulations of adult mDRG neurons, whereas POU4F1 and ISL1 are more widely expressed across

subpopulations. The TFs may thus collaborate in subpopulation-specific ways to drive regulatory programs specific to each of the neuronal subpopulations. In vitro differentiation of human pluripotent cells using POU4F1 (and either NGN1 or NGN2),[8] reprogramming of human fibroblasts using a TF cocktail of multiple subpopulation-restricted TFs including ISL2,[116] and characterization of sensory neurons induced from human pluripotent stem cells uniformly expressing ISL1 and POU4F1,[11] suggest that a transcriptional program similar to mDRG neurons may be at work in humans.

Several hDRG-expressed TFs, cofactors, splicing regulators, and mRNA-binding proteins (both spliceosomal and nonspliceosomal) show neural enrichment, but are only weakly hDRG enriched (**Figs. 3A and B**). Several of these are potentially involved in hDRG function and are described in Supplementary Table 2 (available online at http://links.lww.com/PAIN/A559).

These analyses give a global view of DRG gene expression in humans compared with a number of other tissues, highlighting that the DRG expression profile is distinct from CNS gene expression patterns[94,104] driven apart during development by well-characterized CNS-specific TFs such as POU3F3 and
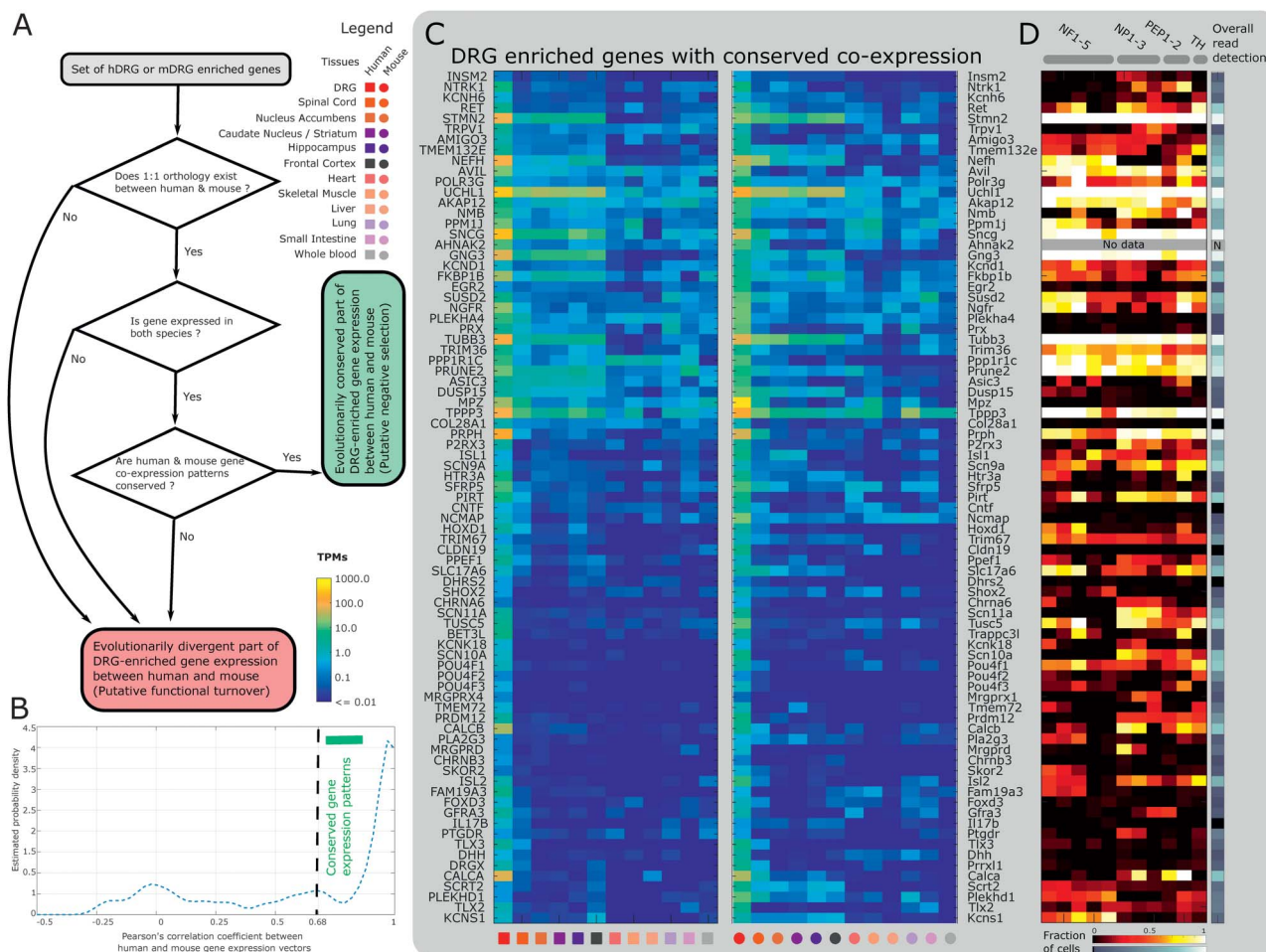
Figure 3. Identification of hDRG-enriched genes. (A) Flowchart for identifying DRG-enriched genes with conserved expression patterns in human and mouse. (B) Estimated density function in DRG-enriched genes for Pearson correlation across vectors of relative abundance in orthologous tissues. (C) The set of hDRG- and mDRG-enriched genes (based on the DRG-enrichment score) with one-to-one orthology and correlated gene expression abundances (based on Pearson correlation) across our panel of analyzed tissues. Many of the 81 genes belong to gene families known to play important functional roles in nociceptors, including GPCRs, ion channels, TFs, and transmembrane proteins. (D) Assessment of expression of hDRG genes in mDRG scRNA-seq data[112] for prediction of potential sensory neuronal subtype expression. DRG, dorsal root ganglia; GPCR, G-protein-coupled receptor; hDRG, human DRG; mDRG, mouse DRG; TF, transcription factor; TPM, transcripts per million.

SOX21 and DRG-specific TFs such as PRDM12, DRGX, and TLX3. Although DRG-specific[58,73,91] and CNS-specific TFs[70,117] have well-appreciated roles in development, their persistent expression into late adulthood in humans has previously not been appreciated, suggesting a role of cell-type maintenance alongside that of the well-characterized developmental lineage drivers. In addition, multiple DRG-enriched TFs (ISL1, ISL2, and DRGX) are known to play a dual role in developmental axonogenesis and adult axon regeneration. Human DRG-specific TFs are also likely involved in regulating the transcription of hDRG-specific genes, although specific TF targeting of DRG-specific genes underlying mammalian regulatory programs is not well understood, except for a small subset of genes profiled in evolutionary development studies.[13,58,73]

### 3.3. Human dorsal root ganglia and mouse dorsal root ganglia–enriched genes and a conserved evolutionary signature

Based on the DRG-enrichment score, we identified 140 protein-coding genes to be hDRG enriched and 141 protein-coding genes to be mDRG enriched. Of these, 128 hDRG-enriched genes and 119 mDRG-enriched genes (identified based on

Section 2.2, **Fig. 4A**) have one-to-one orthologs in the other species. We then looked at the correlation between the human and mouse ortholog gene expression across the analyzed tissue panel to identify whether the gene in the other species also had a DRG-enriched expression profile, reported in **Table 2**. When we analyzed the distribution of correlation values, we found a large population of highly correlated relative abundance, suggesting that a large fraction of these genes are under negative evolutionary selection in their regulatory regions, causing conserved gene expression profiles across tissues (**Fig. 4B**). Eighty-one of 128 (63.3%) human genes have correlation scores above 0.68 (the peak in the distribution before the primary mode). All 81 of these genes in both species have the highest relative abundance in DRG across the panel of analyzed tissues, showing a clear conservation of DRG enrichment. Among the genes we identified with a common pattern of DRG enrichment in mouse and human (**Fig. 4C**), are several that have not been previously studied in DRG biology. These include UCHL1 (neuroepidermal marker of itch[54]), SKOR2 (a developmental corepressor in the cerebellum[74]), TRIM67 (shown to be involved in RAS-mediated signaling in the cerebellum[122]), BET3L (particle transport complex member identified in mDRG single-cell studies[56] but not
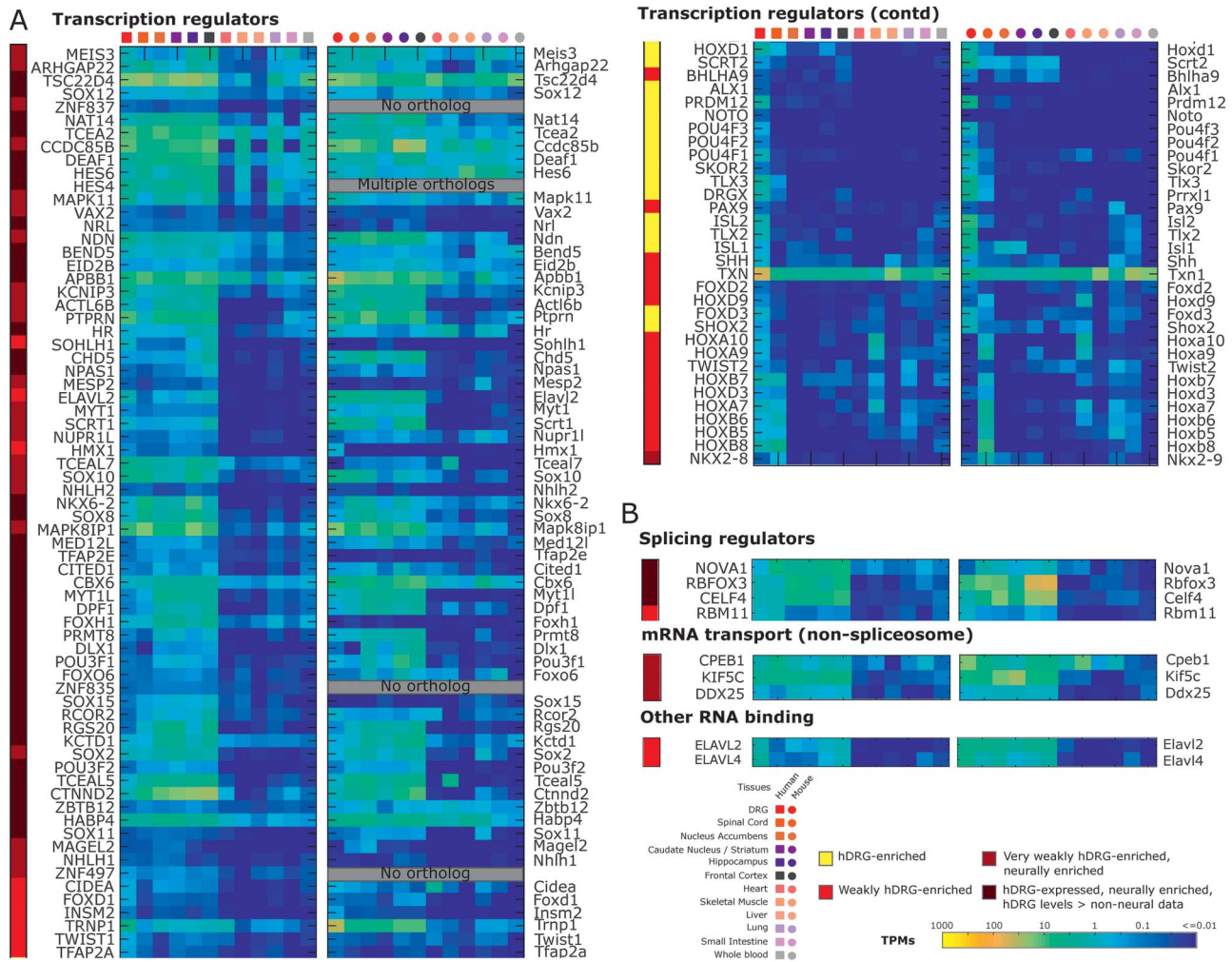
**Figure 4.** Human DRG–enriched genes that potentially regulate the transcriptomic landscape. Orthologous human and mouse gene expression in the corresponding tissues is depicted as heatmaps. Differentially expressed genes that are hDRG enriched or neural enriched are shown. Gene families characterized are TFs (A), splicing factors, mRNA transport molecules, and RNA-binding proteins are shown (B). DRG, dorsal root ganglia; hDRG, human DRG; TF, transcription factor; TPM, transcripts per million.

functionally studied), and TLX2 (whose paralog TLX3 is well characterized as a key DRG-enriched TF).

Several DRG-enriched genes show subpopulation-restricted expression in the mDRG (expression profiles in mDRG neuronal subpopulations based on Usoskin et al.[112] are shown for all DRG-enriched genes in **Fig. 4D** and Supplementary Table 3, available online at http://links.lww.com/PAIN/A559), raising the possibility of a similar expression profile in sensory neuronal subpopulations in the hDRG. Among the more well-studied genes are NTRK1, TRPV1, and CALCA (mRNA found in both peptidergic and nonpeptidergic neurons), ASIC3 (neurofilament and peptidergic neurons), MRGPRD (nonpeptidergic neurons), and PRDM12 (peptidergic, nonpeptidergic, and TH-positive neurons).

