Sex Differences in Nociceptor Translatomes Contribute to Divergent Prostaglandin Signaling in Male and Female Mice

Supplement 1

Suppl. Figure S1: *Ptgds expression is higher in female cortical neurons.* **A**, Ptgds is highly co-localized with neuronal marker NeuN in the mouse brain. **B**, Ptgds expression in mouse cortical neurons is higher in females compared to males (* Unpaired t-test, $t = 3.699$, df = 4, p-value= 0.0209). A total of 100 neurons per slice (3 slices per animal) were used to quantify Ptgds. Panel A Scale bar = 500 μm; panel B scale bar = 50 μm.

Suppl. Table S1A: Raw transcripts per million for INPUT and respective percentiles.

Suppl. Table S1B: Quantile normalized transcripts per million for INPUT and respective statistics.

Suppl. Table S2A: Raw transcripts per million for IP and respective percentiles.

Suppl. Table S2B: Quantile normalized transcripts per million for IP and respective statistics.

Suppl. Table S3: Differentially expressed genes in INPUT.

Suppl. Table S4: Differentially translated mRNAs in IP.

See separate Excel file for Tables S1-S4.

Suppl. Table S5: Functional analysis of differentially expressed INPUT genes. (ND: No data available).

Suppl. Table S6: Functional analysis of differentially translated IP mRNAs. (ND: No data available)

Suppl. Methods and Materials

Animals

All animal procedures were approved by the Institutional Animal Care and Use Committee of University of Texas at Dallas protocol number 14-04.

Nav1.8Cre/Rosa26fstrap mice**:** Rosa26fsTRAP mice were purchased from The Jackson Laboratory (stock #022367). Transgenic mice expressing Cre recombinase under the control of the Scn10a (Nav1.8) promoter were obtained initially from Professor John Wood (University College London) but are commercially available from Infrafrontier (EMMA ID: 04582). Initial studies demonstrated that the introduction of the Cre recombinase in heterozygous animals does not affect pain behavior, and their DRG neurons have normal electrophysiological properties (1). Nav1.8-cre mice on a C57BL/6J genetic background were maintained and bred at the University of Texas at Dallas. Upon arrival, Rosa26^{fsTRAP} mice were crossed to Nav1.8cre to generate the Nav1.8-TRAP mice that express a fused EGFP-L10a protein in Nav1.8 positive neurons. TRAP experiments were performed using male only and female only Nav1.8- TRAP littermates 8–12 weeks old. Mice were group housed (4 maximum) in nonenvironmentally enriched cages with food and water ad libitum on a 12 h light-dark cycle. Room temperature was maintained at $21 \pm 2^{\circ}$ C. Immunohistochemistry and ELISA assays were performed on C57BL/6J mice. Behavioral experiments were performed on C57BL/6J mice and ICR mice (we did not observe any strain differences).

TRAP

Nav1.8-TRAP male and female mice (3 mice per sex were pooled for each of 4 biological replicates) were decapitated under isoflurane anesthesia and DRGs rapidly dissected in icecold dissection buffer (1× HBSS (Invitrogen, 14065006), 2.5 mM HEPES-NaOH [pH 7.4], 35 mM Glucose, 5mM MgCl2, 100 μg/ml cycloheximide, 0.2 mg/ml emetine). DRGs were transferred to ice-cold Precellys Tissue homogenizing CKMix tube with lysis buffer (20 mm

HEPES, 12 mm MgCl2, 150 mm KCl, 0.5 mm DTT, 100 μg/ml cycloheximide, 20 μg/ml emetine, 80 U/ml SUPERase IN, Promega, 1 μl DNase, and protease inhibitor). The lysate was prepared by homogenizing the samples using Precellys® Minilys Tissue Homogenizer at 10 second intervals for a total of 80 seconds, in the cold room (4°C). Samples were then centrifuged at 2000 × g for 5 min to prepare postnuclear fractions. Next, 1% NP-40 and 30 mM 1,2-dihexanoyl-sn-glycero-3-phosphocholine were added and samples centrifuged at 15,000 × g for 10 min to generate a postmitochondrial fraction. A 200 μl sample of this fraction was saved for use as INPUT (bulk RNA-sequencing), and the remaining was incubated with protein G-coated Dynabeads (Invitrogen) bound to 50 μg anti-GFP antibodies (HtzGFP-19F7 and HtzGFP-19C8, Memorial Sloan Kettering Centre) for 3 h at 4°C with end-over-end mixing. Anti-GFP beads were washed with high salt buffer (20 mm HEPES, 5 mm MgCl2, 350 mm KCl, 1% NP-40, 0.5 mm DTT, and 100 μg/ml cycloheximide), and RNA was eluted from all samples using the Direct-zol kit (Zymo Research) according to the manufacturer's instructions. RNA yield was quantified using a Nanodrop system (Thermo Fisher Scientific), and RNA quality was determined by fragment Analyzer (Advanced Analytical Technologies).

