SUPPLEMENTARY INFORMATION

Genetically-Regulated Gene Expression in the Brain Associated With Chronic Pain: Relationships With Clinical Traits and Potential for Drug Repurposing

Johnston *et al.*

S-PrediXcan

Transcriptomic Imputation (TI) involves using a reference dataset where gene expression (RNA-seq) and genotype data is available for the same individuals (e.g. the Genotype-Tissue Expression (GTEx) project) to build tissue-specific predictor models. These predictor models can then be used to predict (impute) genetically-regulated gene expression (GREX) in a separate cohort of individuals where genotyping data is available. These imputed GREX values can then be tested for association with a trait of interest in a transcriptome-wide association study (TWAS). A range of different TI methods are available, of which S-PrediXcan is one. S-PrediXcan (1) is an extension of PrediXcan (2) where individual-level data is not needed– PrediXcan models use elastic net regression to choose and weight SNPs associated with gene expression to include in any given GREX predictor model, which are then used to predict the transcriptome. Then regression coefficients of the trait of interest on imputed GREX for each gene is calculated. In contrast, S-PrediXcan directly computes the gene-level association results, using GWAS summary statistic output. We compare the formulae for PrediXcan and S-PrediXcan (adapted from (1)) below to illustrate this difference.

$$T_g = \sum_{l \in Model} w_{lg} X_l$$

Formula 1: PrediXcan calculation of predicted transcriptome (TI). W_{lg} = predictor model weights, X_l = genotype.

$$Y = T_g \gamma + \varepsilon$$

Formula 2: A second step and phenotype information is needed to perform TWAS (test for trait Y association with predicted transcriptome) and give gene-level association results. T_g = imputed transcriptome.

$$Z_g \approx \sum_{l \in Model} w_{lg} \frac{\sigma_L}{\sigma_G} \frac{\beta_L}{SE(\beta)}$$

Formula 3: S-PrediXcan directly computes gene level association results (Z_g) without need for individual-level information. $\frac{\sigma_L}{\sigma_G}$ = reference genotype dataset (e.g. 1000 Genomes), which replaces the need for

genotype data (X_l in formula 1), $\frac{\beta_L}{SE(\beta)}$ = GWAS summary results Z scores, replacing individual-level phenotype information.

FUMA GENE2FUNC

FUMA (Functional Mapping and Annotation of Genome-Wide Association Studies) (3) is a web-based suite of tools for use in characterizing and prioritizing findings from GWAS summary statistics. The subset of tools within GENE2FUNC, a division of FUMA, specifically takes gene-level results as input, and this is what we use in this study to further characterize our 89 unique significant gene findings from S-PrediXcan analysis. Gene set enrichment refers to testing for overrepresentation of classes of genes compared to what would be expected by chance. FUMA performs hypergeometric tests to test if genes of interest (our 89 gene findings) are overrepresented in pre-defined gene sets. These gene set definitions are from MsigDB (4,5), Wikipathways (6), and reported genes from the GWAS catalog (7), and multiple testing correction is performed within category. Hypergeometric testing involves using the hypergeometric distribution to calculate probability of 'successes' from a specific sample size, approximating to a Fisher's exact test. In practice this can be viewed as a 2x2 table:

	Not MCP-GREX gene	MCP-GREX gene
In gene set	A	C
Not in gene set	B	D

Whether the value C (number of our gene findings that are in the gene set, e.g., positional gene set chr3p21) is an overrepresentation (enrichment) can then be determined through calculating a p value using the hypergeometric distribution. Choosing 'background genes' as part of FUMA input, as described in the main manuscript, allows us to populate the first column. For FUMA GENE2FUNC tutorial see also https://fuma.ctglab.nl/tutorial#gene2func .

DrugBank

Another tool available in FUMA GENE2FUNC is automatic searches of DrugBank (9), giving information on whether genes of interest are also annotated as drug targets, and if so which medications or compounds are associated with each gene. DrugBank is provides information of FDA-approved drugs in addition to drugs currently undergoing FDA approval processes, and as of 2018 more than 4,000 drug targets (including proteins, DNA, and RNA) are included in the database. FUMA assigns DrugBank drug identification number(s) to gene results if the UniProt ID (i.e., protein product of the gene) is listed as one of the targets of the drug.

