Research Paper

PAIN®

Genetic pathway analysis reveals a major role for extracellular matrix organization in inflammatory and neuropathic pain

Marc Parisien[a], Alexander Samoshkin[a,b], Shannon N. Tansley[a], Marjo H. Piltonen[a], Loren J. Martin[c], Nehme El-Hachem[a], Concetta Dagostino[d,e], Massimo Allegri[f,g], Jeffrey S. Mogil[a,h], Arkady Khoutorsky[a,i], Luda Diatchenko[a,*]

Abstract

Chronic pain is a debilitating and poorly treated condition whose underlying mechanisms are poorly understood. Nerve injury and inflammation cause alterations in gene expression in tissues associated with pain processing, supporting molecular and cellular mechanisms that maintain painful states. However, it is not known whether transcriptome changes can be used to reconstruct a molecular pathophysiology of pain. In the current study, we identify molecular pathways contributing to chronic pain states through the analysis of global changes in the transcriptome of dorsal root ganglia, spinal cord, brain, and blood in mouse assays of nerve injury– and inflammation-induced pain. Comparative analyses of differentially expressed genes identified substantial similarities between 2 animal pain assays and with human low-back pain. Furthermore, the extracellular matrix (ECM) organization has been found the most commonly regulated pathway across all tested tissues in the 2 animal assays. Examination of human genome-wide association study data sets revealed an overrepresentation of differentially expressed genes within the ECM organization pathway in single nucleotide polymorphisms most strongly associated with human back pain. In summary, our comprehensive transcriptomics analysis in mouse and human identified ECM organization as a central molecular pathway in the development of chronic pain.

Keywords: Chronic pain, Next-generation sequencing, Differentially expressed genes, Extracellular matrix, GWAS, Bioinformatics

1. Introduction

Chronic pain is a debilitating condition that serves no obvious biological function, in contrast with acute pain triggered at injury, protecting the organism from further damage by activating withdrawal behaviors. More than 20% of the population worldwide suffer from chronic pain, making it the leading cause of disability in humans.[56] The mechanisms underlying the development of chronic pain are not well understood. Accordingly, available medications for chronic pain target the symptoms of the disease and not the underlying pathology.[66] This treatment strategy results in inadequate pain relief under most circumstances and is accompanied by a high prevalence of side effects.[23,49,69]

Two major types of pain states are inflammatory pain (IP), caused by tissue inflammation or injury, and neuropathic pain (NP), caused by damage to or dysfunction of the nervous system itself. Inflammatory pain and NP share several underlying neuronal mechanisms, including increased excitability of primary afferents, reduced spinal inhibitory tone, and the involvement of higher brain centers. Despite the progress in uncovering neuronal alterations along the pain pathway, the genetic programs and molecular mechanisms supporting these changes are not well known.

Preclinical studies in animal models have shown that pain states caused by inflammation or nerve injury are accompanied by massive changes in gene expression along the pain pathway, including in the dorsal root ganglion (DRG), spinal cord (SC), and supraspinal brain areas.[25,29,30,72] Long-lasting changes in gene expression are believed to underlie the development of a sustained pain state by mediating the molecular and cellular mechanisms causing sensitization of peripheral and central pain circuits and the generation of ectopic discharge, both hallmarks of chronic pain.[17, 43, 47] Complex pathophysiology of chronic pain assumes a model of multigene contribution to the reorganization of the somatosensory system in persistent pain states. Identification of these complex gene interactions and subthreshold contributions to pain phenotypes, as well as more global gene expression patterns, require unique bioinformatics approaches and the generation of high-quality gene expression data sets of different pain models and tissues.

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

M. Parisien and A. Samoshkin contributed equally to this work.

[a] *Alan Edwards Centre for Research on Pain, McGill University, Montréal, QC, Canada,* [b] *School of the Clinical Medicine, University of Cambridge, Cambridge, United Kingdom,* [c] *Department of Psychology, University of Toronto, Mississauga, ON, Canada,* [d] *Department of Medicine and Surgery, University of Parma, Parma, Italy,* [e] *Study In Multidisciplinary Pain Research (SIMPAR), Parma, Italy,* [f] *Italian Pain Group, Milan, Italy,* [g] *Pain Therapy Service, Policlinico Monza Hospital, Monza, Italy, Departments of* [h] *Psychology and,* [i] *Anesthesia, McGill University, Montréal, QC, Canada*

**Corresponding author. Address: Alan Edwards Centre for Research on Pain Genome Building, Room 2201, 740 Dr. Penfield Avenue, Montreal, Quebec, Canada H3A 0G1. Tel.: 514-398-2878; fax: 514-398-8900. E-mail address: luda.diatchenko@mcgill.ca (L. Diatchenko).*

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.painjournalonline.com).

PAIN 160 (2019) 932–944

http://dx.doi.org/10.1097/j.pain.0000000000001471

The current understanding of gene expression signatures in pain states is largely based on studies using a single pain assay and is limited to 1 or 2 tissues within the pain pathway. Recent single-cell profiling or cell type–specific transcriptomic studies provide important insights into changes in distinct cell types under pathological conditions; however, these studies are lacking a relevance to dynamics of pathological changes.[24,51,61] Here, we undertook a more comprehensive approach to study the gene expression landscape in pain states by performing transcriptomic analysis of 4 tissues (DRG, SC, brain, and blood) in 2 mouse pain assays modeling inflammatory or NP. Our deep RNA sequencing (>50 M reads) at tissue-wide levels allowed for detection of lowly expressed genes and quantitation of minute differential gene expression fold changes even at the mixed-cell–type sample. A pathway analysis of differentially expressed genes (DEGs) identified the extracellular matrix (ECM) organization as the most commonly regulated category across pain assays and tissues.

2. Results

2.1. Transcriptomics of mouse pain models

To study the pathophysiology of pain states at the molecular level, we performed next-generation deep sequencing of RNAs extracted from tissues relevant to pain processes in 2 mouse pain assays (**Fig. 1A**). Two well-characterized assays were used: complete Freund's adjuvant (CFA) as an assay of IP and spared nerve injury (SNI) as an assay of NP. The tissues were extracted 3 days after CFA injection and 7 days after SNI. These time points were selected based on behavioral data (Fig. S1, available at http://links.lww.com/PAIN/A710; legend available at http://links.lww.com/PAIN/A724): CFA injection–induced hypersensitivity peaks at day 3 (Fig. S1A, available at http://links.lww.com/PAIN/A710; legend available at http://links.lww.com/PAIN/A724), whereas SNI-induced hypersensitivity develops to its maximal extent at day 7 (Fig. S1B, available at http://links.lww.com/PAIN/A724). From each mouse (n = 3/assay), 4 tissues were extracted after injury: DRG (L3-L5), dorsal horn of the lumbar SC, whole brain, and whole blood. The poly(A) mRNA was purified and sequenced using the Illumina HiSeq2000 sequencing platform. Quality of deep sequencing data was assessed in a 4-fold manner: (1) byRNA fragment mapping rates on the mouse genome (Fig. S2, available at http://links.lww.com/PAIN/A710; legend available at http://links.lww.com/PAIN/A724); (2) by triplicates for within-group gene expression correlations (Fig. S3, available at http://links.lww.com/PAIN/A710; legend available at http://links.lww.com/PAIN/A724); (3) by gene expression clustering (**Fig. 1B**); and (4) by principal component analyses (PCAs) (**Fig. 1C**). All approaches confirmed the high quality of RNA purification (RIN 9.3 ± 0.5, minimum 7.8) and sequencing (mapped reads 77.4 M ± 24.8 M, minimum 45.7 M). Gene expression was visualized in a pileup plot (Fig. S4A, available at http://links.lww.com/PAIN/A710; legend available at http://links.lww.com/PAIN/A724). Differential gene expression was assessed in a volcano plot (Fig. S4B, available at http://links.lww.com/PAIN/A710; legend available at http://links.lww.com/PAIN/A724), where π-values represent *P*-value-weighted fold changes.[11] Data for differential expression of genes for all tissues and assays can be found in Table S1 (available at http://links.lww.com/PAIN/A711; legend available at http://links.lww.com/PAIN/A724), Table S2 (available at http://links.lww.com/PAIN/A712; legend available at http://links.lww.com/PAIN/A724), Table S3 (available at http://links.lww.com/PAIN/A713; legend available at http://links.lww.com/PAIN/A724), Table S4 (available at http://links.lww.com/PAIN/A714; legend available at http://links.lww.com/PAIN/A724), Table S5 (available at http://links.lww.com/PAIN/A715; legend available at http://links.lww.com/PAIN/A724), Table S6 (available at http://links.lww.com/PAIN/A716; legend available at http://links.lww.com/PAIN/A724), Table S7 (available at http://links.lww.com/PAIN/A717; legend available at http://links.lww.com/PAIN/A724), and Table S8 (available at http://links.lww.com/PAIN/A718; legend available at http://links.lww.com/PAIN/A724).

Unsupervised clustering of RNA-Seq samples based on gene expression correlation first distinguished the nervous system tissue samples from the whole blood (**Fig. 1B**). In the nervous

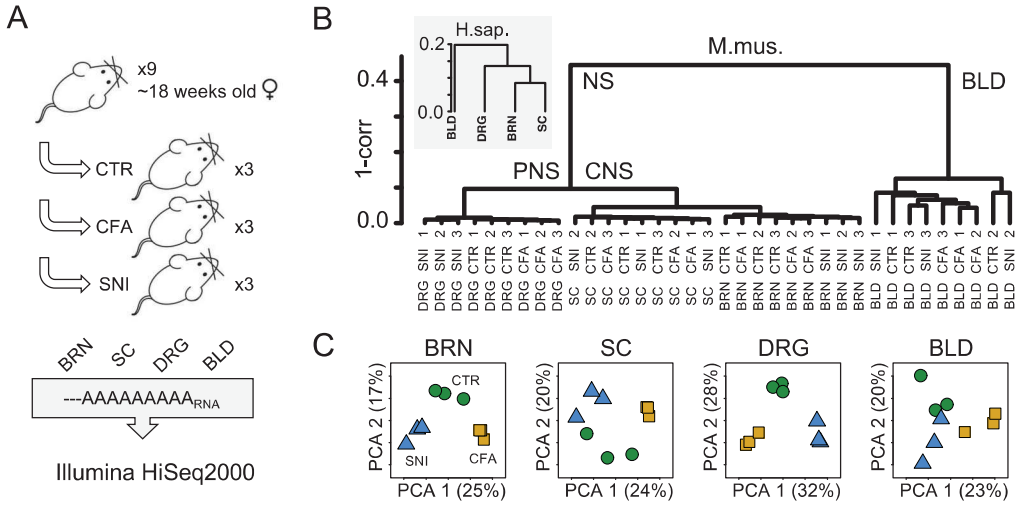

Figure 1. Overview of transcriptomics of mouse pain assays. (A) Experimental workflow. A total of 9 mice were split into 3 groups: control (CTR), treated with complete Freund's adjuvant (CFA), and underwent spared nerve injury (SNI). For each mouse, the following 4 tissues were collected: whole brain (BRN), spinal cord (SC), dorsal root ganglion (DRG), and whole blood (BLD). From each tissue, next-generation sequencing was performed on Illumina HiSeq2000. (B) Unsupervised clustering analyses of mouse transcriptome (M.mus.). The clustering was made based on one minus Pearson's correlation coefficient in gene expression between the samples. The clustering tree shows first branching of the nervous system (NS) tissues from the blood (BLD). Under NS, the clustering tree shows branching of the tissues from the central nervous system (CNS) from the peripheral nervous system (PNS). Similar clustering is observed in humans (H.sap., grayed inset). (C) Tissue-based principal component analyses (PCA) of mouse transcriptome. CTR (green circles), CFA (yellow boxes), SNI (blue triangles).