Immunohistochemistry

Animals were anesthetized with isoflurane (4%) and euthanized by decapitation and tissues were flash frozen in OCT on dry ice. Sections of DRG (20 μm) and brain (20 μm) were mounted onto SuperFrost Plus slides (Thermo Fisher Scientific) and immediately fixed in ice-cold formalin (10%) for 15 min followed by dehydration in 50% ethanol, 70% ethanol and 100% ethanol at room temperature for 5 min each. Tissues were briefly air dried and boundaries were drawn around each section using a hydrophobic pen (ImmEdge H-4000). Once dry, the sections were blocked for at least 1 h in 10% normal goat serum with 0.3% TX-100. DRG slices were stained with peripherin, NeuN and Ptgds overnight at 4°C and brain slices were stained with NeuN and Ptgds for 2 h at room temperature. The peripherin antibody (P5117) was obtained from Sigma-Aldrich, NeuN (MAB377) was obtained from Millipore Sigma and Ptgds (ab182141) from Abcam. Sections were then washed and incubated with respective Alexa Fluor secondary antibodies for 1 h at room temperature. Sections were washed, air dried, and then coverslipped with Prolong Gold Antifade reagent (Fisher Scientific; P36930).

DRG images were taken using an Olympus FluoView 1200 confocal microscope, using the same settings for all images. Analysis of DRG images was done using ImageJ version 1.48 (National Institutes of Health, Bethesda, MD). The values plotted on Figure 5E were obtained by calculating the corrected total cell fluorescence (CTCF) using the following formula: CTCF= Integrated Density – (Area of selected cell X Mean fluorescence of background readings).

Brain images were taken using Olympus vs120 virtual slide microscope, using the same settings for all images. Analysis of brain images was performed using Olympus cellSens software. An ROI of the same size was placed over 100 randomly selected cortical neurons per section to measure the mean gray intensity value. 3 sections were analyzed per animal. Background fluorescence was measured similarly using a negative control that was only exposed to blocking solution and secondary antibodies (no primary). After subtracting background values of negative control, we averaged all intensity values (first per slice and second per animal) to obtain the values plotted on Suppl. Figure 1.

Ptgds validation and estrous cycle assessment

A separate TRAP experiment was conducted where we validated the expression of Ptgds bound to ribosome. We followed the same protocol as described above (TRAP) with two exceptions: we used only one mouse per replicate, and we tracked the female estrous cycle.

The stage of the cycle was determined by vaginal lavage followed by cytological evaluation according to a previous published protocol (2).

ELISA assay

PGD2 levels in the DRG were evaluated using Prostaglandin D2-MOX Express ELISA Kit (Cayman, 500151) following the manufacturer's instructions. The same number of DRGs (collected from all levels cervical, thoracic and lumbar) were used per sample in the ELISA assay.

Injections

PGE2 (Cayman, 14010) was diluted in sterile PBS and injected with a volume of 25 μl via a 30.5-gauge needle and given intraplantarly.

AT-56 (Tocris, 3531) was diluted in 10% DMSO and 30% cyclodextrine and administered intraperitoneally.

Behavior testing

All behavioral experiments were performed between 8:00 A.M. and 6:00 P.M. Facial grimacing was evaluated using the Mouse Grimace Scale (MGS) as described previously (3). Mechanical paw withdrawal thresholds were measured using the up-down method (4) with calibrated Von Frey filaments (Stoelting).

Library generation and sequencing

After purifying the RNA, cDNA libraries were prepared with total RNA Gold library preparation (with ribosomal RNA depletion) for all samples according to the manufacturer's instructions (Illumina). Quality control was performed for RNA extraction and cDNA library preparation steps with Qubit (Invitrogen) and High Sensitivity NGS fragment analysis kit on the Fragment Analyzer (Agilent Technologies). After standardizing the amount of cDNA per sample, the libraries were sequenced on Illumina NextSeq500 sequencing machine with 75-bp single-end reads. mRNA library preparation and sequencing was done at the Genome Center in the University of Texas at Dallas Research Core Facilities.