Connectivity Map (CMap) Analysis

To query CMap L1000 (gene expression) and search for perturbation signatures (e.g., gene expression patterns) with significant connectivity score, two lists of genes must be prepared. The first list is genes up-regulated in the trait of interest, and the second is genes down-regulated in the trait of interest. Query results include results for all perturbagens (i.e., not just drug compounds), and for all connectivity score magnitude (i.e., non-significant results). We retained results for perturbation type 'compound', that passed internal CMap quality control, that had a connectivity score significantly different from the null, and that had a known mechanism of action/ non-missing mechanism of action information. A connectivity score represents the similarity between perturbation signatures, and ranges from -1 to 1. A perturbagen with a raw connectivity score of 1 therefore 'matches' MCP associated gene expression changes perfectly, and a perturbagen with a negative connectivity score suggests the compound may have therapeutic potential in chronic pain as it represents an 'opposite' gene expression change profile to chronic pain.

Input gene lists for CMap query may be reduced before query submission, as genes must have valid Entrez IDs and be present within the database search space. In the table below we indicate genes in the up

and down-regulated lists and whether they were included in the final query. All genes except one (NUP43) were included in a single list as direction of effect (up or down regulation) was consistent across tissues in S-PrediXcan analyses. NUP43 was both up and down-regulated, and so was excluded from the query.

PheWAS

Phenome-wide association study (pheWAS) is a method for testing for association between a trait of interest, which can be genetic, gene-expression based, or phenotypic, and many phenotypic traits simultaneously (the phenome). This can be thought of as a 'reverse' of GWAS, with each individual component of the phenome representing a single SNP, and the trait of interest (e.g., MCP-GREX) representing the GWAS phenotype. In our analyses we impute MCP-GREX in 13 brain tissues and whole blood for BioMe participants, and test for association between this value and our phenome (consisting of 1,000+ phecodes), using the pheWAS R package (10) with adjustment for age, sex, and the first five genotype-derived principal components. For recent review of pheWAS methods in general see (11).

Summary Transcriptome Wide Association (S-TWAS) Analysis

To attempt replication of our S-PrediXcan findings, we carried out S-TWAS (12). S-TWAS is a different method by which elastic net models are applied to summary statistics to identify genes where cis-regulated expression is associated with complex traits. This was carried out using the FUSION software package (Gusev et al), 1000 Genomes reference data, and the pre-computed predictive models for GTEx v8 brain tissue and whole blood available at http://gusevlab.org/projects/fusion/ . We included models for genes with significant heritability and for all populations as recommended (see http://gusevlab.org/projects/fusion/ section: GTEx v8 multi-tissue expression) and retained results where the model applied was an elastic net and genes were not within the MHC region. We then applied Bonferroni correction within-tissue as in the main analysis, finding 97 unique gene-tissue associations comprised of 57 unique genes. 43/229 of our significant S-PrediXcan gene-tissue associations have an available gene-tissue model in S-TWAS; of these, we replicate 39 / 43 gene-tissue associations at the tissue-wide significance level, all with consistent direction of effect between the two analyses. At the experiment-wide significance level, 17 /18 experiment-wide-significant gene-tissue associations are replicated in S-TWAS analyses. We carried out a Fisher's exact test that indicated significant levels of enrichment of S-PrediXcan gene-tissue associations in S-TWAS results ($p < 2.2x10^{-16}$).

Brain regions implicated in pain

The 'neuromatrix' a.k.a. 'pain matrix', first described by Melzack (13), encompasses a large network of brain regions that respond to painful stimuli – this network includes the insula, thalamus, somatosensory cortex, and the anterior cingulate cortex. However, pain processing, and the development of chronic pain, have been associated with a much wider range of brain regions. It is also important to note that many studies in pain and the brain include study of nociception, of noxious stimuli responses, of acute and potentially induced pain, of emotional and psychological factors associated with pain and chronic pain, the study of specific chronic pain condition states, and studies in rodents. Recent detailed review can be found at (14–18), along with recent IASP redefinition of pain to emphasize its distinction from nociception alone (19).