system branch, the clustering distinguished samples of the central nervous system (CNS: brain and SC) from the peripheral nervous system (PNS: DRG). In the CNS sub-branch, all SC samples clustered together, as did all brain samples. We verified that the first 2 clustering branches, nervous system vs blood, and CNS vs PNS were also present in the gene expression data of the same 4 tissues in humans (**Fig. 1B**, inset; microarray data from Ref. 5). Using the top 50 most DEGs, we performed within-tissue PCAs to assess that the 3 pain states—control, CFA, and SNI—were distinguishable in each tissue (**Fig. 1C**). Brain, SC, and DRG presented clear pain model distinctions with only 2 principal components, whereas blood presented a more complex pain profile, requiring higher dimensionality PCAs for better separation. In all, the first 2 PCAs can explain only 17% of within-tissue gene expression variance in the brain, while up to 32% of the variance was explained in the DRG.

2.2. Replication of published data

Next, we evaluated the replication of DEG patterns observed in our study with DEGs in various tissue-assay pairs from already-published data sets (**Fig. 2**; Table S9, available at http://links.lww.com/PAIN/A719; legend available at http://links.lww.com/PAIN/A724; and Table S10, available at http://links.lww.com/PAIN/A720; legend available at http://links.lww.com/PAIN/A724). We assessed all related genome-wide experiments for which data are available in NCBI's Gene Expression Omnibus and ready for prompt analyses with geo2r.[15,19] Expression fold changes and associated *P*-values were converted into π-values for comparison with our work. We found 3 such data sets (**Fig. 2**). Data set GSE18803[13] compares RNA expression data in SNI vs sham at day 7 after SNI in the SC of rats (**Fig. 2A**). Data set GSE15041[67] compares expression data in SNI vs sham in DRG at day 7 after SNI in rats (**Fig. 2B**). Finally, data set GSE38859[57]

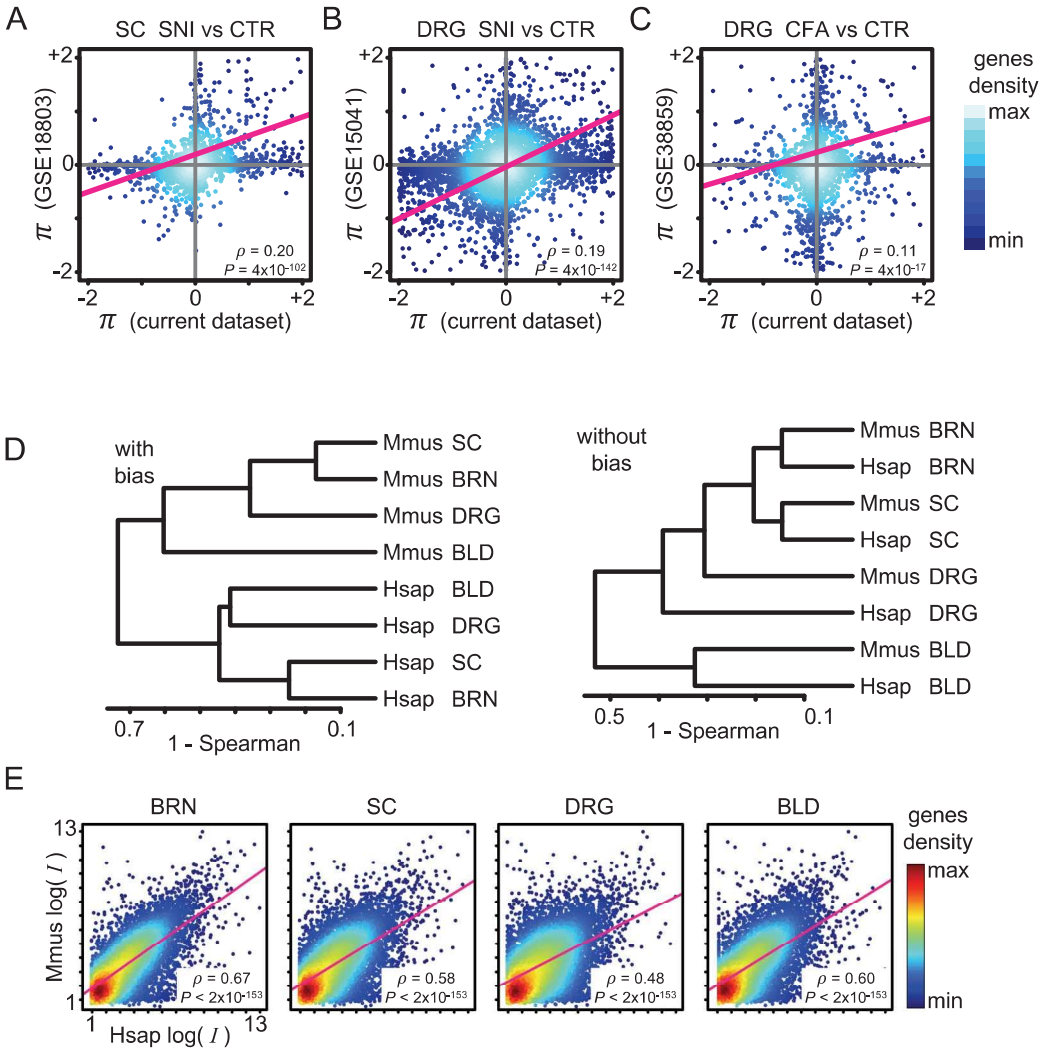

Figure 2. Correlation between the DEGs in current data set with previously published genome-wide rodent studies, and correlation between mouse and human transcriptomes. (A–C) DEG is measured as π-value. Each dot is a gene. Pink line obtained from linear regression. Correlation quantified using rank-based Spearman's coefficient (ρ) with accompanying *P*-value (*P*). Color coding suggests gene density. The corresponding tissues of current data set has been compared with (A) rat spinal cord 7 days SNI vs sham (GEO set GSE18803), (B) rats DRG ipsilateral 7 days SNI vs sham (GEO set GSE15041), and (C) rats DRG 3 days CFA vs sham (GEO set GSE38859). (D) Clustering of mouse (Mmus) and human (Hsap) transcriptomes. The clustering is based on Spearman's correlation coefficient of gene expression between pairs of transcriptomes. Unsupervised clustering leads to first branching by species, suggesting strong influence of species-specific transcriptomes (left panel). Removal of specie-specific transcriptomes leads to pairings by tissue types instead (right panel). (E) Transcriptome gene expression correlation by tissue type between human and CTR mouse. For each tissue, rank-based Spearman's coefficient (ρ) and *P*-value (*P*) are shown. Each dot is a gene. Colour coding suggests gene density. Pink line obtained from linear regression. Mouse RNA-Seq gene expression has been log2 transformed to match human microarray log-based gene expression intensity (I). CFA, complete Freund's adjuvant; DEG, differentially expressed gene; DRG, dorsal root ganglion; SNI, spared nerve injury

compares expression data in CFA vs sham at day 3 in rats (**Fig. 2C**). Because these data sets were compiled using rats, we expected limited replication of DEGs because of interspecies gene expression differences (Table S9, available at http://links.lww.com/PAIN/A719; legend available at http://links.lww.com/PAIN/A724; **Fig. 2**). Furthermore, different quantified platforms were used, ie, RNA-Seq for mouse and microarray for rat. In all 3 genome-wide replication attempts, nevertheless, rank-based Spearman correlation coefficients ranged from 0.11 to 0.20, with accompanying *P*-values between 10^{-17} and 10^{-142}.

Analysis of summary data from a previously published large meta-analysis of DEGs from multiple microarray studies of pain[30] was then used to further quantify replication between our findings and previously published data (Table S10, available at http://links.lww.com/PAIN/A720; legend available at http://links.lww.com/PAIN/A724). The meta-analysis compiled studies that were performed using both mice and rats. It identified a total of 79 genes with significant expression fold change in at least 4 independent microarray experiments. Details of neuropathic and inflammation pain assays were also slightly differing (eg, SNI vs chronic constriction injury). Replication of the surveyed genes yielded a total of 65 committed fold change calls, for which a total of 57 calls agree on the direction of the fold change, thus generating 8 mismatched calls. The *P*-value for this level of success rate, assuming probability to make a correct fold change direction call is one-half (1/2), was evaluated using the hypergeometric test that yielded $P = 3.2 \times 10^{-10}$. Thus, our deep-sequencing data were of high-quality and replicated previous studies.

2.3. Relevance of mouse transcriptome study for humans

We next examined the relevance of using mouse transcriptome data to gain knowledge about humans. A data set containing human gene expression data in the same 4 tissues probed herein[5] allowed for a direct comparison of transcriptomes between control mice and humans, on a per-tissue basis (**Fig. 2**). Unsupervised clustering of the transcriptomes yielded a dendrogram in which the tissues did not cluster by type, but rather by species (**Fig. 2D**, left panel). Removal of species-specific expressed genes and retaining of genes expressed in both species yielded an unsupervised clustering in which tissues clustered by identity (**Fig. 2D**, right panel). In all cases, blood was distinguishable from nervous tissues, and CNS (brain and SC) clustered together apart from PNS (DRG). Comparison of transcriptomes on a per-tissue basis was also performed (**Fig. 2E**). All rank-based correlations of gene expression were statistically significant, with *P*-values below an estimated 2.2×10^{-16}. Spearman correlation coefficients varied from 0.48 in DRG to 0.67 in the brain, indicating strong correspondences between human and mouse transcriptomes, in each tissue. Once again, because the transcriptomes have been quantified using different technologies (RNA-Seq for mouse and microarray for human), it is expected that at least some differences in gene expression arose from the varying accuracies and signal-to-noise ratios of the platforms, notwithstanding species-specific gene expression.