We also identified a gene set that has undetectable gene expression in mDRG scRNA data set, many of which are known to be expressed in mammalian Schwann cells, detailed in Supplementary Table 4 (available online at http://links.lww.com/PAIN/A559). Known peripheral SGC markers are either weakly expressed in our hDRG data sets (such as PAX3 or PAX7) or also expressed in other tissues (such as glial fibrillary acidic protein) and are thus not DRG enriched. Fibroblast markers such as S100A4 are highly expressed in DRG, but are also generically expressed, leading to low DRG-enrichment scores.

We identified that many of the gene products of hDRG-enriched genes can be assigned a handful of molecular roles including transcriptional activators or repressors, receptor and receptor binding, kinases or phosphatases, ion channels, and transmembrane domain containing proteins. Based on the Enrichr tool,[53] we performed enrichment analysis for biological processes in the Gene Ontology (GO) database[3] (Supplementary Table 5, available online at http://links.lww.com/PAIN/A559) and molecular pathways in the Reactome database[18] (Supplementary Table 6, available online at http://links.lww.com/PAIN/A559), which were enriched in our gene set (the Fisher exact test, $P < 0.01$ after multiple testing correction). Gene or protein functional ontologies are well known to be incomplete,[44] necessitating us to look at DRG-enriched genes in nonenriched ontology terms and the literature, and showing the diverse roles played by these gene products. Several of these are already well studied in DRG biology. CALCB is a paralog of CALCA with similar function. MRGPR family members are known to play a key role in itch sensation.[57] PIRT is a well-known regulatory subunit of TRPV1,[48] whereas PRDM12 is a key TF in neural sensory lineages.[13] Querying the Reactome database yielded enrichment for acetylcholine and nicotinic receptor signaling centered around 3 specific receptor subunits, β3, α6, and α9. The CHRNA6 gene is potentially involved in

**Table 2**

**Conservation of expression of hDRG- and mDRG-enriched genes between species.**

High correlation ( >= 0.68 )

| Human gene | Mouse gene | Pearson's R |
|---|---|---|
| TRIM36 | Trim36 | 0.99 |
| SUSD2 | Susd2 | 0.88 |
| INSM2 | Insm2 | 0.96 |
| ISL2 | Isl2 | 1.00 |
| PLEKHA4 | Plekha4 | 0.99 |
| PRDM12 | Prdm12 | 1.00 |
| GFRA3 | Gfra3 | 0.98 |
| SLC17A6 | Slc17a6 | 0.74 |
| TRPV1 | Trpv1 | 0.96 |
| AHNAK2 | Ahnak2 | 0.99 |
| KCNK18 | Kcnk18 | 1.00 |
| PLEKHD1 | Plekhd1 | 0.98 |
| FGF4 | Fgf4 | 1.00 |
| ISL1 | Isl1 | 1.00 |
| DRGX | Prrxl1 | 1.00 |
| CLDN19 | Cldn19 | 0.99 |
| PPP1R1C | Ppp1r1c | 0.94 |
| TMEM72 | Tmem72 | 0.98 |
| NGFR | Ngfr | 1.00 |
| FAM19A3 | Fam19a3 | 0.82 |
| MRGPRD | Mrgprd | 0.93 |
| PRUNE2 | Prune2 | 0.87 |
| NTRK1 | Ntrk1 | 1.00 |
| OC90 | Oc90 | 0.97 |
| P2RX3 | P2rx3 | 1.00 |
| EGR2 | Egr2 | 0.72 |
| POU4F3 | Pou4f3 | 1.00 |
| KCNS1 | Kcns1 | 0.89 |
| CHRNA6 | Chrna6 | 1.00 |
| POU4F2 | Pou4f2 | 1.00 |
| ASIC3 | Asic3 | 0.70 |
| KCND1 | Kcnd1 | 0.93 |
| POU4F1 | Pou4f1 | 1.00 |
| SCN10A | Scn10a | 1.00 |
| SCN9A | Scn9a | 0.93 |
| DHH | Dhh | 0.99 |
| GNG3 | Gng3 | 0.86 |
| SFRP5 | Sfrp5 | 0.99 |
| TLX3 | Tlx3 | 1.00 |
| SNCG | Sncg | 0.99 |
| PPM1J | Ppm1j | 0.93 |
| NCMAP | Ncmap | 0.94 |
| TPPP3 | Tppp3 | 0.98 |
| KCNH6 | Kcnh6 | 0.98 |
| DUSP15 | Dusp15 | 0.97 |
| CHRNB3 | Chrnb3 | 0.97 |
| NMB | Nmb | 0.99 |
| AKAP12 | Akap12 | 0.92 |
| UCHL1 | Uchl1 | 0.96 |
| POLR3G | Polr3g | 0.95 |
| AVIL | Avil | 1.00 |
| KCNA10 | Kcna10 | 0.95 |
| NEFH | Nefh | 0.90 |
| TMEM132E | Tmem132e | 0.95 |
| MPZ | Mpz | 0.98 |
| AMIGO3 | Amigo3 | 0.95 |
| PRPH | Prph | 1.00 |
| TRIM67 | Trim67 | 0.99 |
| FOXD3 | Foxd3 | 0.97 |
| STMN2 | Stmn2 | 0.73 |
| PPEF1 | Ppef1 | 0.99 |
| SKOR2 | Skor2 | 1.00 |
| TUSC5 | Tusc5 | 1.00 |
| COL28A1 | Col28a1 | 0.98 |
| HTR3A | Htr3a | 1.00 |
| FKBP1B | Fkbp1b | 0.88 |
| TUBB3 | Tubb3 | 0.96 |
| SHOX2 | Shox2 | 0.98 |
| DHRS2 | Dhrs2 | 0.96 |
| PTGDR | Ptgdr | 0.70 |
| RET | Ret | 0.99 |
| TLX2 | Tlx2 | 1.00 |
| PRX | Prx | 0.93 |
| PIRT | Pirt | 0.99 |
| HOXD1 | Hoxd1 | 0.98 |
| SCN11A | Scn11a | 1.00 |
| BET3L | Trappc3l | 1.00 |
| CNTF | Cntf | 0.96 |
| CLRN2 | Clrn2 | 0.98 |
| IL17B | Il17b | 0.87 |
| CALCB | Calcb | 1.00 |
| CALCA | Calca | 1.00 |
| PLA2G3 | Pla2g3 | 0.83 |
| SCRT2 | Scrt2 | 0.95 |
| MRGPRX1, MRGPRX4 | Mrgprx1, Mrgprb4, Mrgprb5, Mrgpra3 | Noted pairs >=0.68 |

Marginal correlation ( >= 0.5 and < 0.68 )

| Human gene | Mouse gene | Pearson's R |
|---|---|---|
| NDST4 | Ndst4 | 0.67 |
| VWA7 | Vwa7 | 0.65 |
| MRGPRE | Mrgpre | 0.60 |
| STAC | Stac | 0.59 |
| IBSP | Ibsp | 0.67 |
| CCDC172 | Ccdc172 | 0.64 |
| FABP7 | Fabp7 | 0.56 |
| HMX1 | Hmx1 | 0.65 |
| KLRG2 | Klrg2 | 0.54 |
| TMEM233 | Tmem233 | 0.50 |
| CDK15 | Cdk15 | 0.56 |
| TRPA1 | Trpa1 | 0.59 |
| FAM70A | Tmem255a | 0.67 |

Poor correlation (<0.5)

| Human gene | Mouse gene | Pearson's R |
|---|---|---|
| CRYBA2 | Cryba2 | 0.33 |
| RDH12 | Rdh12 | -0.10 |
| TMEM239 | Tmem239 | 0.18 |
| TNFSF12-TNFSF13 | Tnfsf12Tnfsf13 | 0.06 |
| GCGR | Gcgr | 0.11 |
| ANGPTL7 | Angptl7 | 0.04 |
| SMCP | Smcp | -0.22 |
| NEFM | Nefm | -0.19 |
| GLDN | Gldn | -0.02 |
| ADAMTS16 | Adamts16 | 0.21 |
| KCNG3 | Kcng3 | 0.08 |
| GPHA2 | Gpha2 | -0.06 |
| FRMD3 | Frmd3 | -0.06 |
| A3GALT2P | A3galt2 | -0.10 |
| ANKRD33 | Ankrd33 | 0.03 |
| CLRN1 | Clrn1 | 0.11 |
| KCTD19 | Kctd19 | -0.13 |
| LHFPL1 | Lhfpl1 | 0.01 |
| GPR142 | Gpr142 | 0.25 |
| QRFPR | Qrfpr | -0.06 |
| C11orf53 | 1810046K07Rik | -0.22 |
| DAPL1 | Dapl1 | -0.03 |
| C1QL4 | C1ql4 | 0.06 |
| SHC4 | Shc4 | 0.31 |
| ASCL4 | Ascl4 | 0.15 |
| GABRA6 | Gabra6 | -0.20 |
| GPX2 | Gpx2 | 0.00 |
| PMP2 | Pmp2 | 0.01 |
| TSPAN8 | Tspan8 | 0.29 |
| KLK8 | Klk8 | -0.20 |
| ANGPT2 | Angpt2 | -0.15 |
| PMFBP1 | Pmfbp1 | -0.02 |
| TRPM8 | Trpm8 | -0.04 |
| KCNG4 | Kcng4 | 0.06 |
| TAC1 | Tac1 | 0.47 |
| GNAT3 | Gnat3 | 0.10 |
| PHLDA2 | Phlda2 | 0.44 |
| F2RL2 | F2rl2 | 0.24 |
| CRYGD | Crygd | NC |
| CCKAR | Cckar | 0.21 |
| TIMP4 | Timp4 | -0.23 |
| HTR3B | Htr3b | 0.43 |
| TECTA | Tecta | -0.05 |
| SPTBN5 | Sptbn5 | 0.45 |
| NEFL | Nefl | -0.18 |
| SERTM1 | Sertm1 | 0.13 |
| TYRP1 | Tyrp1 | -0.06 |
| MIA | Mia | -0.34 |
| TPBGL | Gm4980 | 0.31 |
| CSHL1 | Gh | -0.09 |
| FGFBP3 | Fgfbp3 | -0.14 |
| ALX1 | Alx1 | 0.38 |
| IL31RA | Il31ra | -0.22 |
| BIRC7 | Birc7 | 0.08 |
| FSTL1 | Fstl1 | 0.32 |
| SERPINE3 | Serpine3 | -0.24 |
| NOTO | Noto | -0.13 |
| CCL1 | Ccl1 | 0.04 |
| CST5 | Cst10 | -0.22 |
| OTOA | Otoa | -0.05 |
| PRL | Prl | -0.12 |
| DLX3 | Dlx3 | 0.07 |
| KLK5 | Klk5 | -0.21 |
| GAL | Gal | 0.04 |
| KRT27 | Krt27 | 0.25 |
| INSC | Insc | 0.08 |
| SYNGR4 | Syngr4 | -0.08 |
| GRXCR2 | Grxcr2 | -0.10 |
| FXYD2 | Fxyd2 | -0.19 |
| FAM178B | Fam178b | -0.09 |
| VSX1 | Vsx1 | 0.18 |
| SOSTDC1 | Sostdc1 | -0.04 |
| PRLH | Prlh | -0.11 |
| CHRNA9 | Chrna9 | -0.12 |

Human DRG-enriched and mDRG-enriched genes are shown in pairs, when they are direct orthologues with the corresponding Pearson correlation. Darker green implies higher correlation.
DRG, dorsal root ganglia; hDRG, human DRG; mDRG, mouse DRG.

mechanical allodynia in mouse, leading to an identification of a novel interaction between α6 nicotinic receptors and P2X3 ion channels, which leads to inhibition of P2X3 channel activity.[118] CHRNA9 (encoding the α9 nicotinic receptor subunit) was recently implicated in development of chemotherapy-induced peripheral neuropathic pain using an antagonist of α9 and α10 nicotinic receptors.[84] Although more work is needed to validate α9 as the mechanistic target for this compound, the very strong enrichment of CHRNA9 in human DRG makes it an interesting target for this prominent form of neuropathic pain.

More interestingly, we identified many genes whose functional roles in DRG have not been well studied (**Fig. 5C**). PPP1R1C, PPM1J, and PPEF1 are phosphatases: PPM1J is known to regulate neurite growth,[2] PPEF1 interacts with calmodulin in a Ca$^{+2}$-dependent manner,[55] and DUSP15 regulates the JNK pathway.[99] AMIGO3[52] is involved in axonal tract development. SUSD2 is known to regulate neurite growth in hippocampal cultures.[72] TPPP3, PLA2G3, and TRIM36[103] (along with TUBB3) are associated with the microtubule cytoskeleton. AKAP12 is a kinase scaffold involved

in regulating cAMP biosynthesis and signaling. TUSC5 is expressed in the PNS and adipose tissue, and potentially performs shared adipose-nervous system functions.[81] PTGDR and IL17B are known to be involved in inflammation response, and POLR3G is also involved in cytokine production and immune response. BET3L is part of the TRAPP protein trafficking complex, whereas NMB is a neuropeptide involved in nociceptive signaling.[69] In addition, we have a very limited understanding of the role of some DRG-enriched neuronally expressed genes. INSM2, SHOX2, SCRT2, and SKOR2 are development-related transcriptional regulators, but their regulatory function in adult DRG is not known. These DRG-enriched genes create a rich set of novel targets for potential exploration in the pain neurobiology space.