Noxious stimuli result in transmission from nociceptors to the dorsal horn through primary afferent Adelta and C fibres, with this information eventually making its way to the dorsal horn, then on to the thalamus, then eventually the somatosensory cortex and periaqueductal gray (reviewed by (20)). Already, 'pain' regions of the brain, even just those associated with nociception (a small factor in the chronic pain experience), can then be listed as the dorsal horn, thalamus, all regions in the somatosensory cortex (Brodmann's areas 3a, 3b, 1 and 2), and the periaqueductal gray. Nociceptive information is then transmitted to higher-order regions associated with affective aspects of pain, and with memory, including the amygdala (21), hypothalamus (22), anterior cingulate cortex (23), prefrontal cortex (24), and nucleus

accumbens (25). Other higher brain regions implicated in pain processing, pain perception, and emotional aspects of (acute) pain, include the insular cortex (26), cerebellum (27), and the ventral tegmental area (28,29).

Chronic and persistent pain more specifically are also associated with many different brain regions. The cerebellum has been implicated in imaging studies of migraine (30), and connectivity between the hypothalamus and limbic system was found to be altered in fibromyalgia (31). Across a range of chronic pain conditions, regions including the prefrontal cortex, anterior cingulate cortex, insula, and thalamus were found to show decreased grey matter volume (reviewed by (32)), and development of chronic low back pain has been associated with shift from nociceptive to emotional brain network involvement in imaging studies (33). In studies of mice changes in gene expression in the insula, hippocampus, and amygdala were associated with endometriosis (34), and in regions including the nucleus accumbens and putamen changes in gray matter volume were associated with Complex Regional Pain Syndrome (CRPS) (35).

Finally, changes and disruption across networks spanning the entire brain have also been implicated in chronic pain (reviewed by (17)), including reorganization of the default mode network across different chronic pain conditions (36). Although GTEx contains samples from 13 brain regions total, all of which have been previously implicated in pain, if not chronic pain, increasing the number of different brain regions sampled would provide greater insight when studying chronic pain, in line with the wide range of brain regions and widespread disruption associated with chronic pain.