2.4. Relevance of differential gene expression in mouse pain models to humans with clinical condition

Our mouse data were further correlated with a complex human pain state, namely back pain. As it has been shown previously that back pain patients manifest symptoms with both neuropathic and inflammatory components,[20] we compared DEGs in blood of humans with back pain to DEGs in blood in both neuropathic and inflammatory mouse pain models (**Fig. 3**). Very significant correlations were observed for both comparisons; however, in the mouse NP model, this correlation was stronger (**Fig. 3A**; Spearman $\rho = 0.47$, $P = 5 \times 10^{-110}$) than in the mouse inflammation pain model (**Fig. 3B**; Spearman $\rho = 0.33$, $P = 3 \times 10^{-39}$). When genes were grouped by pathways, 80% of differentially expressed pathways in humans displayed correlation coefficients with SNI equal to or greater than 0.35, whereas only 40% of CFA differentially expressed pathways showed correlation coefficients equal or greater than 0.35 (**Fig. 3C**). Thus, we observed both a neuropathic and an inflammatory component at the level of DEGs and differentially expressed pathways in the human back pain profile. However, in this back pain cohort, the neuropathic component was stronger (Kolmogorov–Smirnov $P < 2 \times 10^{-16}$).

2.5. Genome-wide analysis of differentially expressed genes in pain assays

We then established the full list of genes differentially expressed in each of the 4 tissues, and for each of the 2 pain assays (**Fig. 4**; Table S11, available at http://links.lww.com/PAIN/A721; legend available at http://links.lww.com/PAIN/A724). In the CFA assay, the number of DEGs varied from 1123 in the brain to 2642 in SC (**Fig. 4A**). A large proportion of DEGs are tissue-specific, and there were only 48 genes that are differentially expressed in all tissues in the CFA assay. In SNI, the number of DEGs varied from

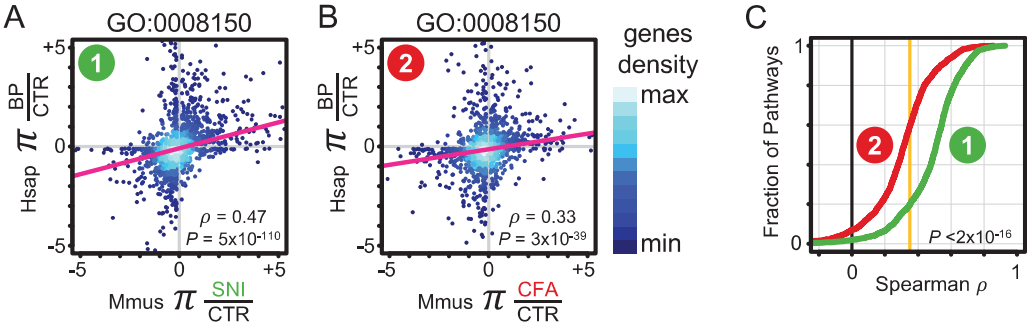

Figure 3. Comparison of differential gene expression in human clinical pain conditions and mouse pain assays. (A and B) Gene expression profile in the blood of human subjects (Hsap) with back pain (BP) were contrasted (π) against those with no pain (CTR). Gene expression profiles in the blood of mice (Mmus) under a pain assay were contrasted (π) against those in the control group (CTR). Each dot is a gene in GO's biological process pathway (GO:0008150). Colour coding reflects gene density. Pink line obtained from linear regression. Rank-based Spearman's correlation coefficient rho (ρ) and associated *P*-value (*P*) are shown. (A) Spared nerve injury pain assays. (B) Complete Freund's adjuvant pain assay. (C) Cumulative fraction of all GO's biological process pathways as a function of their Spearman's ρ. Kolmogorov–Smirnov test *P*-value between the 2 curves is shown. Black vertical line indicates $\rho = 0$, while gold vertical line $\rho = 0.35$. CFA, complete Freund's adjuvant; SNI, spared nerve injury.

968 in brain up to 3194 in DRG (**Fig. 4B**), but only 18 genes were differentially expressed in all 4 tissues.

Differential expression of genes was found conserved between tissues. For CFA, all regression slopes were positive, and percent variance explained (PVE) varied between 1% and 11% (**Fig. 4C**). For SNI, we also found positive regression slopes (except for one), while PVE varied between 0% and 8% (**Fig. 4D**). Most correlated tissue pair in CFA was between DRG and blood (PVE 11%), while in SNI between DRG and SC (PVE 8%).

Furthermore, there was a large overlap in DEGs between CFA and SNI. The percentage of shared DEGs was found to range from 24% in the brain up to 44% in SC (**Fig. 4E**). Unexpectedly, these DEG overlaps within tissues for different pain assays were substantially higher in magnitude than between tissues within one pain model. The large number of shared DEG points to possible common pathways that are activated by inflammation (CFA) and nerve injury (SNI). All tissues considered, as much as 49% of DEGs, were shared between CFA and SNI.

In addition, of the 800 genes already characterized in the context of pain research (see Materials and Methods for source details), a total of 452 known pain genes (56%) of the total 8687 DEGs in our assays overlapped (**Fig. 4F**; Table S12, available at http://links.lww.com/PAIN/A722; legend available at http://links.lww.com/PAIN/A724). This represents a 3.1-fold overrepresentation of observed known differentially expressed pain genes over what would be expected by chance alone, with associated enrichment *P*-value of 2×10^{-129}.

2.6. Tissue-specific pathways involved in distinct pain processes

We next developed a star plot that tracks differential expression in each mouse pain assay, for each gene, in a tissue-based manner

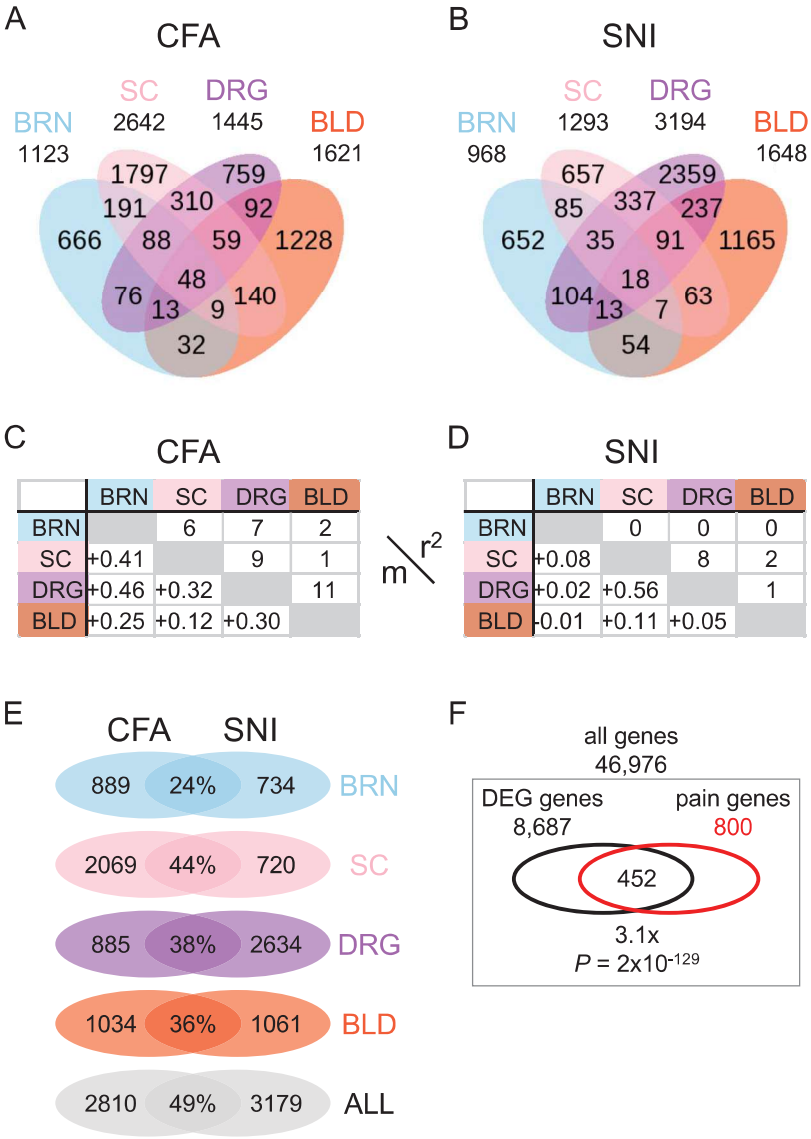

Figure 4. Differentially expressed genes in pain assays and in pain-relevant tissues. (A and B) Four-way Venn diagram showing shared DEGs between tissues in CFA (A) or SNI (B) pain assays. (C and D) Linear regression of DEGs for all tissue pairs. Shown are percent variance explained (Pearson's r^2) rounded to nearest integer (upper right corner), and regression's slope (m) with slope's sign emphasized (lower left corner), for CFA (C) or SNI (D) pain assays. (E) Two-way Venn diagrams showing shared DEGs between CFA and SNI pain assays, for each tissue, and all tissues combined (ALL). Counts at intersection expressed as percentage of the smallest values between the 2 tissue-specific DEGs counts. (F) Two-way Venn diagram showing overlap of DEGs in any tissue or pain mouse assay vs those known to contribute to pain states. Shown at the bottom is observed to expected enrichment ratio (x) of number of DEG among known pain genes, with accompanying *P*-values. CFA, complete Freund's adjuvant; DEGs, differentially expressed genes; SNI, spared nerve injury