### 3.4. Identifying potential pharmacological targets among tissue-restricted genes in the human dorsal root ganglia

Because a primary goal of our analysis is the identification of potential therapeutics, we performed additional analysis on our
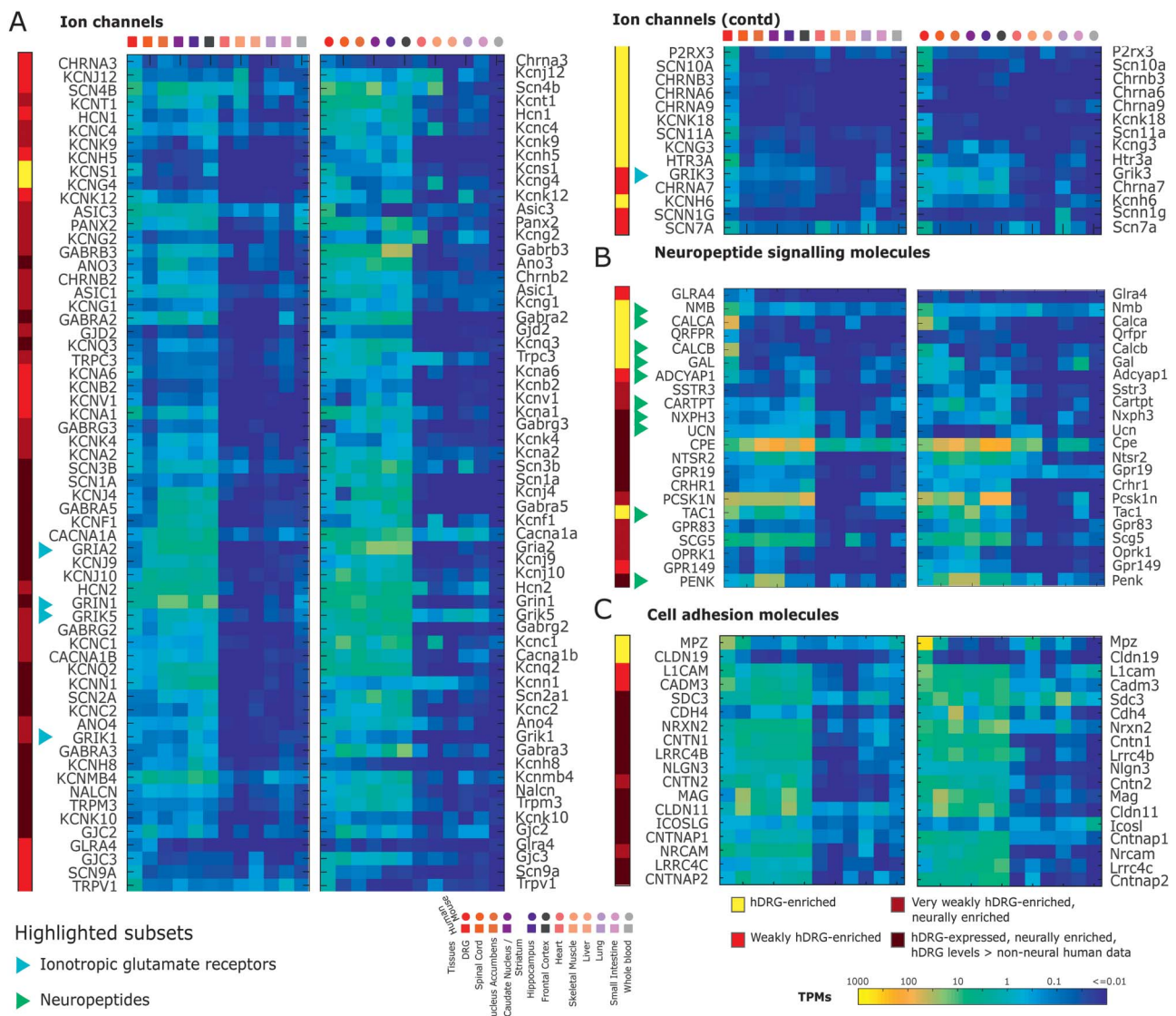


**Figure 5.** Human DRG–enriched genes for pharmacological targets: ion channels, neuropeptides, and cell adhesion molecules. Orthologous human and mouse gene expression in the corresponding tissues is depicted as heatmaps. Differentially expressed ion channels (A), neuropeptide signaling (B), and cell adhesion molecules (C) show several candidate genes for pharmacological targeting, including several neuropeptides. DRG, dorsal root ganglia; hDRG, human DRG; TPM, transcripts per million.

**Table 3**

**Known drug–hDRG-enriched gene product interactome from DGIdb database.**

| Gene | Role: drug {reference: PUBMED ID/drug database} |
|---|---|
| CALCA | Inhibitor: OLCEGEPANT {TTD, DrugBank, 17665333}; |
| CHRNA6 | Agonist: CYTISINE {DrugBank, 25517706, 11553677}; and NICOTINE {DrugBank, 11006350, 19560048, 20081230, 20400469}; <br> Partial agonist: VARENICLINE {DrugBank, 16766716}; <br> Unknown/NA: ETHANOL {DrugBank, 19270242, 23641218}; HEXAMETHONIUM {GuideToPharmacologyInteractions}; and MECAMYLAMINE {GuideToPharmacologyInteractions}; |
| CHRNA9 | Agonist: NICOTINE {DrugBank, 11006350, 19560048, 20081230, 20400469, GuideToPharmacologyInteractions, TdgClinicalTrial, TEND}; <br> Agonist/inhibitor: TETRAETHYLAMMONIUM {DrugBank, 12824448}; <br> Unknown/NA: ETHANOL {DrugBank, 19270242, 23641218}; LOBELINE {DrugBank}; MUSCARINE {GuideToPharmacologyInteractions}; and STRYCHNINE {GuideToPharmacologyInteractions}; |
| CHRNB3 | Agonist: NICOTINE {DrugBank, 11006350, 19560048, 20081230, 20400469}; <br> Unknown/NA: ETHANOL {DrugBank, 19270242, 23641218}; |
| HTR3A | Agonist: CISAPRIDE {DrugBank, 10328995, 1600046, 9697104, 8093733}; METOCLOPRAMIDE {DrugBank, 2955843, GuideToPharmacologyInteractions}; antagonist: ALOSETRON {DrugBank, 10776833, 20136586, 11752352, 10354345, 10848662, 10429744, 11060667, 18555935, 15061683, 17138606, 14596662, GuideToPharmacologyInteractions, TdgClinicalTrial, TEND}; ALOSETRON_HYDROCHLORIDE {ChEMBL}; AMOXAPINE {DrugBank, 1666997}; ARIPIPRAZOLE {DrugBank, 17848919}; ARVERAPAMIL {TTD}; CHLOROPROCAINE {DrugBank, 17016423, 17139284}; CLOZAPINE {DrugBank, 17848919, PharmGKB}; DDP-225 {TTD}; DOLASETRON {DrugBank, 11752352, 14635083, 1397053, 14586600, 9506240, 9341357, 8141114, 15740177, 9257083, 7493546, TTD, TdgClinicalTrial, TEND}; DOLASETRON_MESYLATE {ChEMBL}; GRANISETRON {ChEMBL, DrugBank, 12943486, 17093363, 11814868, 10882387, 11752352, 17379085, 10412832, 15740177, 16044252, TTD, GuideToPharmacologyInteractions, TdgClinicalTrial, TEND}; GRANISETRON_HYDROCHLORIDE {ChEMBL}; LERISETRON {TTD}; MEMANTINE {DrugBank, 11403963}; METHADONE {DrugBank, 19131665}; MIRTAZAPINE {DrugBank, 10446734, 11752352, 17587360, 10446735, 8930006, 11607047, TdgClinicalTrial, TEND}; OLANZAPINE {DrugBank, 17848919, TdgClinicalTrial}; ONDANSETRON_HYDROCHLORIDE {ChEMBL}; PALONOSETRON_HYDROCHLORIDE {ChEMBL}; PROCAINE {DrugBank, 7539114, 11752352, TTD}; RENZAPRIDE {TTD, DrugBank, 16696817}; ROCURONIUM {DrugBank, 10648343, TdgClinicalTrial}; TUBOCURARINE {DrugBank, 12243766, 10066903, 16723497, 1691468, 7684066, GuideToPharmacologyInteractions}; VORTIOXETINE_HYDROBROMIDE {ChEMBL, 21486038}; and ZIPRASIDONE {DrugBank, 17848919}; <br> Agonist/antagonist: ONDANSETRON {ChEMBL, DrugBank, 11752352, 11919526, 17269994, 9506240, 11972287, 10065930, 19320280, 8588861, 15740177, 11763467, 20098535, 12032025, 10517265, GuideToPharmacologyInteractions, TdgClinicalTrial, TEND, TTD}; <br> Binder: LOXAPINE {DrugBank}; and TRIMIPRAMINE {DrugBank, 8863001}; <br> Unknown/NA: {3H}; GRANISETRON {GuideToPharmacologyInteractions}; {3H}; RAMOSETRON {GuideToPharmacologyInteractions}; {3H}; (S)-ZACOPRIDE {GuideToPharmacologyInteractions}; AZASETRON {GuideToPharmacologyInteractions}; BILOBALIDE {GuideToPharmacologyInteractions}; CILANSETRON {DrugBank, 16898618, 15757394, GuideToPharmacologyInteractions, TdgClinicalTrial}; COCAINE {GuideToPharmacologyInteractions}; DILTIAZEM {GuideToPharmacologyInteractions}; ESMIRTAZAPINE {TdgClinicalTrial}; ETHANOL {DrugBank, 19270242}; GINKGOLIDE_B {GuideToPharmacologyInteractions}; PALONOSETRON {GuideToPharmacologyInteractions, TdgClinicalTrial, TEND}; PICROTOXININ {GuideToPharmacologyInteractions}; PUMOSETRAG {TdgClinicalTrial}; QUETIAPINE {DrugBank, 17848919}; QUIPAZINE {GuideToPharmacologyInteractions}; RAMOSETRON {GuideToPharmacologyInteractions}; RICASETRON {GuideToPharmacologyInteractions}; (R)-ZACOPRIDE {GuideToPharmacologyInteractions}; (S)-ZACOPRIDE {GuideToPharmacologyInteractions}; TAPENTADOL {DrugBank, 10592235}; TEDATIOXETINE {TdgClinicalTrial}; TROPISETRON {GuideToPharmacologyInteractions}; and VORTIOXETINE {GuideToPharmacologyInteractions, TdgClinicalTrial}; |
| KCNK18 | Unknown/NA: ARACHIDONIC_ACID {GuideToPharmacologyInteractions}; DESFLURANE {ChEMBL, 15562060, 16192517, 16402123, 17227289}; ENFLURANE {ChEMBL, 16402123, 16973644, 17227289}; HALOTHANE {ChEMBL, 11605899, 11886861, 16402123, 17227289, 20217349, 20519544}; ISOFLURANE {ChEMBL, 11605899, 16402123, 17227289}; and SEVOFLURANE {ChEMBL, 16402123, 21788314, 22226507}; |
| MRGPRX1 | Unknown/NA: CHLOROQUINE {GuideToPharmacologyInteractions}; |
| NTRK1 | Agonist: AMITRIPTYLINE {DrugBank, 19549602}; antagonist: IMATINIB {DrugBank, 17579194, 15832750, 17582306, 16052979, 16782438, TdgClinicalTrial, TEND}; inhibitor: LESTAURTINIB {TTD, TdgClinicalTrial}; REGORAFENIB {DrugBank, MyCancerGenomeClinicalTrial}; unknown/NA: DOVITINIB {ClearityFoundationClinicalTrial}; and ENMD-2076 {TdgClinicalTrial}; |
| P2RX3 | Unknown/NA: BZATP {GuideToPharmacologyInteractions}; |
| QRFPR | Unknown/NA: QRFP {GuideToPharmacologyInteractions}; |
| RET | Antagonist/inhibitor: CABOZANTINIB {DrugBank, 21606412, MyCancerGenomeClinicalTrial, MyCancerGenome, TALC}; inhibitor: AMUVATINIB {MyCancerGenome, TALC, TdgClinicalTrial}; AT9283 {TALC}; IMATINIB {DrugBank, 16782438, TdgClinicalTrial, TEND}; LENVATINIB {TALC}; MOTESANIB {MyCancerGenome, TALC, CIViC, 21422803}; PONATINIB {DrugBank, 23526464, MyCancerGenomeClinicalTrial}; REGORAFENIB {ChEMBL, DrugBank, MyCancerGenome, MyCancerGenomeClinicalTrial, TALC}; SORAFENIB {DrugBank, 16507829, 17016424, 15466206, MyCancerGenome, TTD, GuideToPharmacologyInteractions}; SORAFENIB_TOSYLATE {ChEMBL}; SUNITINIB {TALC, DoCM, 21470995, GuideToPharmacologyInteractions, TdgClinicalTrial, |