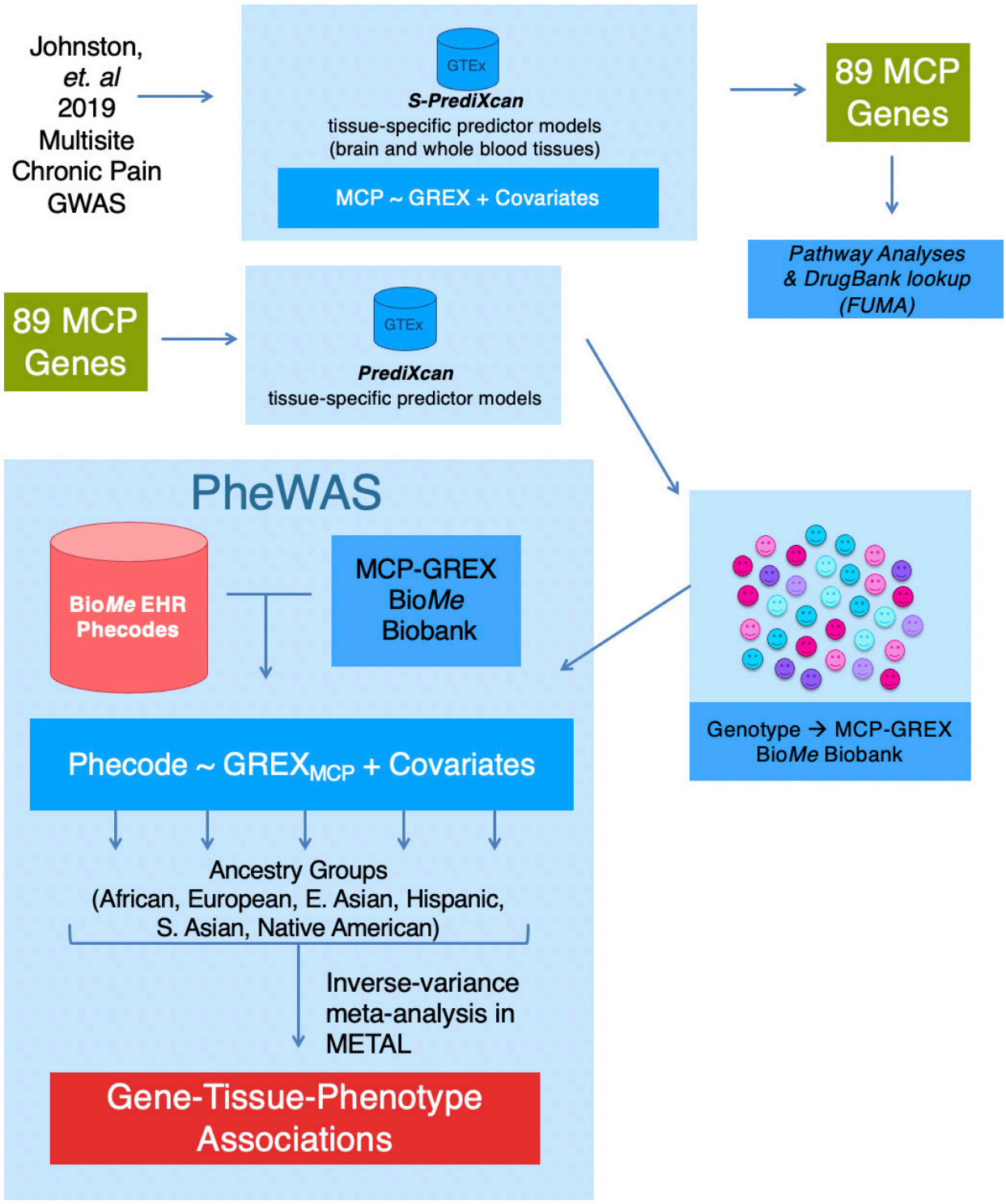

Figure S1: Schematic showing how analyses were combined.

Tissue	p-value Threshold
Amygdala	1.79×10^{-05}
Anterior cingulate cortex BA24	1.41×10^{-05}
Caudate basal ganglia	9.99×10^{-06}
Cerebellar Hemisphere	8.69×10^{-06}
Cerebellum	7.36×10^{-06}
Cortex	9.09×10^{-06}
Frontal Cortex BA9	1.10×10^{-05}
Hippocampus	1.36×10^{-05}
Hypothalamus	1.37×10^{-05}
Nucleus accumbens basal ganglia	1.03×10^{-05}
Putamen basal ganglia	1.13×10^{-05}
Spinal cord cervical c-1	1.54×10^{-05}
Substantia nigra	1.95×10^{-05}
Whole Blood	6.89×10^{-06}

Table S1: Within-tissue p-value Bonferroni significance thresholds.

up-regulated genes	down-regulated genes
ECM1	**SNRPC**
TARS2	**SEMA3B**
GPX1	**AMT**
GMPPB	**VPS33B**
CELSR3	RP11-24H2.3
NMT1	**FUBP1**
RPRD2	**CSK**
C6orf106 (ILRUN)	ZNF501
UHRF1BP1	**RBM6**
SDCCAG8	**MST1**
SUOX	**RPS26**
RNF123	MRPS21
GPR27	**P4HTM**
CEP170	**INTS1**
SP4	RP11-160H22.5
MON1B	**ZNF197**
PTK2	**CTBP2**
SLC25A13	**SEMA3F**