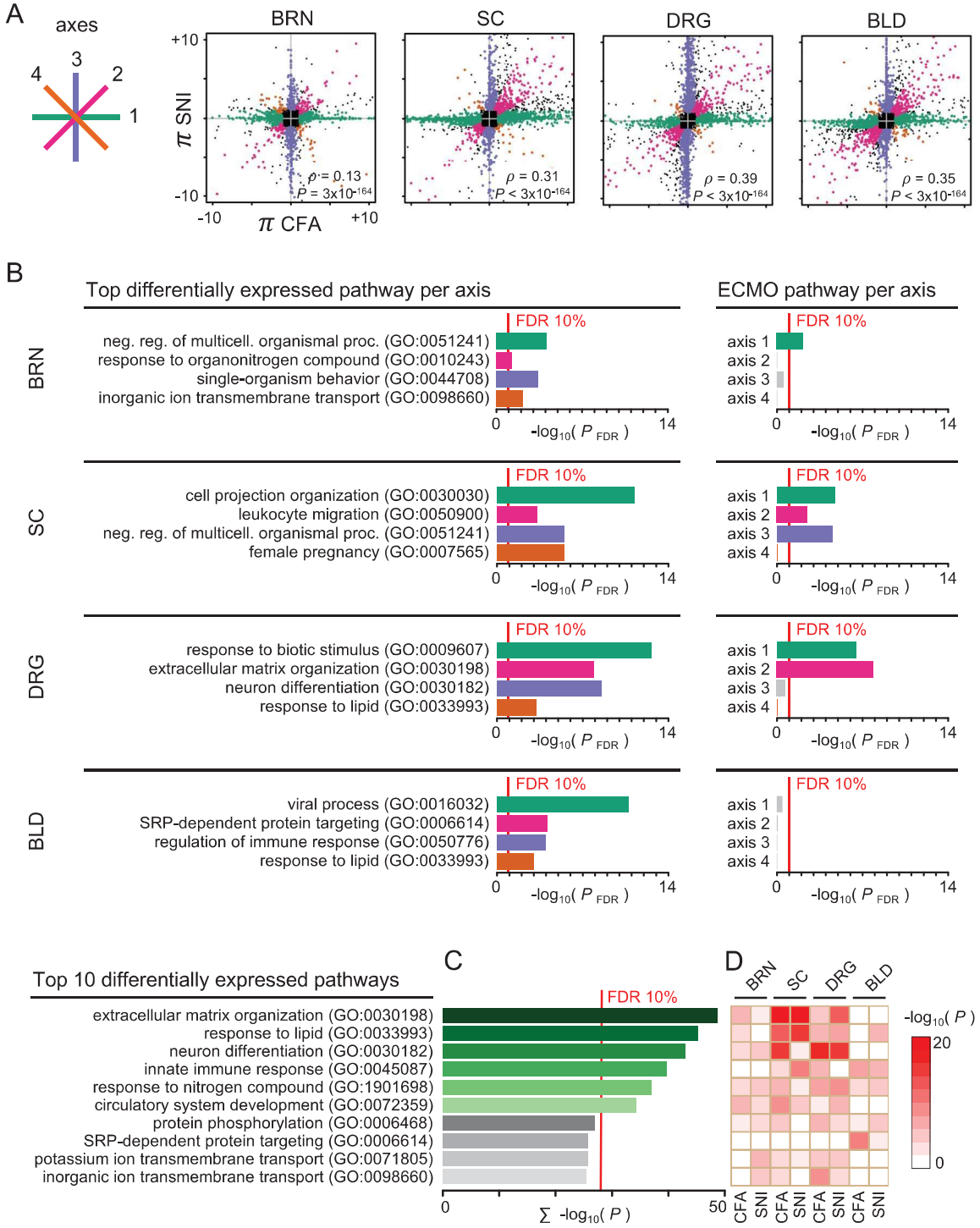

Figure 5. Tissue-specific pathways contributing to distinct pain processes. (A) Star plots depict DEGs between CFA and SNI pain assays, in a tissue-specific manner. Each dot is a gene. Four axes are defined and colored accordingly: CFA-only DEGs (axis 1, green), DEGs correlated between CFA and SNI models (axis 2, pink), SNI-only DEGs (axis 3, purple), and DEGs anticorrelated between CFA and SNI models (axis 4, orange). (B) For each tissue's axes, the pathway with the best FDR-corrected GSEA *P*-value is shown in a bar plot (left). Vertical red lines mark statistical significance at FDR 10% level. Also shown is the ECM organization pathway (GO:0030198, right). (C) Bar plot showing top 10 differentially expressed pathways, summed over all axes of all tissues. Vertical red lines mark statistical significance at FDR 10% level. (D) Heatmap of contribution of each tissue (top) and pain models (bottom) to the top 10 differentially expressed pathways. CFA, complete Freund's adjuvant; DEG, differentially expressed gene; ECM, extracellular matrix; GESA, gene set enrichment analyses; SNI, spared nerve injury.

(**Fig. 5**). We defined each star plot by the 4 axes representing the correlation between DEGs in SNI and CFA assays (**Fig. 5A**; and Fig. S5, available at http://links.lww.com/PAIN/A710; legend available at http://links.lww.com/PAIN/A724). The star plot helped highlight genes that were coregulated between CFA and SNI, again accentuating shared pathophysiology; rank-based Spearman correlations varied from 0.13 in the brain to 0.39 in DRG. The lists of genes identified from the star plots allowed for gene set enrichment analyses (GSEA) (Table S13, available at http://links.lww.com/PAIN/A723; legend available at http://links.lww.com/PAIN/A724). The pathway that showed the most significant enrichment *P*-value is reported for each of the 4 axes, in each tissue-based star plot (**Fig. 5B**).

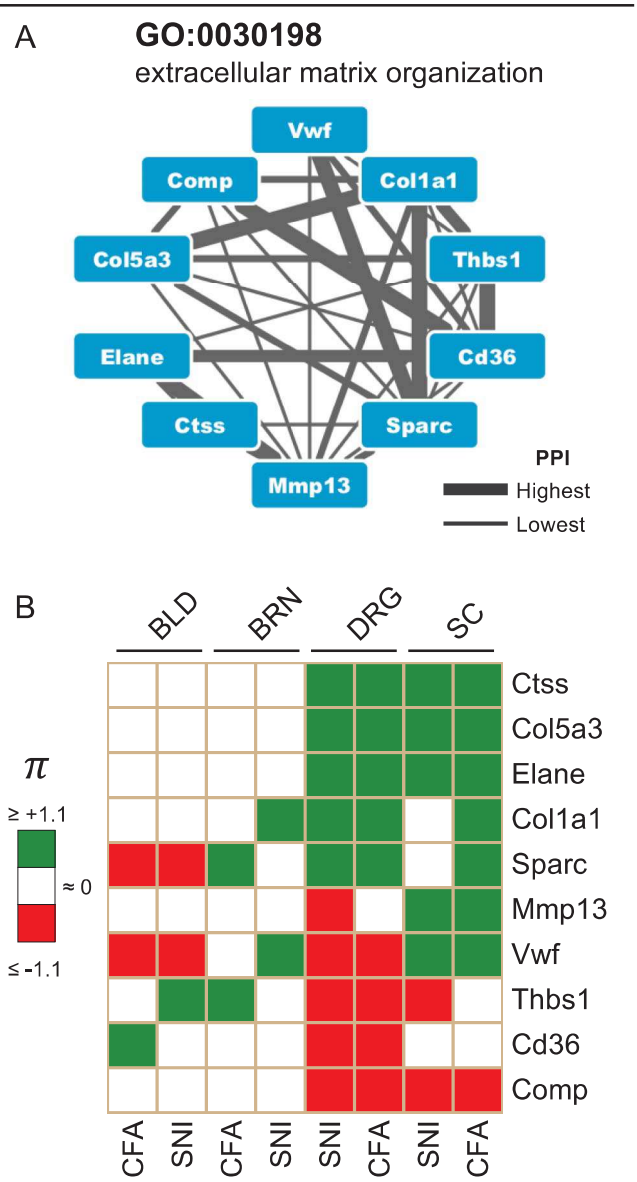

Figure 6. Protein–protein interaction network of the ECM organization pathway. (A) Top 10 genes most differentially expressed in the ECM organization pathway, organized in a protein–protein interaction (PPI) network. Edge confidence, from highest to lowest, is proportional to experimental evidence for interacting protein pair as provided in the STRING database. Genes are CD36 molecule (thrombospondin receptor)—*Cd36*; collagen, type I, alpha 1—*Col1a1*; collagen, type V, alpha 3—*Col5a3*; cartilage oligomeric matrix protein—*Comp*; cathepsin S—*Ctss*; elastase, neutrophil expressed—*Elane*; matrix metallopeptidase 13 (collagenase 3)—*Mmp13*; secreted protein, acidic, cysteine-rich (osteonectin)—*Sparc*; thrombospondin 1—*Thbs1*; and von Willebrand factor—*Vwf*. (B) Heatmap of expression fold change of the top 10 genes across all tissues (top) and pain models (bottom). Expression fold change is measured in π-value units and dichotomized: increased expression compared with control is in green, while decreased is in red. ECM, extracellular matrix

Further analyses for pathway enrichment yielded a list of top 10 differentially expressed pathways, summed over all axes of all tissues (**Fig. 5C**). The ECM organization pathway emerged as the top differentially expressed pathway (GO:0030198). The ECM organization pathway is a biological process that manages the assembly, maintenance, and disassembly of the ECM. We find this pathway to be enriched in 6 different axes (of 16) (**Fig. 5B**): in brain in CFA; in SC, axes 1, 2, and 3; and in DRG, axes 1 and 2. Importantly, when ECM organization seemed significantly enriched in 2 different axes of the same tissue, 2 nonoverlapping subsets of genes of these axes contribute to it (Table S13, available at http://links.lww.com/PAIN/A723; legend available at http://links.lww.com/PAIN/A724).

The top 10 differentially expressed pathways were then re-examined for their impact on tissues and pain assays (**Fig. 5D**). Tissues of the nervous system (brain, SC, and DRG) showed the greatest contributions from the ECM organization pathway, in both CFA and SNI.

2.7. Gene network in the extracellular matrix organization pathway

We next identified the top 10 most DEGs in the ECM organization pathway from their impact, defined as the sum of $|\pi|$ values from all assays and all tissues of differential expression (**Fig. 6**). Proteins encoded by these genes were found to be in a highly-connected network of protein–protein interactions (**Fig. 6A**, STRING database[59]). The proposed network is quality-controlled, built on experimental evidence, text mining, and knowledge transfer between species for orthologous pairs.[70] In the network, all proteins displayed moderate-to-strong evidence of interaction to at least one other protein in the network. We found differential expression of these genes to be tissue- and assay-specific, although it could reflect timing of their differential expression (**Fig. 6B**). Dorsal root ganglion and SC tissues displayed the most differential expression of these genes. Clustering of the genes suggested that *Thbs1*, *Cd36*, and *Comp* were mostly down-regulated, whereas the other 7 proteins were mostly upregulated. We observed that the brain featured the least number of DEGs, whereas the highest number of DEGs was in the DRG. Interestingly, a search in the GeneRIF database[40] for co-occurrences of the gene's name and "pain" revealed that many of the ECM organization genes in the network are already characterized in the context of pain research: *Sparc*,[39,60] *Col1a1*,[1,33] *Mmp13*,[64] *Ctss*,[8,10,11,34,76,77] *Elane*,[68,78] and *Comp*.[16,28]

2.8. Human genome-wide association study results on back pain are enriched for genes in extracellular matrix organization pathway

We next examined the correlation of our mouse DEGs data with genome-wide association study (GWAS) pain data sets. We tested for overrepresentation of single nucleotide polymorphisms (SNPs) from GWAS data sets within human orthologues of mouse DEGs in the set of SNPs with strongest association with back pain (**Fig. 7**; and Fig. S6, available at http://links.lww.com/PAIN/A710; legend available at http://links.lww.com/PAIN/A724).