Table 3 (*continued*)

| Gene | Role: drug {reference: PUBMED ID/drug database} |
|------|-------------------------------------------------|
| | TEND}; SUNITINIB_MALATE {ChEMBL}; and VANDETANIB {ChEMBL, MyCancerGenome, MyCancerGenomeClinicalTrial, TALC, TTD, ClearityFoundationClinicalTrial, DoCM, 20065189, TdgClinicalTrial}; Unknown/NA: APATINIB {GuideToPharmacologyInteractions}; AZD1480 {CIViC, 23056499}; DOVITINIB {ClearityFoundationClinicalTrial}; ENMD-2076 {TdgClinicalTrial}; and LINIFANIB {GuideToPharmacologyInteractions, TdgClinicalTrial}; |
| SCN10A | Blocker: DIBUCAINE_HYDROCHLORIDE {ChEMBL, 8798723, 9742684}; Inhibitor: BENZOCAINE {DrugBank, 17967784, 17016423, 17139284, 16174788, 19661462, TdgClinicalTrial}; BUPIVACAINE {DrugBank, 20685573, 17016423, 17139284, TdgClinicalTrial}; CHLOROPROCAINE {DrugBank, 17016423, 8418714, 17139284, TdgClinicalTrial, TEND}; COCAINE {DrugBank, 17016423, 17139284, TdgClinicalTrial}; DYCLONINE {DrugBank, 9476974, 17016423, 17139284, 10355012, TdgClinicalTrial}; HEXYLCAINE {DrugBank, 17016423, 17139284, TdgClinicalTrial}; LEVOBUPIVACAINE {DrugBank, 11094008, 17016423, 16418020, 17139284, 10969308, TdgClinicalTrial}; LIDOCAINE {DrugBank, 17077153, 19088384, 20585544, TdgClinicalTrial, TEND}; MEPIVACAINE {DrugBank, 17016423, 20044988, 17139284, TdgClinicalTrial}; ORPHENADRINE {DrugBank, 10344632}; OXYBUPROCAINE {DrugBank, 17016423, 17139284, 19429093, TdgClinicalTrial}; PROCAINE {DrugBank, 17016423, 17139284, 9768788, 8581044, 2167458, TdgClinicalTrial}; PROPARACAINE {DrugBank, 9476974, 17016423, 17139284, TdgClinicalTrial}; ROPIVACAINE {DrugBank, 17016423, 17139284, 10781449, TdgClinicalTrial}; TETRACAINE {TTD}; unknown/NA: BATRACHOTOXIN {GuideToPharmacologyInteractions}; DIBUCAINE {TdgClinicalTrial}; LACOSAMIDE {DrugBank, 10592235}; and TETRODOTOXIN {TdgClinicalTrial}; |
| SCN11A | Inhibitor: COCAINE {DrugBank, 17016423, 17139284, TdgClinicalTrial, TEND}; and ZONISAMIDE {DrugBank, 20025128, 19948168, 15511691, 18433351, 20001433, 14704463, 19557119, TdgClinicalTrial}; Unknown/NA: SAFINAMIDE {TdgClinicalTrial}; and TETRODOTOXIN {TdgClinicalTrial}; |
| TUBB3 | Inhibitor: IXABEPILONE {DrugBank, 18321240, 18945860, 18378531, TdgClinicalTrial}; Unknown/NA: CYT997 {DrugBank}; DOCETAXEL {ClearityFoundationBiomarkers}; EPOTHILONE_B {DrugBank, 17016423, 17139284}; and PACLITAXEL {CIViC, 21576762, ClearityFoundationBiomarkers}; |

Only a handful of hDRG-enriched genes have known drugs that effectively target them, suggesting that further research is required into evaluating hDRG-enriched gene products as therapeutic targets. Drug databases are referenced within curly braces. Numbers within curly braces indicate related Pubmed publication IDs.
hDRG, human dorsal root ganglia.

DRG-enriched and neurally enriched gene sets to mine for known or new drug targets. We first used our hDRG-enriched gene set to explore the drug–gene product interaction database DGIdb[115] for targets of known drugs (**Table 3**). This identifies genes for which ligand interactions are known, potentially identifying new drugs that can be repurposed or redesigned to target pain. We identified a subset of gene–ligand interactions from DGIdb, but these were mostly well-known pain targets with ligands moving into or already in clinical trials. For instance, CALCA (*CGRP*), P2RX3 ($P_2X_3$), NTRK1 (*TrkA*), SCN10A (*Nav1.8*), and SCN11A (*Nav1.9*) are all well-known genes involved in pain, most of which have robust ongoing clinical candidate development. Others in this list included nicotinic receptors that were mentioned above but most of the ligands identified lack specificity for these targets. The HTR3A gene, which encodes the 5-HT3 ligand gated ion channel, is enriched in hDRG and many ligands of this receptor subtype are known. Despite a strong case for the role of this receptor in chronic pain states,[49] the pain-relieving potential of antagonists of this receptor has never been clarified in humans, although 1 small clinical trial in neuropathic pain suggests efficacy of antagonists.[62]

Only 1 gene in this list clearly stood out as novel, QRFPR, which encodes the pyroglutamylated RFamide peptide receptor. Interestingly, although the QRFPR receptor has not been widely studied in the context of pain or sensory biology, the RFamide peptides also interact with the Mas-related G-protein-coupled receptor (GPCR) family of receptors[36] suggesting unappreciated potential complexity in the sensory responses to these peptides in the DRG through their cognate receptor. We also note that although these known drug/receptor interactions mentioned above are of potential interest for drug repurposing, they do not encompass potential transcriptional changes in the hDRG that would likely accompany nerve injury leading to neuropathic pain and therefore represent a theoretically incomplete list. However,

we are also not aware of a single study that has evaluated transcriptome-wide changes in the DRG of human pain patients.

The short list of known ligand to gene interactions identified from the DRG-enriched data set suggests that many DRG-specific genes have not been explored as drug targets. We thus characterized several pharmacologically important gene families including kinases, phosphatases, ion channels, GPCRs, neuropeptides and associated receptors, and CAMs. We also considered that drug targets can include drugs that do not cross the BBB and we therefore included hDRG-expressed genes that had strong neuronal expression but did not show expression in nonneuronal tissues. These genes may be targets for drugs that do not cross the BBB, if DRG targeting is the desired outcome.

Among ion channels, we identified several ion channels with well-known DRG-specific expression and roles but also identified several others (**Fig. 5A**), such as nicotinic receptor subunits α6 and α9 (mentioned above) and the potassium channel KCNH6. A much larger set of ion channel genes showed enrichment in neuronal tissues with preserved expression in hDRG, and in many cases also in mDRG. Among these were many additional potassium channel genes and strong expression for kainate receptor subtypes, in particular in the hDRG. Several neuropeptide and neuropeptide receptors have been targets of new drugs developed in the past decade.[33] Some candidate drugs targeting these molecules have not been studied because of low bioavailability (because of BBB and other factors),[23] which is not necessarily a challenge when targeting the DRG. NMB, CALCA, CALCB, and GAL were identified to be hDRG enriched, as was the receptor QRFPR, which was mentioned above. GLRA4, which is a glycine receptor subunit, was expressed in hDRG but not mDRG (**Fig. 5B**). Cell adhesion molecules are also increasingly viewed as pharmacological targets because of their involvement in processes such as inflammation and cell signaling.[50] We identified several neuronally enriched CAMs

such as CNTN1, NLGN3, NRCAM, and LRRC4C, but none of these were specific to hDRG (**Fig. 5C**). Dorsal root ganglia–enriched MPZ and CLDN19 are expressed in Schwann cells, but have gene products that are present at neuron–Schwann cell junctions.

Among GPCRs, the Mas-related GPCR family (including known DRG-specific genes in mouse and human) showed DRG enrichment (as expected) as did several other GPCRs including QRFPR, CCKAR, and F2RL2 (**Fig. 6A**), whereas AGTR2 is notably low in both hDRG and mDRG. We also identified several DRG-enriched (OR7E101P and OR51E2) or pan-neurally enriched (OR2L13 and OR7A5) olfactory receptors. Although we cannot assess if they are enriched in human olfactory epithelium because of a lack of transcriptome data for this tissue, we find that the mouse orthologs for OR7A5 and OR51E2 are not expressed in the mouse olfactory epithelium. Although the role of olfactory receptors in nonolfactory tissue is under debate, some have been shown to be upregulated in nerve injury models,[1] suggesting that these receptors may have functions outside their canonical tissue. A broader list of GPCRs were expressed in hDRG and also found in other CNS tissues. Interestingly, these included a large number of orphan GPCRs, suggesting an unmined pharmacological resource for sensory and pain research. Finally, we assessed kinases (**Fig. 6B**) and phosphatases (**Fig. 6C**). Notable among the kinases were NTRK1 and RET, receptor tyrosine kinases for NGF and GDNF, respectively, both of which are well-known to be DRG enriched in adult human and mouse. Another weakly enriched receptor tyrosine kinase in hDRG was INSRR, an insulin receptor–related receptor. All other kinases we identified were expressed in hDRG but also in other human CNS regions. Several neuron-enriched phosphatases were also detected, including tyrosine phosphatases PTPRN, PTPRN2, and PTPRZ1, but none of these showed a clear enrichment in hDRG.

To delve deeper into identifying DRG-enriched genes as putative therapeutic targets, we looked at experimentally validated protein–protein interactions in StringDB.[27] We limited this search to interactions within and between the set of DRG-enriched genes and the gene families with drug-targeting potential, in a bid to identify pathways with strong therapeutic potential. Based on this network (Supplementary Fig. 6, available online at http://links.lww.com/PAIN/A559) and intersectional analysis with the KEGG pathway database,[46] we identify several well-understood pathways whose components are partially
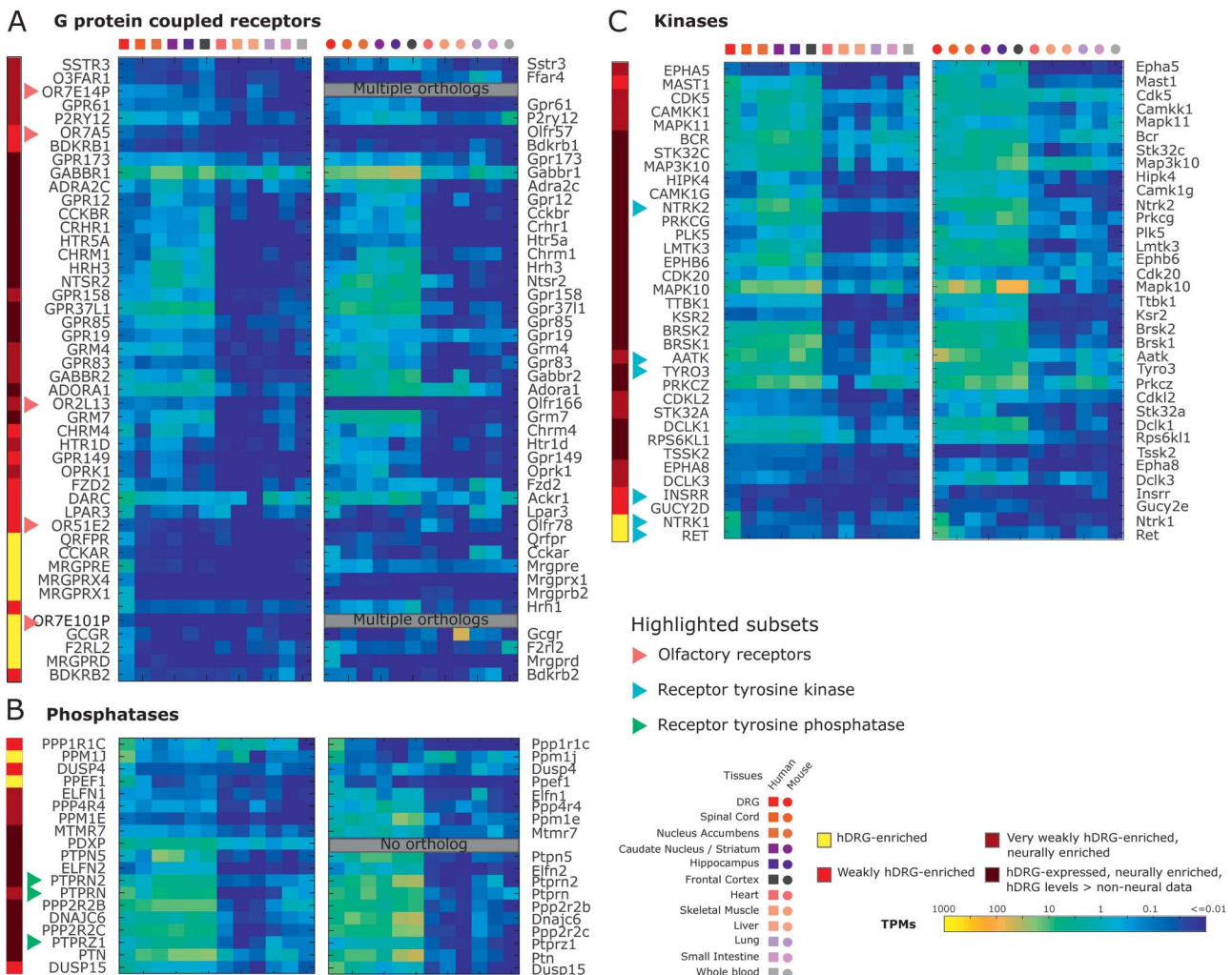


**Figure 6.** Human DRG–enriched genes for pharmacological targets: GPCRs, phosphatases, and kinases. Orthologous human and mouse gene expression in the corresponding tissues is depicted as heatmaps. Differentially expressed GPCRs (A), phosphatases (B), and kinases (C) show several candidate genes, including several receptor tyrosine kinases and phosphatases, and olfactory receptors for pharmacological targeting. DRG, dorsal root ganglia; GPCR, G-protein-coupled receptor; hDRG, human DRG.

identified, including the neurotrophin signaling, cAMP signaling, retrograde endocannabinoid signaling, and axon-guidance pathways. We detailed the neurotrophin signaling pathway (**Fig. 7**), with focus on signaling through the NGF receptor because NGF-targeted therapies are in clinical trials for pain but have an uncertain future because of rare, severe side effects.[5] We find that most gene level abundances for the corresponding proteins in this signaling pathway are present in the hDRG and some are either hDRG enriched (NTRK1 and NGFR), or weakly hDRG enriched (MAPK11, SHC2, ARHGDIA, and ARHGDIG). These NGF signaling pathway components may be drug targets to manipulate this pathway in the hDRG with relative specificity.