PRKAR2A	**NUDT18**
UFL1	**UBA7**
SCAMP2	C15orf57
TSKU	**ZNF35**
LANCL1	**MAU2**
GRK4	**ACADL**
MST1R	**PACSIN3**
KLHDC8B	**UBOX5**
TSPYL4	**HEXIM1**
LIN28B-AS1	**KCNH2**
MPI	ERICH2
KNDC1	**LATS1**
GINM1	**DCAKD**
FASTKD5	AC007405.6
NELFA	**LLGL1**
PPP6C	SLC38A3
ZNF23	ZNF502
DNAH11	**RAD51**
BAK1	**URM1**
IL23A	FAM180B
DNMT3B	ACSF3
SCAI	RP11-147L13.11
CDK14	LINC01671
KIF3B	**CYB561D2**
RP11-147L13.8	**COX11**
S100A1	
NUP43	NUP43
SHMT1	

Table S2: Genes included in CMap query are shown in bold.

Ancestry Group	N_{min}	N_{max}
East Asian	732	846
African American	6105	7514
Southeast Asian	477	572
Hispanic American	8373	10324
Native American	69	83
European American	8262	9483

Table S3: Minimum and maximum sample sizes per ancestry group included in PheWas analyses. Minimum and maximum sample sizes refer to the MCP-GREx-phecode association testing minimum and maximum number of participants. This varies as exclusion criteria vary per phecode.

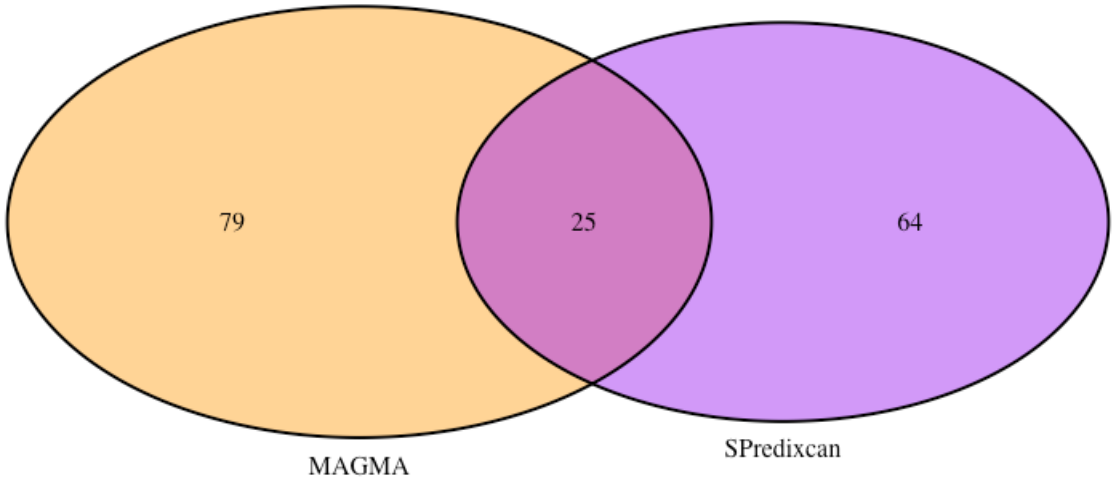

Figure S2: Venn diagram of number of genes shared and distinct in S-PrediXcan results compared to MAGMA results.

MAGMA	Shared	S-PrediXcan
BAI2	MRPS21	CELSR3
PABPC4	RPRD2	SEMA3B
FAM212B	TARS2	AMT
PRPF3	ECM1	RP11-24H2.3
RABGAP1L	CEP170	FUBP1
AC092782.1	SDCCAG8	C6orf106 (ILRUN)
NRXN1	LANCL1	CSK
SLC4A10	GPX1	ZNF501
CPS1	RNF123	SUOX
ERBB4	GMPPB	MST1
SPHKAP	UBA7	RPS26
EOMES	MST1R	GPR27
DHX30	RBM6	P4HTM
ARIH2OS	SEMA3F	INTS1
LAMB2	SNRPC	RP11-160H22.5
CCDC71	UHRF1BP1	MON1B
C3orf84	NUP43	ZNF197
CCDC36	SP4	CTBP2
RP11-3B7.1	SLC25A13	PTK2
RHOA	NUDT18	PRKAR2A
TCTA	TSKU	UFL1
NICN1	VPS33B	SCAMP2
DAG1	DCAKD	GRK4
BSN	NMT1	C15orf57
AMIGO3	KIF3B	KLHDC8B
IP6K1		TSPYL4
CDHR4		ZNF35
TRAIP		LIN28B-AS1
CAMKV		MAU2
MON1A		ACADL
RBM5		PACSIN3
GNAT1		MPI
EIF4E3		KNDC1
ROBO2		GINM1
BBX		FASTKD5
MSL2		UBOX5
PCCB		HEXIM1