We applied this methodology to analyze GWAS results for human back pain, the most common human musculoskeletal pain condition. We used data from the UK Biobank project, for which genotyping and medical records were collected for about 500,000 people.[2,58] Enrichment analyses showed that SNPs within the genes that are differentially expressed in CFA/SNI are overrepresented in top differentially expressed pathways (**Fig. 7**). A heatmap shows which tissues and pain assays exhibit the highest association with the human back pain phenotype. For the ECM organization pathway, genes differentially expressed in the brain and in the SC in an inflammation profile, and in DRG in a neuropathic profile, showed marked enrichment of top GWAS SNPs in these genes.

3. Discussion

To further our understanding of pain processes at the molecular level, and to assess in an unbiased fashion the potential roles of

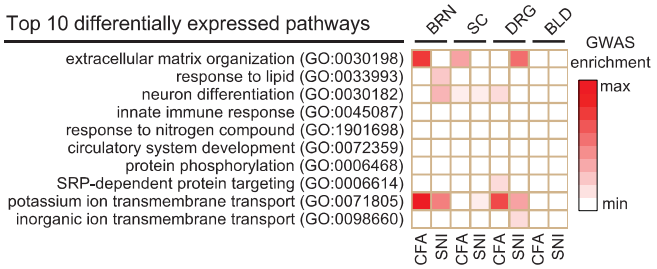

Figure 7. Enrichment of SNPs in back pain GWAS for the ECM organization pathway. Heatmap shows an enrichment of SNPs with strongest association with human back pain condition in loci of genes whose orthologs in mice are differentially expressed in experimental pain assays. The results are presented for the top 10 differentially expressed pathways in a tissue-specific fashion (top) and in a pain-specific fashion (bottom). ECM, extracellular matrix; GWAS, genome-wide association study.

contributing biological processes, we performed genome-wide next-generation transcriptomics analyses of pain-relevant tissues in 2 mouse pain assays (**Fig. 1**). This approach provides a deep analysis of transcriptomics of pain states. An alternative single-cell sequencing approach generates a detailed picture of changes in various cell types at the transcriptome level. However, a compromise is made between the number of samples and the sequencing depth or coverage. Typical results of single-cell analyses report expression data for only about 3500 genes (with large SDs),[65] a direct consequence of the sequencing depth of about 2×10^5 mapped reads per sample.[24] The deep sequencing obtained in our study (10^7 mapped reads compared to 10^5) allows for detecting low expressed genes, but most importantly, to assign statistical significance of minute differences in expression fold changes even in the mixed-cell–type population. Once identified, genes of interest can be checked for expression in specific cell types in single-cell sequencing resources, in DRG,[65] or in SC.[24]

We performed pathway analysis of DEGs across tissues and assays. We identified several pathways known to be involved in pain mediation, such as response to lipids, neuron differentiation, and innate immune response. Strikingly, the most regulated pathway in our analysis was ECM organization. The top 10 DEGs within the ECM organization category encode proteins with diverse functions including structural proteins such as collagen and cartilage components (*Col5a3*, *Comp*, *Col1a1*, and *Sparc*), proteins involved in cell-to-cell and cell-to-matrix interactions (*Vwf*, *Thbs1*, and *Cd36*), and enzymes involved in ECM remodelling (*Ctss*, *Elane*, and *Mmp13*). The observed diversity in DEGs within the ECM suggests that identification of the ECM organization as the most commonly regulated pathway in pain assays reflects a central role of this biological domain as a whole and is not related to a specific protein or group of proteins.

After identifying the top 10 biological pathways associated with pain states in mouse assays, we attempted to correlate these data with human GWAS of back pain (**Fig. 7**). We observed overrepresentation of the genes corresponded to the top SNPs of the back pain GWAS results within the ECM organization pathway. The statistical enrichment was observed for the SNPs situated within the genes identified in the inflammatory assay, detected both in the brain and in the SC, and in the neuropathic assay, detected in the DRG. Thus, we obtained evidence for a major contribution of the ECM organization pathway to the molecular pathophysiology of back pain and identified critical elements of this contribution at multiple levels: (1) the top associated SNPs, from the human GWAS, (2) the contributing genes in which these SNPs reside, (3) the contributing pathways, as genes are organized in modules, (4) the tissues in which transcriptional changes and responses occur, and (5) the type of contributing pain profile, inflammation, or neuropathic. Although the association analysis was restricted to 10 pathways, it recognized a new genetic pathway contributing to human back pain and was able to identify combined impacts of individual SNPs despite their modest effect sizes. Most of the DEGs within the ECM organization pathway in the mouse assays of pain state were detected in the dorsal horn of the SC and DRG tissues, with some DEGs also observed in the brain. We thus anticipate that ECM remodeling detected from expression profiling would impact behavior at the whole organism level.[12]

The ECM is composed of a variety of structural proteins, creating the extracellular scaffold surrounding cells in the tissue.[7,42] It also includes molecules mediating cell–cell and cell–matrix interaction, and enzymes involved in ECM remodeling. In the nervous system, ECM components provide not only the structural support for neuronal and nonneuronal cells, but also regulate synapse formation and function, and modulate neuronal excitability.[4,46,52] In the cortex, hippocampus, and amygdala, the ECM has been demonstrated to restrict neuroplasticity by inhibiting activity-dependent structural reorganization of presynaptic and post-synaptic compartments,[55] suppressing the lateral mobility of AMPA receptors in the membrane[21] and modulating neuronal excitability.[4] Accordingly, disruption of the ECM enhanced synaptic plasticity, acquisition of memories, and cognitive flexibility and extinction.[46,50,71,74] Despite this progress in the understanding key roles of ECM in synaptic plasticity and memory formation, the role of the ECM in regulation of nociceptive circuits and the development of chronic pain is not well known.

Removal of inhibitory constraints is a paramount mechanism to increase the excitability of nociceptive circuits in pathological pain conditions.[14,45] Repeated or intense activation of sensory neurons in response to peripheral tissue injury leads to a long-lasting increase in the excitability of SC circuits (a process known as central sensitization), resulting in a significant amplification of peripheral inputs. Although central sensitization results from a combination of mechanisms, the most prominent one is a reduction of inhibitory tone. Several mechanisms have been proposed to mediate the reduction of inhibitory tone in the SC in pathological pain conditions, including (1) decreased expression of neuron-specific K^+-Cl^- cotransporter-2 (KCC2) in spinal neurons, leading to a change in chloride gradient, and thereby attenuated inhibitory synaptic currents in response to GABA/glycine[14]; (2) decrease in the activity of spinal inhibitory neurons[62]; and, (3) reduced descending inhibition.[44] In light of our results and previous studies in the cortex and hippocampus, disinhibition of ECM-mediated suppression of spinal and peripheral pain circuits might be a novel potential mechanism promoting painful hypersensitive states. Modulation of ECM composition and density could impact nociception not only by direct effect on neuronal activity but also through T-cell and macrophage infiltration, as well as nerve and blood vessel sprouting, all factors potentially contributing to chronic pain.[26]

We noted a number of limitations of this study. A single mouse sex was used in our study to simplify interpretation and reduce heterogeneity, given well-known sex differences in pain processing.[41] Female mice were chosen since women represent the clear majority of chronic pain patients.[41] Furthermore, we compared transcriptomics of mice subjected to chronic inflammation and nerve injury with matching home cage controls and not vehicle/sham groups. This experimental design is suboptimal for animal preclinical studies, but because we were comparing mouse data

sets with human transcriptomics and GWAS, we reasoned that this approach would generate data that better match human pain studies, where healthy individuals serve as controls. Finally, we used adult mice (18 weeks) to better mimic the situation in human population where the disease is more common in adults. However, further aging mice may display yet unique molecular processes in pain assays, reflecting aging human aging processes. Recognizing our limitations, we stressed that we consider our results as a discovery that is needed to be followed up and replicated in different populations, including males and older mice, and different conditions, including better differentiating between pain and injury.

In summary, our catalog of genes differentially expressed in 2 mouse pain assays will be a valuable resource for the pain research field, enabling study of the molecular basis underlying pain processing. We showed that changes in transcriptome in neuropathic and IP states in animal models display substantial overlap between each other and with human LBP. Thus, these results capture shared molecular pathophysiological mechanisms between neuropathic and IP states. Furthermore, these results validate mouse pain assays for studying human pain conditions, although this does not exclude or quantify species specificity in pain mechanisms. Finally, and most importantly, we identified the ECM organization pathway as a major contributor to both pain etiologies, including its major contribution to risk of developing back pain at the level of human genetic variability.

4. Materials and methods

4.1. Statistics

Statistical tests used are indicated where the tests are performed. Correction for multiple testing is applied whenever multiple tests were involved, along with a statement indicating which method of correction is used, based on a fair assessment of how correlated the multiple tests are (eg, Bonferroni, Benjamini–Hochberg or FDR, etc). Threshold for statistical significance is P-value ≤ 0.05 for each test, unless specifically indicated at the location of the test.

4.2. Animals

For whole transcriptome studies, nine 18-week-old BALB/c female mice were randomly assigned into 3 groups of 3 mice: (1) control (CTR, naive), (2) IP induced with CFA, and (3) SNI for modelling NP. For chondroitinase ABC studies, both male and female mice were used. All animals used in this experiment were bred on site, and littermates were introduced into the study as soon as they matured to 8 weeks of age. All mice were housed in standard polycarbonate cages in groups of 3 or 4 same sex, in a temperature-controlled (20 ± 1°C) environment (14:10-hour light/dark cycle; lights on at 07:00 hours); tap water and food (Harlan Teklad 8604; Teklad Diets, Madison, WI) were available ad libitum.