### 3.5. Putative axonal localization of mRNAs in human dorsal root ganglia sensory neurons

mRNA may be specifically compartmentalized inside somatic or axonal compartments of very long pseudounipolar sensory neurons. Gene expression changes in proximal vs distal portions of peripheral nerves in nerve-injured rats[43] have shown a clear signal of differential gene expression, and abundant evidence

links axonally localized mRNAs in mouse DRG neurons to axon growth, regeneration, and nociceptive plasticity.[45,64,65,68,119] Gene expression in peripheral nerves and the DRG in humans have only been contrasted with respect to peripheral neuropathies,[97] and putative localization of mRNAs to the axons of hDRG neurons has not been examined in general. We analyzed the tibial nerve RNA-seq experiments from the GTex consortium,[17] contrasting it with our hDRG RNA-seq profiles and rodent scRNA-seq data sets of DRG neurons[112] and sensory neuron axonal compartment mRNA profiles,[35] to identify transcripts that might be potentially axonally transported in hDRG sensory neurons.

Transcription factor protein sequences typically contain NLS sequences that promote retrograde transport to the nucleus. It has thus been hypothesized that anterograde transport of mRNA to axons, stimulus-induced local translation, and retrograde transport of protein to the nucleus are a potential mechanism for growth cone, or nerve ending to nucleus signaling[42] and similar signaling may promote nociceptor sensitization.[65] In the mDRG, several proteins have been shown to be locally translated in the axonal compartments of sensory neurons. These include the
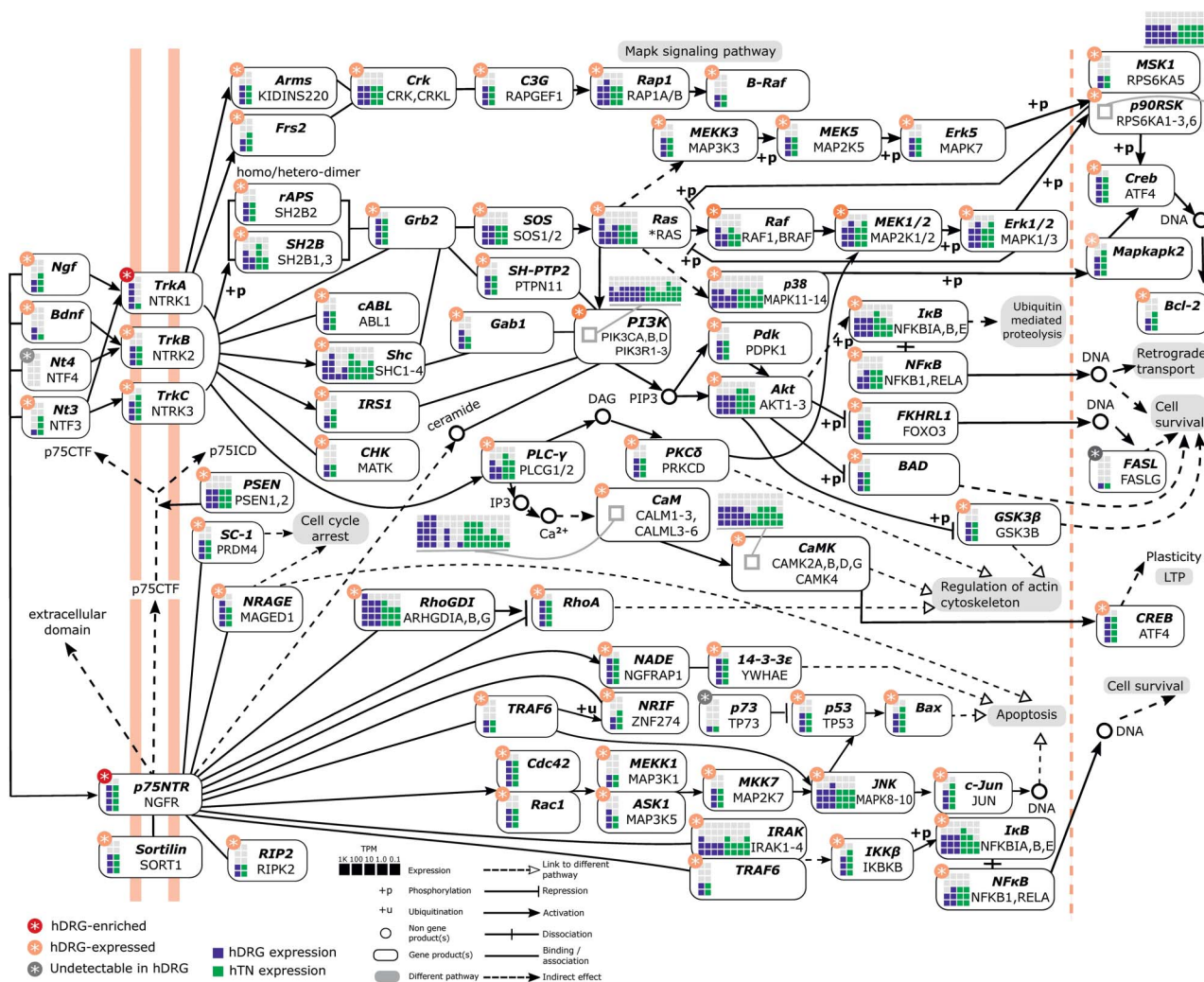


**Figure 7.** Enrichment of NGF/TrkA signaling components in hDRG. The NGF neurotrophin signaling pathway, based on the KEGG database, showing hDRG and human tibial nerve (hTN) expression and hDRG enrichment for members of the pathway. Several signaling molecules in this pathway are expressed or enriched in the hDRG compared with other tissues analyzed in this study (figure based on KEGG visualization with permission from Kanehisa Laboratories who retain copyright). KEGG protein group IDs are in italicized boldface, with associated genes written below the ID. If the associated gene name is lexically identical to the protein group ID, then it is not shown. hDRG, human dorsal root ganglia; TPM, transcripts per million.

enzyme *Ranbp1*, the TFs *Stat3* and *Creb*, filament protein *Vim*, and nuclear transport protein *Importin-β*.[7,65,92] Although some of these have been shown to be axonally translated under specific circumstances such as axon growth or repair, under normal conditions local translation may not take place, but baseline amounts of mRNA would still be detectable in the axon. Gumy et al.[35] have previously shown that in uninjured axons of sensory neurons sourced from rats, mRNA from inflammatory and immune response genes is present. Therefore, we looked for expression of the corresponding genes (Ranbp1, Stat3, Creb1, Vim, and Kpnb1) in the mDRG scRNA-seq data set and find that these are expressed in subpopulations of these neurons (**Fig. 8A**). We further find that orthologs are expressed in both the hDRG and hTN (**Fig. 8A**), suggesting that these mRNAs are potentially axonally transported in humans.

To expand our analysis more broadly, we decided to examine genes known to be axonally transported in healthy mammalian DRG. mRNA in axonal compartments of sensory neurons has been profiled in genome-wide fashion in both mouse[68] and rats,[35] and meta-analysis of these data sets has also been performed.[47] Gumy et al.[35] generated microarray profiles of the axonal compartment in adult rat axons from DRG explants that contain the largest gene set. From this, we identified 2261 human genes that have an annotated, unambiguous ortholog with these rat genes (gene list presented on our web site). These genes' expression levels were queried in hDRG and hTN and found to be highly expressed in both hDRG and hTN with many genes being expressed in the 10–100 TPM range, as shown in the distribution of gene abundances (**Fig. 8B**). This analysis supports the conclusion
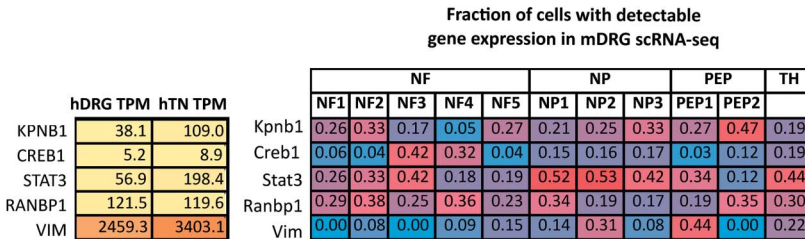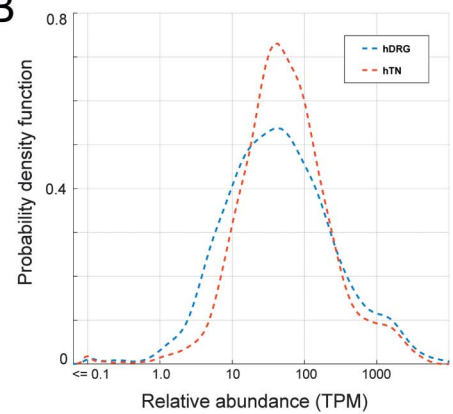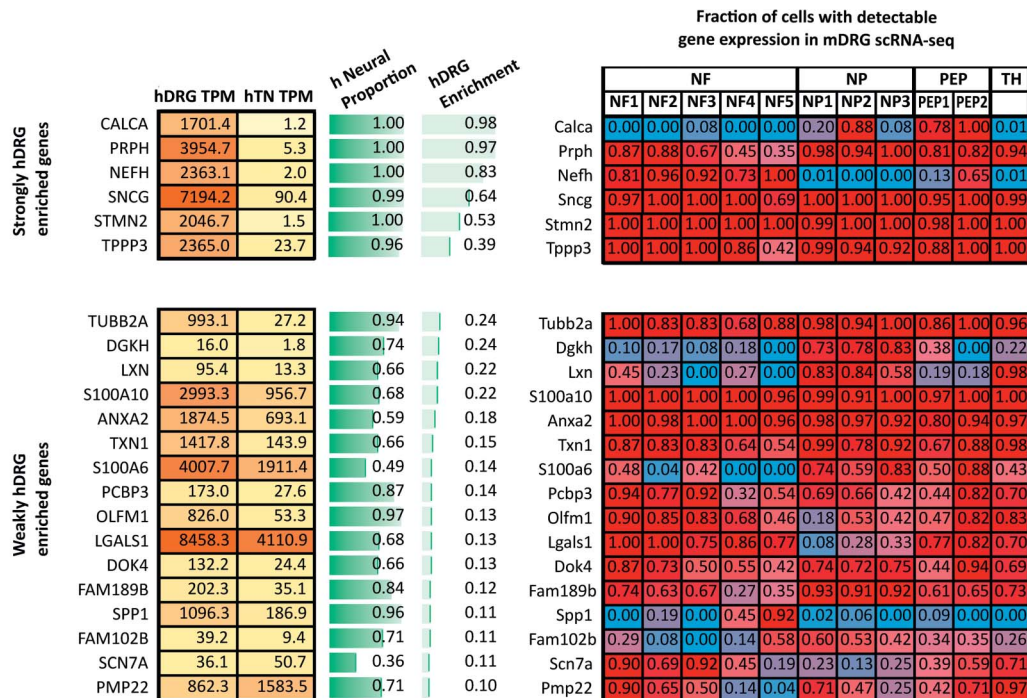


**Figure 8.** Identification of axonally transported mRNA in the human tibial nerve (hTN) transcriptome. (A) Several key genes that have been shown in the literature to be locally translated in the axon and retrograde transported show gene expression in hDRG and hTN samples, and also show expression in an mDRG neuronal subpopulation. (B) A large majority of genes with mRNAs that were detected to be axonally transported in rat DRG neurons have human orthologs that are expressed (>1.0 TPM) in both the hDRG and hTN, as shown based on the estimated probability density of the relative abundances (in TPM) in the respective tissues. (C) We identified several strongly and weakly enriched hDRG genes that are also expressed in hTN, and in mDRG neuronal subpopulations. Their potential for expression in nonneuronal cells is shown as the h neural proportion score. DRG, dorsal root ganglia; hDRG, human DRG; hTN, human tibial nerve; mDRG, mouse DRG; TPM, transcripts per million.

that axonally transported mRNAs are conserved from rat to human.