STAG1		KCNH2
MAML3		NELFA
GABRB2		ERICH2
UQCC2		LATS1
IP6K3		AC007405.6
C6orf106		PPP6C
FHL5		LLGL1
LIN28B		ZNF23
FYN		SLC38A3
LAMA2		ZNF502
KATNA1		DNAH11
SDK1		RAD51
GRM3		URM1
FOXP2		BAK1
FAM120A		FAM180B
PHF2		IL23A
ASTN2		DNMT3B
DNM1		SCAI
EXD3		CDK14
MLLT10		ACSF3
ZRANB1		RP11-147L13.8
JAKMIP3		RP11-147L13.11
NCAM1		LINC01671
RERG		CYB561D2
EFNB2		S100A1
NUMB		COX11
PPP1R13B		SHMT1
ATXN1L		
IST1		
ZNF821		
ACBD4		
HEXIM2		
C17orf58		
ASXL3		
DCC		
ILF3		
ATP13A1		
ZNF101		
SLC24A3		

TM9SF4		
ASXL1		
C20orf112		

Table S4: Genes shared and distinct in MAGMA results compared to S-PrediXcan results.

ENSGID	Gene Symbol	Tissue	Z (S-Pred.)	Z (S-TWAS)	P Raw (S-Pred.)	P Adj. (S-Pred.)	P Raw (S-TWAS)	P Adj. (S-TWAS)
ENSG00000197728	RPS26	Amygdala	-4.71	-4.8	2.53E-06	7.04E-03	1.57E-06	9.66E-04
ENSG00000136448	NMT1	Anterior_cingulate_cortex_BA24	5.89	6.03	3.83E-09	1.35E-05	1.60E-09	1.39E-06
ENSG00000197728	RPS26	Anterior_cingulate_cortex_BA25	-4.84	-4.82	1.27E-06	4.50E-03	1.43E-06	1.24E-03
ENSG00000178802	MPI	Anterior_cingulate_cortex_BA26	4.5	4.48	6.76E-06	2.39E-02	7.62E-06	6.63E-03
ENSG00000166260	COX11	Anterior_cingulate_cortex_BA27	-4.39	-4.39	1.11E-05	3.92E-02	1.15E-05	1.00E-02
ENSG00000173540	GMPPB	Caudate_basal_ganglia	6.15	6.43	7.77E-10	3.88E-06	1.31E-10	1.77E-07
ENSG00000197728	RPS26	Caudate_basal_ganglia	-4.94	-4.92	7.80E-07	3.90E-03	8.48E-07	1.15E-03
ENSG00000173540	GMPPB	Cerebellar_Hemisphere	6.07	5.92	1.30E-09	7.49E-06	3.28E-09	5.62E-06
ENSG00000197728	RPS26	Cerebellar_Hemisphere	-4.86	-4.94	1.15E-06	6.62E-03	7.83E-07	1.34E-03
ENSG00000115365	LANCL1	Cerebellar_Hemisphere	4.49	4.75	6.97E-06	4.00E-02	1.99E-06	3.41E-03
ENSG00000233276	GPX1	Cerebellum	5.23	6.53	1.69E-07	1.15E-03	6.53E-11	1.40E-07
ENSG00000197728	RPS26	Cerebellum	-5.15	-5.02	2.58E-07	1.75E-03	5.17E-07	1.11E-03
ENSG00000139531	SUOX	Cerebellum	5.03	5.15	4.88E-07	3.31E-03	2.63E-07	5.62E-04
ENSG00000164068	RNF123	Cerebellum	4.94	5.05	7.94E-07	5.38E-03	4.37E-07	9.34E-04
ENSG00000266472	MRPS21	Cerebellum	-4.92	-4.64	8.86E-07	6.01E-03	3.49E-06	7.46E-03