4.3. Mouse pain assays

4.3.1. Complete Freund's adjuvant assay

Complete Freund's adjuvant (50%; Sigma) was injected subcutaneously in a volume of 20 μL into the left and right plantar hind paws using a 100-μL microsyringe with a 30-gauge needle. Three days after CFA, mice were decapitated, and blood was collected along with the brain and cerebellum, DRG (ipsilateral L3-L5), and ipsilateral dorsal horn of SC. Blood was collected in RNAprotect Animal Blood tubes (Qiagen, Germantown, MD) and stored according to manufacturer's recommendations. Blood and tissue samples were kept at −80°C until RNA extraction.

4.3.2. Spared nerve injury assay

Spared nerve injury, an experimental nerve injury designed to produce NP, was performed under isoflurane/oxygen anesthesia as described previously.[54] Briefly, using an operating microscope (×40), the 3 terminal branches of the sciatic nerve (tibial, sural, and common peroneal) were exposed. The tibial and common peroneal nerves were cut, after tight ligation with 6.0 silk, "sparing" the sural nerve. The incisions were closed in layers using interrupted sutures (6-0 Vicryl) and wound clips. The animal recovered on a thermostatically controlled heating pad (carefully monitored to prevent overheating) until ambulatory as per standard operating procedures. To reduce the overall number of mice required, the SNI surgery was performed in a bilateral fashion (ie, left and right side). Seven days after the surgery, blood and other tissues (L3-L5 DRG, dorsal horn of the SC, and brain) were collected and stored as for CFA-treated mice.

4.3.3. Von Frey testing

The up–down method of Dixon[9] was used to estimate 50% withdrawal thresholds using nylon monofilaments (Stoelting Touch Test), calibrated monthly. All experiments took place during the light cycle, between 9:00 and 16:00 hours. Mice were placed in custom-made Plexiglas cubicles (5 × 8.5 × 6 cm) on a perforated metal floor and were permitted to habituate for at least 1 hour before testing. Filaments were applied to the plantar surface of the hind paw for 1 second, and responses were recorded. Two consecutive measures were taken on each hind paw at each time point and averaged. Curves were plotted using GraphPad Prism v.7.

4.4. RNA extraction

Total RNA from nervous tissues was isolated using RNeasy Lipid Tissue Mini Kit, whereas RNA from blood cells was isolated using the RNeasy Mini Kit, including DNaseI treatment (all from Qiagen), according to the manufacturer's instructions. Total RNA was quantified using the NanoDrop 2000 (Thermo Scientific), and the RNA quality was assessed with the 2100 Bioanalyzer (Agilent Technologies).

4.5. Whole transcriptome sequencing (RNA-Seq)

All RNA-sequencing procedures were performed by Genomics Platform facility (Institute for Research in Immunology and Cancer, Montreal, Canada). Transcriptome libraries were generated from 1 μg of total RNA using the Kapa RNA–stranded Sample Prep Kit (KK8400; KAPABiosystems) following the manufacturer's protocols. Briefly, poly-A mRNA was purified using poly-T oligo-attached magnetic beads using 2 rounds of purification. During the second elution of the poly-A RNA, the RNA was fragmented and primed for cDNA synthesis. During cDNA synthesis, dUTP was incorporated in the second-strand synthesis, where dUTP-containing strand was selectively degraded. Adenylation of the 3′ ends and ligation of adapters were performed after the manufacturer protocol. Enrichment of DNA fragments that have adapter molecules on both ends was performed using 10 cycles of PCR amplification using the KAPA PCR mix and Illumina-adapted primers cocktail.

4.6. Sequencing

Paired-end (2 × 100 bp) sequencing was performed using the Illumina HiSeq2000 machine running TruSeq v3 chemistry. Details of the Illumina sequencing technologies can be found at https://www.illumina.com/techniques/sequencing/rna-sequencing.html.

4.7. Processing of RNA-Seq data

RNA-Seq data have been trimmed with Trimmomatic v0.32,[6] then mapped on UCSC's mouse genome version mm10 (grabbed from ftp://ussd-ftp.illumina.com/Mus_musculus/UCSC/mm10/) using tophat v2.0.11[27] and bowtie v1.0.0.[32] Differential gene expression detected at gene level from experimental triplicates using cuffdiff v2.2.1.[63] The RNA-Seq data have been deposited in NCBI's Gene Expression Omnibus[19] and are accessible through GEO accession number GSE111216. We used a level of statistical significance at alpha 0.05 to define DEGs $|\pi| \geq$ critical value of 1.11.[73] Gene names from HUGO Gene Nomenclature Committee (HGNC).[75]

RNA fragment raw counts ranged from 61 M to 195 M, with an average of 93 M reads (SD of 30 M) (Fig. S2, available at http://links.lww.com/PAIN/A710; legend available at http://links.lww.com/PAIN/A724). After adapter trimming, the 100 nts paired-end reads had a mean length of 94 nts (SD of 13 nts), retaining the specificity of paired-end sequencing. The average ratio of mapped to trimmed reads was 89% (SD 4%). Within-group gene expressions was all strongly correlated: *P*-values for statistical significance for all correlated pairs were estimated to lie below 2.2×10^{-16}, whereas squared Pearson's correlation coefficients r^2 ranged from 0.79 to 0.98 (Fig. S3, available at http://links.lww.com/PAIN/A710; legend available at http://links.lww.com/PAIN/A724). Tissues that display highest correlated gene expression were DRG (mean r^2 0.98), followed by SC (0.97) and whole brain (0.96). Less correlated is whole blood (0.84). The CFA pain model showed marginally more robustly correlated gene expression (mean r^2 0.95) than SNI (0.93).

4.8. Replication of publicly available data

Replication of differential expression was attempted using genome-wide data sets, available in the GEO database,[19] as well using a survey of differentially expressed, but pain-centric genes.[30] A total of 79 genes were identified through significant fold change in at least 4 different microarray experiments. Of the study's initial list of 79 genes, 65 were matched to HGNC's official gene names. For each gene surveyed, information about expression fold change (FC) direction (up, down) has been compiled in 2 tissues, namely SC and DRG, and for 2 pain models, namely IP and NP. Here, we have been careful to interpret absence of change direction as not necessarily evidence for absence of change, therefore leading to a two-state model of expression fold change per gene. For replication, we first determined in our data the direction of change of expression, up or down, based on the sign of the change and from associated significance *P*-value $P < 0.05$ (the fold change was unassigned if *P*-value is insignificant). For instance, we find the gene cathepsin S, Ctss, upregulated in SNI in DRG ($\log_2$ FC = +1.2, $P = 5 \times 10^{-5}$). Then, for each gene, we compiled the number of times (up to 4; 2 tissues in 2 pain models) that the survey and our results committed to call a fold change direction in corresponding tissue-pain model pairs. For instance, the gene Ctss has 3 corresponding committed calls: DRG and SC in NP, and SC in IP. We also tracked fold change mismatches, or conflicts, in committed calls. For instance, gene Tac1 (tachykinin precursor 1) has previously been found to be upregulated in DRG in NP but is found to be downregulated in our work.

4.9. Human and mouse transcriptomes comparison

Human gene expression data from the naive tissues studied in the current data set were downloaded from[5]: http://xavierlab2.mgh.harvard.edu/EnrichmentProfiler/download.html. Blood RNA samples from human subjects diagnosed with low-back pain were collected in University of Parma hospital at the first visit and at the 3-month follow-up from the acute episode. The enrolled subjects reported pain for less than 8 weeks and did not have other pain episodes within the last 6 months. The subjects with previous history of cancer or vertebral fracture were excluded. Blood was collected by venipuncture into Tempus Blood tubes (Applied Biosystems, Beverly, MA). Total RNA was extracted using Maxwell 16 LEV simplyRNA Blood Kit (Promega, Madison, WI) and quantified using the NanoDrop 1000 (Thermo Scientific, Waltham, MA). The RNA quality was assessed with the 4200 TapeStation (Agilent Technologies, Santa Clara, CA). Blood transcriptomics data for control subjects were from GEO set GSE90081.[53] RNA-Seq data were mapped using STAR aligner,[18] and differential expression of genes was assessed using FeatureCount[35] followed by Deseq2.[36]

4.10. Species-specific gene expression

Identification of species-specific gene expression was performed by counting in each species how many tissues expresses the gene (from N = 0-4), then computing the bias $|N_{mouse}\text{-}N_{human}|$; a bias of 2 or more defines a gene that is species-specific.

4.11. Star plot

In the star plot, the 4 axes were defined: axis 1 identified genes that were exclusively differentially expressed in the CFA pain model (green; $|\pi_{CFA}| \geq 2$ and $|\pi_{SNI}| \leq 1$); axis 3, genes in SNI only (purple; $|\pi_{SNI}| \geq 2$ and $|\pi_{CFA}| \leq 1$); axis 2, genes in both CFA and SNI with same +/+ or −/− fold change sign (pink; $|\pi_{CFA}| \geq 2$ or $|\pi_{SNI}| \geq 2$, and $2.0 \geq \pi_{CFA}/\pi_{SNI} \geq 0.5$); and finally axis 4, genes in both CFA and SNI, but this time with opposite ± or −/+ fold change sign (orange; $|\pi_{CFA}| \geq 2$ or $|\pi_{SNI}| \geq 2$, and $2.0 \geq -\pi_{CFA}/\pi_{SNI} \geq 0.5$). Because each axis required statistical significance at the alpha 0.01 level with $|\pi| \geq 2$,[73] identified genes represented a confident set to perform further analyses.

4.12. Gene set enrichment analyses

Gene Ontology's (GO) biological processes pathways have been used for analysis.[3,22] For each star plot, the most significant pathway was reported for each of the 4 axes. Once a pathway was reported, genes that were members of that pathway were removed from the list of genes and the GSEA was repeated; this reduced the likelihood that pathways with similar or related functions be reported. Because we aimed for biological specificity, we limited analyses to pathways with 1000 genes or less. In GO, pathways are organized in a hierarchical fashion; we discarded pathways that have more than 5 generations of children, again favoring pathways with more defined biological functions. Finally, we required that a selected pathway has no more than 50% overlap in gene composition with all those pathways previously selected, promoting functional diversity, but more importantly, preventing closely related pathways to be overrepresented during the selection process.