Mapping and TPM quantification:

RNA-seq read files (fastq files) were checked for quality by FastQC (Babraham Bioinformatics, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and read trimming was done based on the Phred score and per-base sequence content. Trimmed Reads were then mapped against the reference genome and transcriptome (Gencode vM16 and GRCm38.p5) using STAR v2.2.1 (5). Relative abundances in Transcripts Per Million (TPM) for every gene of every sample was quantified by stringtie v1.3.5 (6). Non-coding genes and mitochondrial genes were removed from the analysis (based on Gencode annotation) and the TPMs for nonmitochondrial coding genes were re-normalized to sum to 1 million.

Order statistics and re-normalization of expression data

In order to identify a set of consistently expressed genes in the transcriptome (INPUT) samples, percentile ranks were calculated on TPMs for each coding gene for each sample. We conservatively chose 15,072 genes that were above the 30th percentile in each INPUT sample, for at least one sex, to be in the set of genes considered consistently detected in the transcriptome. Quantile normalization was then performed based on the set of all coding genes.

The IP (translatome) analysis was only performed for the 15,072 consistently transcriptomeexpressed genes. In order to identify a set of consistently expressed genes in the translatome (IP) samples, percentile ranks were calculated on TPMs for each of the 15,072 consistently transcriptome-expressed coding genes for each sample. We then chose 12,542 genes out of those 15,072 genes to be consistently detected in the translatome based on whether their expression was on or above the 15th percentile (out of 15,072 genes) in each IP sample, for at least one sex.

The percentile thresholds for choosing consistently transcriptome-expressed and translatomeexpressed genes were conservatively estimated by identifying thresholds that would eliminate genes with consistently low or no detected reads.

Differential expression analysis

We first calculated the log2-fold change (based on median TPMs) for each consistently transcriptome-expressed coding gene in the INPUT, and for each consistently translatomeexpressed coding gene in the IP. It is calculated as follows:

$$
\text{LFC}_i^{(T)} = log_2\left(\frac{\mu_{i,f}^{(T)} + 0.01}{\mu_{i,m}^{(T)} + 0.01}\right)
$$

where for gene i, LFC(T)i is the log2-fold change, $\mu(T)$ i, fand $\mu(T)$ i, m are the median TPMs in females and males respectively, with 0.01 as the smoothing factor.

We used strictly standardized mean difference (SSMD) (6, 7) to discover genes with systematically altered expression percentile ranks between males and females. SSMD is the difference of means controlled by the variance of the sample measurements. We used SSMD as a measure of effect size since it is appropriate for smaller sample sizes while simultaneously controlling for within-group variability. It is calculated as follows:

$$
SSMD_i^{(P)} = \frac{\mu_{i,f}^{(P)} - \mu_{i,m}^{(P)}}{\sqrt{v_{i,f}^{(P)} + v_{i,m}^{(P)}} + 0.01}
$$

where for gene i, SSMD(P)i is the strictly standardized mean difference, with $\mu(P)$ i, f, v(P)i, f and $\mu(P)$ i,m, $\nu(P)$ i,f are the means and variances of gene TPM percentile ranks in females and males respectively, under the assumption that covariance is 0, and with 0.01 as the smoothing factor.

For calculating the overlap in distribution between the qTPMs, in animals sampled from different sexes, we further calculated Bhattacharyya distance (8), which is used to calculate the amount of overlap in the area under the curve of the two sample distributions (corresponding to each sex) in order to identify the best candidates for the DE gene set. Unlike SSMD, BD does not make assumptions of equal variance in the two compared samples, and thus, is useful for comparing distributions of gene relative abundance (in TPMs). It is calculated as follows:

$$
BD_i^{(Q)} = \frac{1}{4} \left(\frac{(\mu_{i,f}^{(Q)} - \mu_{i,m}^{(Q)})^2}{v_{i,f}^{(Q)} + v_{i,m}^{(Q)} + 0.01} \right) + \frac{1}{4} \ln \left(\frac{1}{4} \left(\frac{v_{i,f}^{(Q)}}{v_{i,m}^{(Q)} + 0.01} + \frac{v_{i,m}^{(Q)}}{v_{i,f}^{(Q)} + 0.01} + 2 \right) \right)
$$