ENSG00000172992	DCAKD	Cerebellum	-4.52	-4.43	6.25E-06	4.24E-02	9.33E-06	1.99E-02
ENSG00000173540	GMPPB	Cortex	5.9	6.32	3.53E-09	1.94E-05	2.66E-10	4.10E-07
ENSG00000197728	RPS26	Cortex	-4.9	-4.9	9.78E-07	5.37E-03	9.44E-07	1.45E-03
ENSG00000115365	LANCL1	Cortex	4.8	5.66	1.59E-06	8.74E-03	1.51E-08	2.33E-05
ENSG00000233276	GPX1	Frontal_Cortex_BA9	6.25	6.29	4.03E-10	1.84E-06	3.10E-10	3.79E-07
ENSG00000197728	RPS26	Frontal_Cortex_BA10	-5.28	-4.94	1.30E-07	5.91E-04	7.66E-07	9.37E-04
ENSG00000172992	DCAKD	Frontal_Cortex_BA11	-4.58	-4.21	4.68E-06	2.13E-02	2.56E-05	3.13E-02
ENSG00000267731	RP11-147L13.8	Frontal_Cortex_BA12	4.47	4.58	7.96E-06	3.63E-02	4.63E-06	5.66E-03
ENSG00000166260	COX11	Hippocampus	-4.37	-4.54	1.27E-05	4.68E-02	5.70E-06	4.83E-03
ENSG00000197728	RPS26	Hypothalamus	-4.87	-4.93	1.11E-06	4.06E-03	8.29E-07	6.53E-04
ENSG00000143369	ECM1	Nucleus_accumbens_basal_ganglia	5.43	5.46	5.70E-08	2.76E-04	4.79E-08	6.35E-05
ENSG00000233276	GPX1	Nucleus_accumbens_basal_ganglia	4.96	5.24	7.03E-07	3.41E-03	1.57E-07	2.08E-04
ENSG00000128891	C15orf57	Nucleus_accumbens_basal_ganglia	-4.76	-4.4	1.93E-06	9.36E-03	1.09E-05	1.44E-02

ENSG00000197728	RPS26	Putamen_basal_ganglia	-5.24	-4.9	1.64E-07	7.28E-04	9.52E-07	1.07E-03
ENSG00000173540	GMPPB	Spinal_cord_cervical_c-1	5.97	5.81	2.40E-09	7.80E-06	6.20E-09	4.69E-06
ENSG00000197728	RPS26	Substantia_nigra	-4.54	-4.61	5.54E-06	1.42E-02	3.96E-06	2.05E-03
ENSG00000173540	GMPPB	Whole_Blood	5.95	5.79	2.70E-09	1.95E-05	7.10E-09	2.15E-05
ENSG00000065060	UHRF1BP1	Whole_Blood	5.04	4.79	4.71E-07	3.42E-03	1.66E-06	5.03E-03
ENSG00000004864	SLC25A13	Whole_Blood	4.94	5.05	7.66E-07	5.55E-03	4.37E-07	1.32E-03
ENSG00000197728	RPS26	Whole_Blood	-4.92	-4.8	8.54E-07	6.19E-03	1.61E-06	4.88E-03
ENSG00000004534	RBM6	Whole_Blood	-4.86	-4.85	1.15E-06	8.30E-03	1.21E-06	3.66E-03
ENSG00000055211	GINM1	Whole_Blood	4.66	4.77	3.17E-06	2.30E-02	1.85E-06	5.60E-03
ENSG00000172992	DCAKD	Whole_Blood	-4.58	-4.51	4.75E-06	3.44E-02	6.35E-06	1.92E-02
ENSG00000167118	URM1	Whole_Blood	-4.52	-4.49	6.32E-06	4.58E-02	7.08E-06	2.14E-02