4.13. Human genome-wide association studies

We performed GWAS on data from the UK Biobank project (application number 20802).[2,58] Back pain was selected as a phenotype of choice because it is present in a large proportion of the population (about 8%). To define cases, we used the field 3571, where subjects answered "no" to the question "Have you had back pains for more than 3 months?," while answering "yes" to the "back pain" category in field 6159 where the question was "In the last month have you experienced any of the following that interfered with your usual activities?" (n = 32,209). The limited time extent is in line with the time span between pain intervention in the mice and their sacrifice. Control individuals were those that answered "none of the above" for the question in field 6159 (n = 163,825). First-hand standard genotyping quality control was performed by the UK BioBank consortium, and is fully documented on their web portal (http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=155580). Genome-wide association study was performed using genotypic data only, using the 500K cohort version. People were discarded based on: failed genotyping QC (as per UK BioBank–heterozygosity rate, genotyping rate, etc.), genetic vs declared sex mismatch, voluntary retraction, and of non-"white British" ancestry. Among individuals with first- and second-degree relatives, we retained the one individual who displayed the best genotyping rate. Kinship was estimated using KiNG v2.1.[37] We used PLINK version v1.90b4.6 64-bit (August 15, 2017) to perform logistic regression,[48] using age, sex, genotyping platform, and the first 5 genetic principal components as covariables. Post-GWAS SNP QC featured at least 1% minor allele frequency, genotyping rate better than or equal to 90%, and *P*-value for departure from Hardy–Weinberg equilibrium no less than 10^{-6}. PLINK was also used to extract linkage disequilibrium information between SNPs and select tag SNPs based on $r^2 \leq 0.5$; enrichment of SNPs in genes was evaluated using these tag SNPs. Genes that span several haploblocks will be represented by each block's tag SNP; this might seem to skew analyses as longer genes will contribute to more SNPs, but usage of tag SNPs guarantee uncoupled GWAS results within long genes as tag SNPs are, by definition, in low LD with one another.

4.14. Genes commonly implicated in pain pathways

The list of genes commonly implicated in pain pathways has been composed from several sources: for human pain genes, we used a combination of Algynomics'/Cogenics Pain Research Panel V2[38] and Pain Research Forum's pain gene resource (https://www.painresearchforum.org/), and Amigo's Gene Ontology term GO:0019233, whereas for mouse pain genes, we used the Pain Genes Database.[31] We also used Ensembl BioMart service to provide for a correspondence between human genes and mouse ones. The pain genes list comprises a total of n = 800 pain genes (Table S12, available at http://links.lww.com/PAIN/A722; legend available at http://links.lww.com/PAIN/A724), with matching human/mouse homologs. All mouse pain genes have human homologs, whereas a few human pain genes have no mouse homologs (n = 21).

4.15. Protein interaction network

We considered data from the STRING database, a repository of interacting proteins.[59] Only interactions in the human taxon (9606) were considered because they were more numerous than that of the mouse (10090) ones.

4.16. Study approval

All mice experimental procedures were approved by the Animal Care Committee at McGill University under protocol number 7869 and were performed in full agreement with the ethical guidelines of the Canadian Council of Animal Care and the guidelines of the Committee for Research and Ethical Issues of the International Association for the Study of Pain.

The human low-back pain cohort is part of a larger study (PainOmics), approved by the Ethical Committee at the hospital university of Parma, protocol number 43543 version 8, and registered on clinicaltrials.gov (NCT02037763). All patients signed a written informed consent before the enrolment.

Conflict of interest statement

The authors have no conflict of interest to declare.

The accession number for the data sets reported in this article is GEO set GSE111216.

Acknowledgments

The authors thank Dr. Samar Khoury for help with data extraction from the UK Biobank. The current study was conducted under UK Biobank application number 20802.
Funding for this work is kindly provided by the Canadian Excellence Research Chairs (CERC) Program (www.cerc.gc.ca) grant CERC09 (to L.D.), and the Rita Allen Foundation and the American Pain Society Award in Pain (to A.K.).
Author contributions: L. Diatchenko, J. S. Mogil, M. Parisien, A. Samoshkin, A. Khoutorsky, M. Allegri, N. El-Hachem, and S. N. Tansley designed the analytical plan and experiments. A. Samoshkin, S. N. Tansley, M. H. Piltonen, and L. J. Martin performed animal experiment, collected, and processed mouse tissues. M. Allegri and C. Dagostino were responsible for human subject collection and samples processing. M. Parisien performed all bioinformatics analyses, except human blood mRNA processed by N. El-Hachem. M. Parisien, A. Samoshkin, A. Khoutorsky, and L. Diatchenko wrote the manuscript. A. Khoutorsky, J. S. Mogil, M. Allegri and L. Diatchenko supervised the project. All the authors read and edited the final manuscript.

Appendix A. Supplemental digital content

Supplemental digital content associated with this article can be found online at http://links.lww.com/PAIN/A710, http://links.lww.com/PAIN/A711, http://links.lww.com/PAIN/A712, http://links.lww.com/PAIN/A713, http://links.lww.com/PAIN/A714, http://links.lww.com/PAIN/A715, http://links.lww.com/PAIN/A716, http://links.lww.com/PAIN/A717, http://links.lww.com/PAIN/A718, http://links.lww.com/PAIN/A719, http://links.lww.com/PAIN/A720, http://links.lww.com/PAIN/A721, http://links.lww.com/PAIN/A722, http://links.lww.com/PAIN/A723, and http://links.lww.com/PAIN/A724.

Article history:
Received 22 October 2018
Received in revised form 6 December 2018
Accepted 14 December 2018
Available online 3 January 2019

References

[1] Abdelaziz DM, Abdullah S, Magnussen C, Ribeiro-da-Silva A, Komarova SV, Rauch F, Stone LS. Behavioral signs of pain and functional impairment in a mouse model of osteogenesis imperfecta. Bone 2015;81:400–6.

[2] Allen NE, Sudlow C, Peakman T, Collins R; Biobank UK. UK biobank data: come and get it. Sci Transl Med 2014;6:224ed4.
[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] Balmer TS. Perineuronal nets enhance the excitability of fast-spiking neurons. eNeuro 2016;3. doi: 10.1523/ENEURO.0112-16.2016.
[5] Benita Y, Cao Z, Giallourakis C, Li C, Gardet A, Xavier RJ. Gene enrichment profiles reveal T-cell development, differentiation, and lineage-specific transcription factors including ZBTB25 as a novel NF-AT repressor. Blood 2010;115:5376–84.
[6] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics 2014;30:2114–20.
[7] Bonnans C, Chou J, Werb Z. Remodelling the extracellular matrix in development and disease. Nat Rev Mol Cell Biol 2014;15:786–801.
[8] Cattaruzza F, Lyo V, Jones E, Pham D, Hawkins J, Kirkwood K, Valdez-Morales E, Ibeakanma C, Vanner SJ, Bogyo M, Bunnett NW. Cathepsin S is activated during colitis and causes visceral hyperalgesia by a PAR2-dependent mechanism in mice. Gastroenterology 2011;141: 1864–74.e1861–1863.
[9] Chaplan SR, Bach FW, Pogrel JW, Chung JM, Yaksh TL. Quantitative assessment of tactile allodynia in the rat paw. J Neurosci Meth 1994;53: 55–63.
[10] Clark AK, Wodarski R, Guida F, Sasso O, Malcangio M. Cathepsin S release from primary cultured microglia is regulated by the P2X7 receptor. Glia 2010;58:1710–26.
[11] Clark AK, Yip PK, Malcangio M. The liberation of fractalkine in the dorsal horn requires microglial cathepsin S. J Neurosci 2009;29:6945–54.
[12] Cobos EJ, Nickerson CA, Gao F, Chandran V, Bravo-Caparrós I, González-Cano R, Riva P, Andrews NA, Latremoliere A, Seehus CR, Perazzoli G, Nieto FR, Joller N, Painter MW, Ma CHE, Omura T, Chesler EJ, Geschwind DH, Coppola G, Rangachari M, Woolf CJ, Costigan M. Mechanistic differences in neuropathic pain modalities revealed by correlating behavior with global expression profiling. Cell Rep 2018;22: 1301–12.
[13] Costigan M, Moss A, Latremoliere A, Johnston C, Verma-Gandhu M, Herbert TA, Barrett L, Brenner GJ, Vardeh D, Woolf CJ, Fitzgerald M. T-cell infiltration and signaling in the adult dorsal spinal cord is a major contributor to neuropathic pain-like hypersensitivity. J Neurosci 2009;29: 14415–22.
[14] Coull JA, Boudreau D, Bachand K, Prescott SA, Nault F, Sik A, De Koninck P, De Koninck Y. Trans-synaptic shift in anion gradient in spinal lamina I neurons as a mechanism of neuropathic pain. Nature 2003;424: 938–42.
[15] Davis S, Meltzer PS. GEOquery: a bridge between the gene expression Omnibus (GEO) and BioConductor. Bioinformatics 2007;23:1846–7.
[16] Denning WM, Woodland S, Winward JG, Leavitt MG, Parcell AC, Hopkins JT, Francom D, Seeley MK. The influence of experimental anterior knee pain during running on electromyography and articular cartilage metabolism. Osteoarthritis Cartilage 2014;22:1111–9.
[17] Dib-Hajj SD, Waxman SG. Translational pain research: lessons from genetics and genomics. Sci Transl Med 2014;6:249sr244.
[18] Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.
[19] Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002;30:207–10.
[20] Freynhagen R, Baron R. The evaluation of neuropathic components in low back pain. Curr Pain Headache Rep 2009;13:185–90.
[21] Frischknecht R, Heine M, Perrais D, Seidenbecher CI, Choquet D, Gundelfinger ED. Brain extracellular matrix affects AMPA receptor lateral mobility and short-term synaptic plasticity. Nat Neurosci 2009;12: 897–904.
[22] Gene_Ontology_Consortium. Gene ontology consortium: going forward. Nucleic Acids Res 2015;43:D1049–1056.
[23] Goshua A, Craigie S, Guyatt GH, Agarwal A, Li R, Bhullar JS, Scott N, Chahal J, Pavalagantharajah S, Chang Y, Couban R, Busse JW. Patient values and preferences regarding opioids for chronic noncancer pain: a systematic review. Pain Med 2017;19:2469–80.
[24] Haring M, Zeisel A, Hochgerner H, Rinwa P, Jakobsson JET, Lonnerberg P, La Manno G, Sharma N, Borgius L, Kiehn O, Lagerstrom MC, Linnarsson S, Ernfors P. Neuronal atlas of the dorsal horn defines its architecture and links sensory input to transcriptional cell types. Nat Neurosci 2018;21:869–80.
[25] 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.
[26] Ji RR, Chamessian A, Zhang YQ. Pain regulation by non-neuronal cells and inflammation. Science 2016;354:572–7.
[27] Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 2013;14:R36.
[28] Kluzek S, Bay-Jensen AC, Judge A, Karsdal MA, Shorthose M, Spector T, Hart D, Newton JL, Arden NK. Serum cartilage oligomeric matrix protein and development of radiographic and painful knee osteoarthritis. A community-based cohort of middle-aged women. Biomarkers 2015; 20:557–64.
[29] Korczeniewska OA, Husain S, Khan J, Eliav E, Soteropoulos P, Benoliel R. Differential gene expression in trigeminal ganglia of male and female rats following chronic constriction of the infraorbital nerve. Eur J Pain 2018;22:875–88.
[30] LaCroix-Fralish ML, Austin JS, Zheng FY, Levitin DJ, Mogil JS. Patterns of pain: meta-analysis of microarray studies of pain. PAIN 2011;152: 1888–98.
[31] Lacroix-Fralish ML, Ledoux JB, Mogil JS. The pain genes database: an interactive web browser of pain-related transgenic knockout studies. PAIN 2007;131:3.e1–4.
[32] Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 2009;10:R25.
[33] Legerlotz K, Jones ER, Screen HR, Riley GP. Increased expression of IL-6 family members in tendon pathology. Rheumatology (Oxford) 2012;51: 1161–5.
[34] Leichsenring A, Backer I, Wendt W, Andriske M, Schmitz B, Stichel CC, Lubbert H. Differential expression of Cathepsin S and X in the spinal cord of a rat neuropathic pain model. BMC Neurosci 2008;9:80.
[35] Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 2014;30:923–30.
[36] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.
[37] Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. Bioinformatics 2010;26:2867–73.
[38] Meloto CB, Bortsov AV, Bair E, Helgeson E, Ostrom C, Smith SB, Dubner R, Slade GD, Fillingim RB, Greenspan JD, Ohrbach R, Maixner W, McLean SA, Diatchenko L. Modification of COMT-dependent pain sensitivity by psychological stress and sex. PAIN 2016;157:858–67.
[39] Millecamps M, Tajerian M, Sage EH, Stone LS. Behavioral signs of chronic back pain in the SPARC-null mouse. Spine (Phila Pa 1976) 2011; 36:95–102.
[40] Mitchell JA, Aronson AR, Mork JG, Folk LC, Humphrey SM, Ward JM. Gene indexing: characterization and analysis of NLM's GeneRIFs. AMIA Annu Symp Proc 2003:460–4.
[41] Mogil JS. Sex differences in pain and pain inhibition: multiple explanations of a controversial phenomenon. Nat Rev Neurosci 2012;13:859–66.
[42] Mouw JK, Ou G, Weaver VM. Extracellular matrix assembly: a multiscale deconstruction. Nat Rev Mol Cell Biol 2014;15:771–85.
[43] Niederberger E, Resch E, Parnham MJ, Geisslinger G. Drugging the pain epigenome. Nat Rev Neurol 2017;13:434–47.
[44] Ossipov MH, Morimura K, Porreca F. Descending pain modulation and chronification of pain. Curr Opin Support Palliat Care 2014;8:143–51.
[45] Petitjean H, Pawlowski SA, Fraine SL, Sharif B, Hamad D, Fatima T, Berg J, Brown CM, Jan LY, Ribeiro-da-Silva A, Braz JM, Basbaum AI, Sharif-Naeini R. Dorsal horn parvalbumin neurons are gate-keepers of touch-evoked pain after nerve injury. Cell Rep 2015;13:1246–57.
[46] Pizzorusso T, Medini P, Berardi N, Chierzi S, Fawcett JW, Maffei L. Reactivation of ocular dominance plasticity in the adult visual cortex. Science 2002;298:1248–51.
[47] Price TJ, Inyang KE. Commonalities between pain and memory mechanisms and their meaning for understanding chronic pain. Prog Mol Biol Transl Sci 2015;131:409–34.
[48] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75.
[49] Riediger C, Schuster T, Barlinn K, Maier S, Weitz J, Siepmann T. Adverse effects of antidepressants for chronic pain: a systematic review and meta-analysis. Front Neurol 2017;8:307.
[50] Romberg C, Yang S, Melani R, Andrews MR, Horner AE, Spillantini MG, Bussey TJ, Fawcett JW, Pizzorusso T, Saksida LM. Depletion of