where for gene i, BD(Q)i is the Bhattacharyya Distance over gene qTPMs, with μ(Q)i,f , v(Q)i,f and $\mu(Q)$ i, m, $\nu(Q)$ i, f are the means and variances of gene qTPMs in females and males respectively, and with 0.01 as the smoothing factor. The Bhattacharyya coefficient BC(Q)i ranges between 0 (for totally non-overlapping distributions) and 1 (for completely identical distributions) and is derived from the Bhattacharyya distance as follows:

$$
BC_i^{(Q)} = e^{-BD_i^{(Q)}}
$$

In our analysis, we used a modified form of the Bhattacharyya coefficient that ranges between 0 (for completely identical distributions) and +1 or -1 (for totally non-overlapping distributions, sign defined by the log-fold change value). It is calculated as follows:

$$
(1 - \mathrm{BC}_i^{(Q)}) \times \mathrm{sgn}(\mathrm{LFC}_i^{(T)})
$$

Motif analysis

For each gene in the up-regulated translation gene list (IP fraction), we created one 5′ UTR and 3' UTR sequence per gene. For each gene, this was done by first extracting the respective UTR sequences from all isoforms from BioMart (7). Isoforms with UTRs <20 bases, or with sequence that is part of an ORF in another isoform were discarded. The remaining isoforms were collapsed into a single sequence in order of their location in the genome, with overlapping regions present exactly once in the sequence. The motif finding tool MEME (8) was used to identify motifs.

Network of interactions

We used STRING database (9) and Cytoscape (10) to generate Figure 3G and visualize interactions between genes differentially expressed and DRG enriched genes.

Single cell data

Single cell mouse DRG sequencing data from previously published work (11) was used to generate Figure 5B. Seurat package 2.2.1 (12) was used to cluster the single-cell data and visualization (t-SNE) (13).

Statistics

Statistical analysis for behavior, image quantification and ELISA assay was done in GraphPad Prism 8. Data visualization was done in Python (version 3.7 with Anaconda distribution). All data are represented as mean \pm SEM. Single comparisons were performed using Student's *t* test, and multiple comparisons were performed using a one-way or two-way ANOVA with *post hoc* tests for between-group comparisons. Statistical results can be found in the figure legends.

Coding for bioinformatics analysis and data visualization was done in Python (version 3.7 with Anaconda distribution).

Figure 1 and other schematic drawings were generated using Biorender (BioRender.com).

Data availability

Sequencing raw data files from this study are available in the NCBI GEO database under the accession number GSE155676.

Supplemental References

1. Stirling LC, Forlani G, Baker MD, Wood JN, Matthews EA, Dickenson AH, et al. (2005): Nociceptor-specific gene deletion using heterozygous NaV1. 8-Cre recombinase mice. *Pain*. 113:27-36.

2. McLean AC, Valenzuela N, Fai S, Bennett SAL (2012): Performing vaginal lavage, crystal violet staining, and vaginal cytological evaluation for mouse estrous cycle staging identification. *J Vis Exp*.e4389-e4389.

3. Langford DJ, Bailey AL, Chanda ML, Clarke SE, Drummond TE, Echols S, et al. (2010): Coding of facial expressions of pain in the laboratory mouse. *Nat Methods*. 7:447-449.

4. Chaplan SR, Bach FW, Pogrel JW, Chung JM, Yaksh TL (1994): Quantitative assessment of tactile allodynia in the rat paw. *J Neurosci Methods*. 53:55-63.

5. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. (2013): STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 29:15-21.

6. Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL (2015): StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature biotechnology*. 33:290-295.

7. Kinsella RJ, Kähäri A, Haider S, Zamora J, Proctor G, Spudich G, et al. (2011): Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database (Oxford)*. 2011:bar030.

8. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. (2009): MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res*. 37:W202-208.

9. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. (2019): STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids research*. 47:D607- D613.

10. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. (2003): Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*. 13:2498-2504.

11. Li CL, Li KC, Wu D, Chen Y, Luo H, Zhao JR, et al. (2016): Somatosensory neuron types identified by high-coverage single-cell RNA-sequencing and functional heterogeneity. *Cell Res*. 26:967.

12. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R (2018): Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 36:411-420.

13. Maaten Lvd, Hinton G (2008): Visualizing data using t-SNE. *Journal of machine learning research*. 9:2579-2605.