It is obvious that some portions of the hTN gene expression are driven by Schwann cell transcriptomes, so we additionally looked at the set of genes that were previously identified to be hDRG enriched (**Fig. 4**), but not putatively Schwann cell expressed (Supplementary Tables 3 and 4, available online at http://links. lww.com/PAIN/A559) to see whether any of these genes were detectable in the hTN. Although most of these genes were not identified as axonally transported in the rat study, we find that several of these genes are detectable in the hTN transcriptome (**Fig. 8C**). The detection of these transcripts shows that an additional subset of hDRG transcripts that are putatively transported into the axonal compartment. Based on our findings, we also looked across all 414 GTEx tibial nerve samples and identified sensory neuronal genes such as PRPH and CALCA, and neuronal markers such as RBFOX2 to be expressed at abundances >10 TPM.

In addition, we identified multiple weakly hDRG-enriched genes (hDRG-enrichment score >0.10) to be also detectable in the hTN transcriptome. Although some of these genes (such as TUBB2A) are known to be expressed in the Schwann cell transcriptome, we show that all these genes are also expressed in mDRG neurons based on the mDRG scRNA-seq data set, and thus the expression in the tibial nerve may be contributed by both neuronal and glial cells (**Fig. 8C**). The potentially sizeable overlap of the transcriptome between neurons and glia in the PNS requires further study at the single-cell level to deconvolute transcriptome contributions from different cell types.

We found that membrane protein PIRT (regulatory subunit of TRPV1)[48,107] is localized only to the hDRG with no detectable expression in the hTN. In addition, we noticed several sensory neuron-specific genes such as MAS-related GPCRs MRGPRX1, MRGPRX3, and MRGPRE have low or undetectable gene expression levels in hTN, but are detected at high levels in the hDRG, suggesting that these mRNAs are not axonally transported. These make them good candidates for neuronal subpopulation level cell body biomarkers for mRNA in situ hybridization studies. The fold change of all human genes between the hDRG and hTN is detailed in the human RNA-seq table on our web site.

## 4. Discussion

Our analysis profiled the hDRG transcriptome, and contrasted it with other human tissue transcriptomes and their corresponding mouse transcriptomes. By contrasting gene expression patterns across tissues in humans, we created comprehensive gene coexpression modules that take into account gene expression in DRG. For each coexpression module, we identified putative transcription regulatory genes (TFs and cofactors). For the coexpression module with genes specifically expressed in the hDRG (and silenced in other analyzed tissues), we identified 13 DRG-specific TFs. We confirmed known mammalian sensory development-related TFs in DRG such as DRGX, TLX3, and PRDM12 in the adult hDRG, and additionally identified several TFs such as POU4F1, POU4F3, HOXD1, and SKOR2. We comprehensively catalogued hDRG-expressed members of gene families known to be functionally important in the DRG: ion channels, receptors, kinases, and RNA-binding proteins.

One important finding of our work is that most DRG-enriched genes in the commonly used laboratory mouse strain, C57BL/6, are accurately predicted by hDRG specificity suggesting strong translational capacity for the mouse model. With the exception of

HMX1 and NOTO, all DRG-enriched TFs have conserved enrichment, which strongly suggests purifying selection, and supports the validity of the mouse as a preclinical model species for sensory biology and pain pharmacological and genetic research. Empirically, this is borne out as hDRG and mDRG whole transcriptome profiles cluster together in our analysis, unlike CNS transcriptomes (**Fig. 1B**).

There are several factors that distinguish our work from previous studies that used high-throughput transcriptome analyses. Our goal was to design an analysis framework to contrast the steady state mRNA profile of different human and mouse tissues, with a focus on identifying gene expression patterns across tissues that can be relevant for understanding DRG biology and guide pain therapeutic discovery. We focused on unbiased gene set enrichment analysis in DRG-specific gene expression patterns for functional analysis, in contrast to Flegel et al.[26] who comprehensively characterized expression patterns of gene families known to be important in sensory tissues a priori. Flegel et al. also pooled data from individual DRG donors in vitro, whereas we separately quantify transcriptome profiles from 3 individual DRG donors, as well as pooling them in silico where required, such that both measures of dispersion and central tendency can be calculated on estimated transcript abundances for further statistical analyses. Sapio et al.[97] contrasted human and mouse DRG gene expression and performed functional analysis of DRG-enriched genes in the context of sensory neuropathies, but we additionally contrast mouse RNA-seq data from a homologous set of tissues to our human tissue panel, thereby building a framework to contrast human and mouse evolutionary divergence of expression patterns. Like the Sapio et al. study, we identify constituent cell types where DRG-expressed genes are most likely to be transcribed. The most comprehensive study inclusive of the DRG for cross-tissue transcriptome comparisons between human and mouse was performed by the Genomics Institute of the Novartis Research Foundation in 2004 using standard Affymetrix microarrays.[104] To address these gaps and to generate our own resource for the field, we profiled the mRNA abundances in the hDRG by performing RNA-seq and contrasted it with relevant human and mouse RNA-seq data sets from public repositories, primarily the GTEx Consortium.[17] Our quantification is based on a genome-wide readout of transcript abundance not limited to a predefined set of probes, has the added benefit of low technical variability with respect to microarray studies,[60] and unlike the Novartis study is analyzed from the perspective of DRG biology, and drug target discovery in the PNS.

In addition to this, we performed several analyses that characterize the hDRG transcriptome in several biologically relevant ways for the first time. Over and above identifying hDRG-specific genes, gene coexpression analysis across human tissues allowed us to identify gene modules with similar coexpression patterns that are putatively coregulated, thus taking the first step in identifying the transcriptional regulatory networks underlying fundamental cell types in the hDRG. Identification of hDRG-enriched TFs allowed us to further identify TFs that may be implicated in controlling such regulatory networks into adulthood. Notably, our result divides the well-studied mammalian pan-CNS–expressed genes[19] into CNS genes that are downregulated in the hDRG and those that are expressed in both the CNS and in the hDRG (such as DOCK3, which had been known to be expressed pan-CNS in mouse[75]). Further studies analyzing single-cell transcriptome profiles and performing TF-binding motif analyses are required for clarifying our understanding of these transcriptional regulatory networks.

Based on our comprehensively catalogued and annotated hDRG mRNA profile, we aim to perform several follow-up studies. Because we specifically sourced L2 lumbar DRG from female donors, we aim to contrast mRNA profiles between both male and female DRG donors, to identify sex-specific differences in DRG gene expression. We aim to identify whether there are region-specific differences in the DRG transcriptome by analyzing DRG from additional spinal levels and that innervate different target tissues. Finally, we aim to perform RNA-seq on cohorts of patients with pain to identify both patient-specific and cohort-specific transcriptomic signatures of chronic pain, if they exist, in the hDRG. The absence of these data sets in the present analysis can be viewed as a weakness of this study, but, to the best of our knowledge, these data sets do not exist or are not publicly available. All our sequencing data are publicly available and can be searched without the need for additional analysis on our resource webpage. This resource can be built on by interested researchers.

An interesting finding of our study is the presence of neuronally expressed genes in the hTN transcriptome, some of which are DRG enriched and neuronally expressed (such as PRPH and CALCA), and others which have been shown to be axonally transported or locally translated in mammalian sensory neurons. We interpret this to mean that RNAs generated in DRG neuron nuclei are potentially trafficked into peripherally projecting axons, likely through binding to RNA-binding proteins that have previously been shown to be abundant in DRG axons.[88,89] This finding opens the possibility of gaining insight into changes in DRG transcriptomes in humans by sampling from peripheral nerves (eg, tibial nerve biopsies) or even from skin biopsies that are routinely taken for evaluation of epidermal nerve fiber densities in patients with neuropathic pain. By comparing transcriptomes of peripheral tissues such as skin biopsies with nerve endings, or nerve biopsies, we aim to eventually be able to build up a "personal genomics" model of the pain transcriptome[113] that can give molecular neurobiological insight into phenotype changes in DRG neurons that drive chronic pain. Insofar as local translation at nociceptor terminals has already been identified as a key regulation step in nociceptor plasticity,[63–65] this insight could develop into a powerful tool for mechanism-based discovery in patients. Additional work is needed to confirm this finding, but our identification of specific transcripts that are likely to be axonally transported will facilitate this work. For instance, our work points to a distinct subset of mRNAs that can be used in in situ hybridization experiments on human nerve biopsies to demonstrate that these transcripts are indeed localized to distal axons.

In summary, we have described the transcriptional landscape of the hDRG in comparison with a large number of other tissues and comprehensively categorized similarities and differences to the most commonly used preclinical pain research model, the laboratory mouse. The work elucidates a number of previously unknown facets of sensory biology (eg, the continuous adult expression of TFs known to be involved in nociceptor development) and identifies new potential drug targets based on unbiased transcriptomic enrichment analysis of the hDRG.

## Conflict of interest statement

The authors have no conflict of interest to declare.

## Acknowledgements

## Appendix A. Supplemental digital content

Supplemental digital content associated with this article can be found online at http://links.lww.com/PAIN/A559.

## References

[1] Abaffy T. Human olfactory receptors expression and their role in non-olfactory tissues-a mini-review. J Pharmacogenomics Pharmacoproteomics 2015;6:1.

[2] Ambjørn M, Dubreuil V, Miozzo F, Nigon F, Møller B, Issazadeh-Navikas S, Berg J, Lees M, Sap J. A loss-of-function screen for phosphatases that regulate neurite outgrowth identifies PTPN12 as a negative regulator of TrkB tyrosine phosphorylation. PLoS One 2013;8:e65371.

[3] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000;25:25–9.

[4] Atluri S, Sudarshan G, Manchikanti L. Assessment of the trends in medical use and misuse of opioid analgesics from 2004 to 2011. Pain Physician 2014;17:E119–128.

[5] Bannwarth B, Kostine M. Targeting nerve growth factor (NGF) for pain management: what does the future hold for NGF antagonists? Drugs 2014;74:619–26.

[6] Bar-Joseph Z, Gifford DK, Jaakkola TS. Fast optimal leaf ordering for hierarchical clustering. Bioinformatics 2001;17(suppl 1):S22–9.

[7] Ben-Yaakov K, Dagan SY, Segal-Ruder Y, Shalem O, Vuppalanchi D, Willis DE, Yudin D, Rishal I, Rother F, Bader M. Axonal transcription factors signal retrogradely in lesioned peripheral nerve. EMBO J 2012;31:1350–63.

[8] Blanchard JW, Eade KT, Szűcs A, Sardo VL, Tsunemoto RK, Williams D, Sanna PP, Baldwin KK. Selective conversion of fibroblasts into peripheral sensory neurons. Nat Neurosci 2015;18:25–35.

[9] Bolstad BM, Irizarry RA, Åstrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 2003;19:185–93.

[10] Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. Bioinformatics 2010;26:2363–7.

[11] Chambers SM, Qi Y, Mica Y, Lee G, Zhang XJ, Niu L, Bilsland J, Cao L, Stevens E, Whiting P. Combined small-molecule inhibition accelerates developmental timing and converts human pluripotent stem cells into nociceptors. Nat Biotechnol 2012;30:715–20.

[12] Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics 2013;14:128.

[13] Chen YC, Auer-Grumbach M, Matsukawa S, Zitzelsberger M, Themistocleous AC, Strom TM, Samara C, Moore AW, Cho LT, Young GT, Weiss C, Schabhuttl M, Stucka R, Schmid AB, Parman Y, Graul-Neumann L, Heinritz W, Passarge E, Watson RM, Hertz JM, Moog U, Baumgartner M, Valente EM, Pereira D, Restrepo CM, Katona I, Dusl M, Stendel C, Wieland T, Stafford F, Reimann F, von Au K, Finke C, Willems PJ, Nahorski MS, Shaikh SS, Carvalho OP, Nicholas AK, Karbani G, McAleer MA, Cilio MR, McHugh JC, Murphy SM, Irvine AD, Jensen UB, Windhager R, Weis J, Bergmann C, Rautenstrauss B, Baets J, De Jonghe P, Reilly MM, Kropatsch R, Kurth I, Chrast R, Michiue T, Bennett DL, Woods CG, Senderek J. Transcriptional regulator PRDM12 is essential for human pain perception. Nat Genet 2015;47:803–8.

[14] Chinwalla AT, Cook LL, Delehaunty KD, Fewell GA, Fulton LA, Fulton RS, Graves TA, Hillier LW, Mardis ER, McPherson JD. Initial sequencing and comparative analysis of the mouse genome. Nature 2002;420:520–62.

[15] Chiu IM, Barrett LB, Williams EK, Strochlic DE, Lee S, Weyer AD, Lou S, Bryman GS, Roberson DP, Ghasemlou N, Piccoli C, Ahat E, Wang V, Cobos EJ, Stucky CL, Ma Q, Liberles SD, Woolf CJ. Transcriptional profiling at whole population and single cell levels reveals somatosensory neuron molecular diversity. Elife 2014:3.