perineuronal nets enhances recognition memory and long-term depression in the perirhinal cortex. J Neurosci 2013;33:7057–65.
[51] Rozenbaum M, Rajman M, Rishal I, Koppel I, Koley S, Medzihradszky KF, Oses-Prieto JA, Kawaguchi R, Amieux PS, Burlingame AL, Coppola G, Fainzilber M. Translatome regulation in neuronal injury and axon regrowth. eNeuro 2018;5. doi: 10.1523/ENEURO.0276-17.2018.
[52] Senkov O, Andjus P, Radenovic L, Soriano E, Dityatev A. Neural ECM molecules in synaptic plasticity, learning, and memory. Prog Brain Res 2014;214:53–80.
[53] Shchetynsky K, Diaz-Gallo LM, Folkersen L, Hensvold AH, Catrina AI, Berg L, Klareskog L, Padyukov L. Discovery of new candidate genes for rheumatoid arthritis through integration of genetic association data with expression pathway analysis. Arthritis Res Ther 2017;19:19.
[54] Shields SD, Eckert WA III, Basbaum AI. Spared nerve injury model of neuropathic pain in the mouse: a behavioral and anatomic analysis. J Pain 2003;4:465–70.
[55] Sorg BA, Berretta S, Blacktop JM, Fawcett JW, Kitagawa H, Kwok JC, Miquel M. Casting a wide net: role of perineuronal nets in neural plasticity. J Neurosci 2016;36:11459–68.
[56] de Souza JB, Grossmann E, Perissinotti DMN, de Oliveira Junior JO, de Fonseca PRB, Posso IP. Prevalence of chronic pain, treatments, perception, and interference on life activities: Brazilian population-based survey. Pain Res Manag 2017;2017:4643830.
[57] Strong JA, Xie W, Coyle DE, Zhang JM. Microarray analysis of rat sensory ganglia after local inflammation implicates novel cytokines in pain. PLoS One 2012;7:e40779.
[58] Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T, Collins R. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015;12:e1001779.
[59] Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, von Mering C. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 2015;43: D447–452.
[60] Tajerian M, Alvarado S, Millecamps M, Dashwood T, Anderson KM, Haglund L, Ouellet J, Szyf M, Stone LS. DNA methylation of SPARC and chronic low back pain. Mol Pain 2011;7:65.
[61] 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.
[62] Torsney C, MacDermott AB. Disinhibition opens the gate to pathological pain signaling in superficial neurokinin 1 receptor-expressing neurons in rat spinal cord. J Neurosci 2006;26:1833–43.
[63] 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.
[64] Tsarouhas A, Soufla G, Katonis P, Pasku D, Vakis A, Spandidos DA. Transcript levels of major MMPs and ADAMTS-4 in relation to the clinicopathological profile of patients with lumbar disc herniation. Eur Spine J 2011;20:781–90.
[65] Usoskin D, Furlan A, Islam S, Abdo H, Lonnerberg P, Lou D, Hjerling-Leffler J, Haeggstrom J, Kharchenko O, Kharchenko PV, Linnarsson S, Ernfors P. Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. Nat Neurosci 2015;18:145–53.
[66] Vardeh D, Mannion RJ, Woolf CJ. Toward a mechanism-based approach to pain diagnosis. J Pain 2016;17(9 suppl):T50–69.
[67] Vega-Avelaira D, Geranton SM, Fitzgerald M. Differential regulation of immune responses and macrophage/neuron interactions in the dorsal root ganglion in young and adult rats following nerve injury. Mol Pain 2009;5:70.
[68] Vicuna L, Strochlic DE, Latremoliere A, Bali KK, Simonetti M, Husainie D, Prokosch S, Riva P, Griffin RS, Njoo C, Gehrig S, Mall MA, Arnold B, Devor M, Woolf CJ, Liberles SD, Costigan M, Kuner R. The serine protease inhibitor SerpinA3N attenuates neuropathic pain by inhibiting T cell-derived leukocyte elastase. Nat Med 2015;21:518–23.
[69] Volkow N, Benveniste H, McLellan AT. Use and misuse of opioids in chronic pain. Annu Rev Med 2018;69:451–65.
[70] von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, Jouffre N, Huynen MA, Bork P. STRING: known and predicted protein-protein associations, integrated and transferred across organisms. Nucleic Acids Res 2005;33:D433–437.
[71] Wang D, Fawcett J. The perineuronal net and the control of CNS plasticity. Cell Tissue Res 2012;349:147–60.
[72] 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. doi: 10.1177/1744806916629048.
[73] Xiao Y, Hsiao TH, Suresh U, Chen HI, Wu X, Wolf SE, Chen Y. A novel significance score for gene selection and ranking. Bioinformatics 2014; 30:801–7.
[74] Xue YX, Xue LF, Liu JF, He J, Deng JH, Sun SC, Han HB, Luo YX, Xu LZ, Wu P, Lu L. Depletion of perineuronal nets in the amygdala to enhance the erasure of drug memories. J Neurosci 2014;34:6647–58.
[75] Yates B, Braschi B, Gray KA, Seal RL, Tweedie S, Bruford EA. Genenames.org: the HGNC and VGNC resources in 2017. Nucleic Acids Res 2017;45:D619–25.
[76] Zhang X, Wu Z, Hayashi Y, Okada R, Nakanishi H. Peripheral role of cathepsin S in Th1 cell-dependent transition of nerve injury-induced acute pain to a chronic pain state. J Neurosci 2014;34:3013–22.
[77] Zhao P, Lieu T, Barlow N, Metcalf M, Veldhuis NA, Jensen DD, Kocan M, Sostegni S, Haerteis S, Baraznenok V, Henderson I, Lindstrom E, Guerrero-Alba R, Valdez-Morales EE, Liedtke W, McIntyre P, Vanner SJ, Korbmacher C, Bunnett NW. Cathepsin S causes inflammatory pain via biased agonism of PAR2 and TRPV4. J Biol Chem 2014;289:27215–34.
[78] Zhao P, Lieu T, Barlow N, Sostegni S, Haerteis S, Korbmacher C, Liedtke W, Jimenez-Vargas NN, Vanner SJ, Bunnett NW. Neutrophil elastase activates protease-activated receptor-2 (PAR2) and transient receptor potential vanilloid 4 (TRPV4) to cause inflammation and pain. J Biol Chem 2015;290:13875–87.