[16] Consortium GO. Gene Ontology annotations and resources. Nucleic Acids Res 2013;41:D530–5.

[17] Consortium GT. The genotype-tissue expression (GTEx) project. Nat Genet 2013;45:580–5.

[18] Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, Caudy M, Garapati P, Gillespie M, Kamdar MR, Jassal B, Jupe S, Matthews L, May B, Palatnik S, Rothfels K, Shamovsky V, Song H, Williams M, Birney E, Hermjakob H, Stein L, D'Eustachio P. The Reactome pathway knowledgebase. Nucleic Acids Res 2014;42:D472–477.

[19] Darmanis S, Sloan SA, Zhang Y, Enge M, Caneda C, Shuer LM, Gephart MGH, Barres BA, Quake SR. A survey of human brain transcriptome diversity at the single cell level. Proc Natl Acad Sci U S A 2015;112:7285–90.

[20] Dart RC, Surratt HL, Cicero TJ, Parrino MW, Severtson SG, Bucher-Bartelson B, Green JL. Trends in opioid analgesic abuse and mortality in the United States. N Engl J Med 2015;372:241–8.

[21] Dawes JM, Antunes-Martins A, Perkins JR, Paterson KJ, Sisignano M, Schmid R, Rust W, Hildebrandt T, Geisslinger G, Orengo C, Bennett DL, McMahon SB. Genome-wide transcriptional profiling of skin and dorsal root ganglia after ultraviolet-B-induced inflammation. PLoS One 2014;9: e93338.

[22] Dreszer TR, Karolchik D, Zweig AS, Hinrichs AS, Raney BJ, Kuhn RM, Meyer LR, Wong M, Sloan CA, Rosenbloom KR, Roe G, Rhead B, Pohl A, Malladi VS, Li CH, Learned K, Kirkup V, Hsu F, Harte RA, Guruvadoo L, Goldman M, Giardine BM, Fujita PA, Diekhans M, Cline MS, Clawson H, Barber GP, Haussler D, James Kent W. The UCSC Genome Browser database: extensions and updates 2011. Nucleic Acids Res 2012;40: D918–923.

[23] Egleton RD, Davis TP. Development of neuropeptide drugs that cross the blood-brain barrier. NeuroRx 2005;2:44–53.

[24] Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95:14863–8.

[25] Felfly H, Muotri A, Yao H, Haddad GG. Hematopoietic stem cell transplantation protects mice from lethal stroke. Exp Neurol 2010;225:284–93.

[26] Flegel C, Schobel N, Altmuller J, Becker C, Tannapfel A, Hatt H, Gisselmann G. RNA-seq analysis of human trigeminal and dorsal root ganglia with a focus on chemoreceptors. PLoS One 2015;10:e0128951.

[27] Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C, Jensen LJ. STRING v9. 1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Research 2013;41:D808–D815.

[28] Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, Liu C, Shi W, Bryant SH. The NCBI BioSystems database. Nucleic Acids Res 2010; 38:D492–496.

[29] Gerhold KA, Pellegrino M, Tsunozaki M, Morita T, Leitch DB, Tsuruda PR, Brem RB, Catania KC, Bautista DM. The star-nosed mole reveals clues to the molecular basis of mammalian touch. PLoS One 2013;8: e55001.

[30] Glusman G, Caballero J, Robinson M, Kutlu B, Hood L. Optimal scaling of digital transcriptomes. PLoS One 2013;8:e77885.

[31] Goswami SC, Mishra SK, Maric D, Kaszas K, Gonnella GL, Clokie SJ, Kominsky HD, Gross JR, Keller JM, Mannes AJ, Hoon MA, Iadarola MJ. Molecular signatures of mouse TRPV1-lineage neurons revealed by RNA-Seq transcriptome analysis. J Pain 2014;15:1338–59.

[32] Gray KA, Seal RL, Tweedie S, Wright MW, Bruford EA. A review of the new HGNC gene family resource. Hum Genomics 2016;10:6.

[33] Griebel G, Holsboer F. Neuropeptide receptor ligands as drugs for psychiatric diseases: the end of the beginning? Nat Rev Drug Discov 2012;11:462–78.

[34] Guan Z, Kuhn JA, Wang X, Colquitt B, Solorzano C, Vaman S, Guan AK, Evans-Reinsch Z, Braz J, Devor M, Abboud-Werner SL, Lanier LL, Lomvardas S, Basbaum AI. Injured sensory neuron-derived CSF1 induces microglial proliferation and DAP12-dependent pain. Nat Neurosci 2016;19:94–101.

[35] Gumy LF, Yeo GS, Tung YCL, Zivraj KH, Willis D, Coppola G, Lam BY, Twiss JL, Holt CE, Fawcett JW. Transcriptome analysis of embryonic and adult sensory axons reveals changes in mRNA repertoire localization. RNA 2011;17:85–98.

[36] Han SK, Dong X, Hwang JI, Zylka MJ, Anderson DJ, Simon MI. Orphan G protein-coupled receptors MrgA1 and MrgC11 are distinctively activated by RF-amide-related peptides through the Gαq/11 pathway. Proc Natl Acad Sci U S A 2002;99:14740–5.

[37] Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, Barnes I, Bignell A, Boychenko V, Hunt T, Kay M, Mukherjee S, Rajan J, Despacio-Reyes G, Saunders G, Steward C, Harte R, Lin M, Howald C, Tanzer A, Derrien T, Chrast J, Walters N, Balasubramanian S, Pei B, Tress M, Rodriguez JM, Ezkurdia I, van Baren J, Brent M, Haussler D, Kellis M, Valencia A, Reymond A, Gerstein M, Guigo R, Hubbard TJ. GENCODE: the reference human genome annotation for the ENCODE Project. Genome Res 2012;22:1760–74.

[38] Hill MO. Diversity and evenness: a unifying notation and its consequences. Ecology 1973;54:427–32.

[39] Hu G, Huang K, Hu Y, Du G, Xue Z, Zhu X, Fan G. Single-cell RNA-seq reveals distinct injury responses in different types of DRG sensory neurons. Sci Rep 2016;6:31851.

[40] Inquimbert P, Bartels K, Babaniyi OB, Barrett LB, Tegeder I, Scholz J. Peripheral nerve injury produces a sustained shift in the balance between glutamate release and uptake in the dorsal horn of the spinal cord. PAIN 2012;153:2422–31.

[41] Isensee J, Wenzel C, Buschow R, Weissmann R, Kuss AW, Hucho T. Subgroup-elimination transcriptomics identifies signaling proteins that define subclasses of TRPV1-positive neurons and a novel paracrine circuit. PLoS One 2014;9:e115731.

[42] Ji SJ, Jaffrey SR. Axonal transcription factors: novel regulators of growth cone-to-nucleus signaling. Dev Neurobiol 2014;74:245–58.

[43] Jiang N, Li H, Sun Y, Yin D, Zhao Q, Cui S, Yao D. Differential gene expression in proximal and distal nerve segments of rats with sciatic nerve injury during Wallerian degeneration. Neural Regen Res 2014;9: 1186–94.

[44] Jiang Y, Clark WT, Friedberg I, Radivojac P. The impact of incomplete knowledge on the evaluation of protein function prediction: a structured-output learning perspective. Bioinformatics 2014;30:i609–16.

[45] Jung H, Yoon BC, Holt CE. Axonal mRNA localization and local protein synthesis in nervous system assembly, maintenance and repair. Nat Rev Neurosci 2012;13:308–24.

[46] Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 2016;44:D457–62.

[47] Kar AN, Lee SJ, Twiss JL. Expanding axonal transcriptome brings new functions for axonally synthesized proteins in health and disease. Neuroscientist 2018;24:111–29.

[48] Kim AY, Tang Z, Liu Q, Patel KN, Maag D, Geng Y, Dong X. Pirt, a phosphoinositide-binding protein, functions as a regulatory subunit of TRPV1. Cell 2008;133:475–85.

[49] Kim YS, Chu Y, Han L, Li M, Li Z, LaVinka PC, Sun S, Tang Z, Park K, Caterina MJ. Central terminal sensitization of TRPV1 by descending serotonergic facilitation modulates chronic pain. Neuron 2014;81: 873–87.

[50] Kiryushko D, Bock E, Berezin V. Pharmacology of cell adhesion molecules of the nervous system. Curr Neuropharmacol 2007;5: 253–67.

[51] Kogelman LJA, Christensen RE, Pedersen SH, Bertalan M, Hansen TF, Jansen-Olesen I, Olesen J. Whole transcriptome expression of trigeminal ganglia compared to dorsal root ganglia in Rattus Norvegicus. Neuroscience 2017;350:169–79.

[52] Kuja-Panula J, Kiiltomäki M, Yamashiro T, Rouhiainen A, Rauvala H. AMIGO, a transmembrane protein implicated in axon tract development, defines a novel protein family with leucine-rich repeats. J Cell Biol 2003; 160:963–73.

[53] Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW, Ma'ayan A. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016;44:W90–97.

[54] Kupczyk P, Holysz M, Gajda M, Reich A, Kobuszewska A, Nowakowska B, Szepietowski J. Expression of UCHL1 gene in the skin of Psoriasis: Neuroepidermal marker of itch. Proceedings of 31st Conference of Polish Dermatological Society, Wroclaw, Poland, May 2016.

[55] Kutuzov MA, Solov'eva OV, Andreeva AV, Bennett N. Protein Ser/Thr phosphatases PPEF interact with calmodulin. Biochem Biophys Res Commun 2002;293:1047–52.

[56] Li CL, Li KC, Wu D, Chen Y, Luo H, Zhao JR, Wang SS, Sun MM, Lu YJ, Zhong YQ, Hu XY, Hou R, Zhou BB, Bao L, Xiao HS, Zhang X. Somatosensory neuron types identified by high-coverage single-cell RNA-sequencing and functional heterogeneity. Cell Res 2016;26: 83–102.

[57] Liu Q, Dong X. The role of the Mrgpr receptor family in itch. Pharmacology of Itch. Alan Cowan, Gil Yosipovitch (eds). Berlin-Heidelberg, Germany: Springer, 2015. p. 71–88.

[58] Lopes C, Liu Z, Xu Y, Ma Q. Tlx3 and Runx1 act in combination to coordinate the development of a cohort of nociceptors, thermoceptors, and pruriceptors. J Neurosci 2012;32:9706–15.

[59] Manteniotis S, Lehmann R, Flegel C, Vogel F, Hofreuter A, Schreiner BS, Altmuller J, Becker C, Schobel N, Hatt H, Gisselmann G. Comprehensive RNA-Seq expression analysis of sensory ganglia with a focus on ion channels and GPCRs in trigeminal ganglia. PLoS One 2013;8:e79523.

[60] Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res 2008;18:1509–17.

[61] Masahira N, Takebayashi H, Ono K, Watanabe K, Ding L, Furusho M, Ogawa Y, Nabeshima Y, Alvarez-Buylla A, Shimizu K, Ikenaka K. Olig2-positive progenitors in the embryonic spinal cord give rise not only to motoneurons and oligodendrocytes, but also to a subset of astrocytes and ependymal cells. Dev Biol 2006;293:358–69.

[62] McCleane GJ, Suzuki R, Dickenson AH. Does a single intravenous injection of the 5HT3 receptor antagonist ondansetron have an analgesic effect in neuropathic pain? A double-blinded, placebo-controlled cross-over study. Anesth Analg 2003;97:1474–8.

[63] Melemedjian OK, Asiedu MN, Tillu DV, Peebles KA, Yan J, Ertz N, Dussor GO, Price TJ. IL-6- and NGF-induced rapid control of protein synthesis and nociceptive plasticity via convergent signaling to the eIF4F complex. J Neurosci 2010;30:15113–23.

[64] Melemedjian OK, Asiedu MN, Tillu DV, Sanoja R, Yan J, Lark A, Khoutorsky A, Johnson J, Peebles KA, Lepow T, Sonenberg N, Dussor G, Price TJ. Targeting adenosine monophosphate-activated protein kinase (AMPK) in preclinical models reveals a potential mechanism for the treatment of neuropathic pain. Mol Pain 2011;7:70.

[65] Melemedjian OK, Tillu DV, Moy JK, Asiedu MN, Mandell EK, Ghosh S, Dussor G, Price TJ. Local translation and retrograde axonal transport of CREB regulates IL-6-induced nociceptive plasticity. Mol Pain 2014;10:45.

[66] Miller JA, Horvath S, Geschwind DH. Divergence of human and mouse brain transcriptome highlights Alzheimer disease pathways. Proc Natl Acad Sci U S A 2010;107:12698–703.

[67] Miller SJ, Jessen WJ, Mehta T, Hardiman A, Sites E, Kaiser S, Jegga AG, Li H, Upadhyaya M, Giovannini M, Muir D, Wallace MR, Lopez E, Serra E, Nielsen GP, Lazaro C, Stemmer-Rachamimov A, Page G, Aronow BJ, Ratner N. Integrative genomic analyses of neurofibromatosis tumours identify SOX9 as a biomarker and survival gene. EMBO Mol Med 2009;1: 236–48.

[68] Minis A, Dahary D, Manor O, Leshkowitz D, Pilpel Y, Yaron A. Subcellular transcriptomics-dissection of the mRNA composition in the axonal compartment of sensory neurons. Dev Neurobiol 2014;74:365–81.

[69] Mishra SK, Holzman S, Hoon MA. A nociceptive signaling role for neuromedin B. J Neurosci 2012;32:8686–95.

[70] Molyneaux BJ, Arlotta P, Menezes JR, Macklis JD. Neuronal subtype specification in the cerebral cortex. Nat Rev Neurosci 2007;8:427–37.

[71] Momin A, Wood JN. Sensory neuron voltage-gated sodium channels as analgesic drug targets. Curr Opin Neurobiol 2008;18:383–8.

[72] Nadjar Y, Triller A, Bessereau JL, Dumoulin A. The Susd2 protein regulates neurite growth and excitatory synaptic density in hippocampal cultures. Mol Cell Neurosci 2015;65:82–91.

[73] Nagy V, Cole T, Van Campenhout C, Khoung TM, Leung C, Vermeiren S, Novatchkova M, Wenzel D, Cikes D, Polyansky AA, Kozieradzki I, Meixner A, Bellefroid EJ, Neely GG, Penninger JM. The evolutionarily conserved transcription factor PRDM12 controls sensory neuron development and pain perception. Cell Cycle 2015;14:1799–808.

[74] Nakatani T, Minaki Y, Kumai M, Nitta C, Ono Y. The c-Ski family member and transcriptional regulator Corl2/Skor2 promotes early differentiation of cerebellar Purkinje cells. Dev Biol 2014;388:68–80.

[75] Namekata K, Enokido Y, Iwasawa K, Kimura H. MOCA induces membrane spreading by activating Rac1. J Biol Chem 2004;279: 14331–7.

[76] Nascimento R, Santiago M, Marques S, Allodi S, Martinez A. Diversity among satellite glial cells in dorsal root ganglia of the rat. Braz J Med Biol Res 2008;41:1011–17.

[77] Niederreither K, Vermot J, Schuhbaur B, Chambon P, Dolle P. Retinoic acid synthesis and hindbrain patterning in the mouse embryo. Development 2000;127:75–85.

[78] Nomaksteinsky M, Kassabov S, Chettouh Z, Stoekle HC, Bonnaud L, Fortin G, Kandel ER, Brunet JF. Ancient origin of somatic and visceral neurons. BMC Biol 2013;11:53.

[79] Okaty BW, Sugino K, Nelson SB. Cell type-specific transcriptomics in the brain. J Neurosci 2011;31:6939–43.

[80] Oldham MC, Horvath S, Geschwind DH. Conservation and evolution of gene coexpression networks in human and chimpanzee brains. Proc Natl Acad Sci U S A 2006;103:17973–8.

[81] Oort PJ, Warden CH, Baumann TK, Knotts TA, Adams SH. Characterization of Tusc5, an adipocyte gene co-expressed in peripheral neurons. Mol Cell Endocrinol 2007;276:24–35.

[82] Owens ND, Blitz IL, Lane MA, Patrushev I, Overton JD, Gilchrist MJ, Cho KW, Khokha MK. Measuring absolute RNA copy numbers at high temporal resolution reveals transcriptome kinetics in development. Cell Rep 2016;14:632–47.

[83] Pachter L. Models for transcript quantification from RNA-Seq. 2011. arXiv preprint arXiv:1104.3889v2.

[84] Pacini A, Micheli L, Maresca M, Branca JJV, McIntosh JM, Ghelardini C, Mannelli LDC. The α9α10 nicotinic receptor antagonist α-conotoxin RgIA prevents neuropathic pain induced by oxaliplatin treatment. Exp Neurol 2016;282:37–48.

[85] Pearson K. Note on regression and inheritance in the case of two parents. Proc R Soc Lond 1895;58:240–2.

[86] Perkins JR, Antunes-Martins A, Calvo M, Grist J, Rust W, Schmid R, Hildebrandt T, Kohl M, Orengo C, McMahon SB. A comparison of RNA-seq and exon arrays for whole genome transcription profiling of the L5 spinal nerve transection model of neuropathic pain in the rat. Mol pain 2014;10:7.

[87] Pevny LH, Sockanathan S, Placzek M, Lovell-Badge R. A role for SOX1 in neural determination. Development 1998;125:1967–78.

[88] Price TJ, Hargreaves KM, Cervero F. Protein expression and mRNA cellular distribution of the NKCC1 cotransporter in the dorsal root and trigeminal ganglia of the rat. Brain Res 2006;1112:146–58.

[89] Price TJ, Rashid MH, Millecamps M, Sanoja R, Entrena JM, Cervero F. Decreased nociceptive sensitization in mice lacking the fragile X mental retardation protein: role of mGluR1/5 and mTOR. J Neurosci 2007;27: 13958–67.

[90] Ramskold D, Wang ET, Burge CB, Sandberg R. An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data. PLoS Comput Biol 2009;5:e1000598.

[91] Rebelo S, Chen ZF, Anderson DJ, Lima D. Involvement of DRG11 in the development of the primary afferent nociceptive system. Mol Cell Neurosci 2006;33:236–46.

[92] Rishal I, Fainzilber M. Retrograde signaling in axonal regeneration. Exp Neurol 2010;223:5–10.

[93] Romero IG, Pai AA, Tung J, Gilad Y. RNA-seq: impact of RNA degradation on transcript quantification. BMC Biol 2014;12:42.

[94] Roth RB, Hevezi P, Lee J, Willhite D, Lechner SM, Foster AC, Zlotnik A. Gene expression analyses reveal molecular relationships among 20 regions of the human CNS. Neurogenetics 2006;7:67–80.

[95] Rouwette T, Sondermann J, Avenali L, Gomez-Varela D, Schmidt M. Standardized profiling of the membrane-enriched proteome of mouse dorsal root ganglia (DRG) provides novel insights into chronic pain. Mol Cell Proteomics 2016;15:2152–68.

[96] Roux J, Rosikiewicz M, Robinson-Rechavi M. What to compare and how: comparative transcriptomics for Evo-Devo. J Exp Zool B Mol Dev Evol 2015;324:372–82.

[97] Sapio MR, Goswami SC, Gross JR, Mannes AJ, Iadarola MJ. Transcriptomic analyses of genes and tissues in inherited sensory neuropathies. Exp Neurol 2016;283:375–95.

[98] Scholz J, Woolf CJ. Can we conquer pain? Nat Neurosci 2002;5(suppl): 1062–7.

[99] Schwertassek U, Buckley DA, Xu CF, Lindsay AJ, McCaffrey MW, Neubert TA, Tonks NK. Myristoylation of the dual-specificity phosphatase c-JUN N-terminal kinase (JNK) stimulatory phosphatase 1 is necessary for its activation of JNK signaling and apoptosis. FEBS J 2010;277:2463–73.

[100] Seok J, Warren HS, Cuenca AG, Mindrinos MN, Baker HV, Xu W, Richards DR, McDonald-Smith GP, Gao H, Hennessy L, Finnerty CC, Lopez CM, Honari S, Moore EE, Minei JP, Cuschieri J, Bankey PE, Johnson JL, Sperry J, Nathens AB, Billiar TR, West MA, Jeschke MG, Klein MB, Gamelli RL, Gibran NS, Brownstein BH, Miller-Graziano C, Calvano SE, Mason PH, Cobb JP, Rahme LG, Lowry SF, Maier RV, Moldawer LL, Herndon DN, Davis RW, Xiao W, Tompkins RG. Inflammation, host response to injury LSCRP. Genomic responses in mouse models poorly mimic human inflammatory diseases. Proc Natl Acad Sci U S A 2013;110:3507–12.

[101] Shannon CE. Communication in the presence of noise. Proc IRE 1949; 37:10–21.

[102] Shiroguchi K, Jia TZ, Sims PA, Xie XS. Digital RNA sequencing minimizes sequence-dependent bias and amplification noise with optimized single-molecule barcodes. Proc Natl Acad Sci U S A 2012; 109:1347–52.

[103] Short KM, Cox TC. Subclassification of the RBCC/TRIM superfamily reveals a novel motif necessary for microtubule binding. J Biol Chem 2006;281:8970–80.

[104] Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, Zhang J, Soden R, Hayakawa M, Kreiman G, Cooke MP, Walker JR, Hogenesch JB. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A 2004;101:6062–7.

[105] Sudmant PH, Alexis MS, Burge CB. Meta-analysis of RNA-seq expression data across species, tissues and studies. Genome Biol 2015;16:287.

[106] Takao K, Miyakawa T. Genomic responses in mouse models greatly mimic human inflammatory diseases. Proc Natl Acad Sci U S A 2015; 112:1167–72.

[107] Tang Z, Kim A, Masuch T, Park K, Weng H, Wetzel C, Dong X. Pirt functions as an endogenous regulator of TRPM8. Nat Commun 2013;4: 2179.

[108] Thakur M, Crow M, Richards N, Davey GI, Levine E, Kelleher JH, Agley CC, Denk F, Harridge SD, McMahon SB. Defining the nociceptor transcriptome. Front Mol Neurosci 2014;7:87.

[109] Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol 2013;31:46–53.

[110] Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 2009;25:1105–11.

[111] Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 2012;7:562–78.

[112] Usoskin D, Furlan A, Islam S, Abdo H, Lönnerberg P, Lou D, Hjerling-Leffler J, Haeggström J, Kharchenko O, Kharchenko PV. Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. Nat Neurosci 2015;18:145–53.

[113] Vellucci R. Heterogeneity of chronic pain. Clin Drug Investig 2012;32 (suppl 1):3–10.

[114] Vilella AJ, Severin J, Ureta-Vidal A, Heng L, Durbin R, Birney E. EnsemblCompara GeneTrees: complete, duplication-aware phylogenetic trees in vertebrates. Genome Res 2009;19:327–35.

[115] Wagner AH, Coffman AC, Ainscough BJ, Spies NC, Skidmore ZL, Campbell KM, Krysiak K, Pan D, McMichael JF, Eldred JM, Walker JR, Wilson RK, Mardis ER, Griffith M, Griffith OL. DGIdb 2.0: mining clinically relevant drug-gene interactions. Nucleic Acids Res 2016;44:D1036–1044.

[116] Wainger BJ, Buttermore ED, Oliveira JT, Mellin C, Lee S, Saber WA, Wang AJ, Ichida JK, Chiu IM, Barrett L, Huebner EA, Bilgin C, Tsujimoto N, Brenneis C, Kapur K, Rubin LL, Eggan K, Woolf CJ. Modeling pain in vitro using nociceptor neurons reprogrammed from fibroblasts. Nat Neurosci 2015;18:17–24.

[117] Whittington N, Cunningham D, Le TK, De Maria D, Silva EM. Sox21 regulates the progression of neuronal differentiation in a dose-dependent manner. Dev Biol 2015;397:237–47.

[118] Wieskopf JS, Mathur J, Limapichat W, Post MR, Al-Qazzaz M, Sorge RE, Martin LJ, Zaykin DV, Smith SB, Freitas K. The nicotinic α6 subunit gene determines variability in chronic pain sensitivity via cross-inhibition of P2X2/3 receptors. Sci Transl Med 2015;7:287ra272.

[119] Willis DE, Twiss JL. Profiling axonal mRNA transport. Methods Mol Biol 2011;714:335–52.

[120] Wu S, Marie Lutz B, Miao X, Liang L, Mo K, Chang YJ, Du P, Soteropoulos P, Tian B, Kaufman AG, Bekker A, Hu Y, Tao YX. Dorsal root ganglion transcriptome analysis following peripheral nerve injury in mice. Mol Pain 2016;12.

[121] Xie W, Schultz MD, Lister R, Hou Z, Rajagopal N, Ray P, Whitaker JW, Tian S, Hawkins RD, Leung D. Epigenomic analysis of multilineage differentiation of human embryonic stem cells. Cell 2013;153:1134–48.

[122] Yaguchi H, Okumura F, Takahashi H, Kano T, Kameda H, Uchigashima M, Tanaka S, Watanabe M, Sasaki H, Hatakeyama S. TRIM67 protein negatively regulates Ras activity through degradation of 80K-H and induces neuritogenesis. J Biol Chem 2012;287:12050–9.

[123] Yin K, Deuis JR, Lewis RJ, Vetter I. Transcriptomic and behavioural characterisation of a mouse model of burn pain identify the cholecystokinin 2 receptor as an analgesic target. Mol Pain 2016;12.

[124] You L, Wu J, Feng Y, Fu Y, Guo Y, Long L, Zhang H, Luan Y, Tian P, Chen L. APASdb: a database describing alternative poly (A) sites and selection of heterogeneous cleavage sites downstream of poly (A) signals. Nucleic Acids Res 2014;43:D59–67.

[125] Young-Pearse T, Ivic L, Kriegstein A, Cepko C. Characterization of mice with targeted deletion of glycine receptor alpha 2. Mol Cell Biol 2006;26:5728–34.

[126] Zaitseva M, Vollenhoven BJ, Rogers PA. In vitro culture significantly alters gene expression profiles and reduces differences between myometrial and fibroid smooth muscle cells. Mol Hum Reprod 2006; 12:187–207.