Table S5: Tissue-wide significant gene-tissue findings in both S-PrediXcan and S-TWAS. Z (S-Pred.) = Z score (S-PrediXcan), Z (S-TWAS) = Z score (S-TWAS), P Raw (S-Pred.) = unadjusted association p-value (S-PrediXcan), P Adj. (S-Pred.) = Bonferroni-corrected (within-tissue) p value (S-PrediXcan), P Raw (S-TWAS) = unadjusted association p value (S-TWAS), P Adj. (S-TWAS) = Bonferroni-corrected (within-tissue) p value (S-TWAS).

	Significant (S-PrediXcan)	**NS (S-PrediXcan)**	TOTAL
Significant (S-TWAS)	17	17	34
NS (S-TWAS)	1	11823	11824
TOTAL	18	11840	11858

Table S6: Contingency table for Fisher's exact test. Significant = gene-tissue results with a p value less than the experiment-wide p value threshold for that analysis (S-PrediXcan or S-TWAS). NS = Non-significant gene-tissue results. Note a total of 11, 858 gene-tissue predictor models are shared between S-PrediXcan and S-TWAS, and this makes up the background in this test.

	Significant MCP	NS MCP	Totals
Significant Toikumo et al	6	32	38
NS Toikumo et al	37	11783	11820
Totals	43	11815	11858*

Table S7: Contingency table for Fisher's exact test. *Total number of gene-tissue associations with models available in both FUSION and S-PrediXcan. MCP = Multisite Chronic Pain (S-PrediXcan analyses), NS = non-significant. Thirty-eight gene-tissue associations from Toikumo et al have prediction models available in both S-PrediXcan and FUSION, and 43 of 229 significant gene-tissue associations for MCP in our analyses have a prediction model in both S-PrediXcan and FUSION.

ENSGID	Gene	Tissue	SPrediXcan_Z	SPrediXcan_P	FUSION_Z	FUSION_P
ENSG00000173540	GMPPB	Brain_Cerebellar_Hemisphere	6.066939899	1.30E-09	4.92625	8.38E-07
ENSG00000173540	GMPPB	Brain_Cortex	5.904840509	3.53E-09	4.9262	8.38E-07
ENSG00000233276	GPX1	Brain_Frontal_Cortex_BA9	6.252951146	4.03E-10	6.6838	2.33E-11
ENSG00000164068	RNF123	Brain_Cerebellum	4.936958906	7.94E-07	8.67813	4.02E-18
ENSG00000197728	RPS26	Brain_Anterior_cingulate_cortex_BA24	-4.844020429	1.27E-06	-4.3525	1.35E-05
ENSG00000197728	RPS26	Brain_Cortex	-4.896032857	9.78E-07	-4.49495	6.96E-06

Table S8: Significant gene-tissue results in MCP S-PrediXcan analyses also significant in FUSION TWAS by Toikumo et al.

Symbol	DrugBank Accession ID	Drug Name & Description
AMT	DB00116	tetrahydrofolic acid, nutritional supplement
	DB00157	NADH, nutritional supplement (some evidence for PD, CFS, AD, CVD benefit)
	DB04789	5-methyltetrahydrofolic acid, nutritional supplement
CSK	DB01254	dasatinib, tyrosine kinase inhibitor, cancer treatment (leukemia)
	DB02010	staurosporine, protein kinase C inhibitor
	DB05075	TG-100801, topically applied kinase inhibitor (macular degeneration)
	DB12010	fostamatinib, spleen tyrosine kinase inhibitor (thrombocytopenia)
FUBP1	DB05786	irofulven, novel anti-cancer compound
GPX1	DB00143	glutathione, nutritional supplement
IL23A	DB05459	briakinumab, anti-IL-12 monoclonal antibody for T-cell-driven autoimmune disease treatment
	DB11834	guselkumab, monclonal antibody for plaque psoriasis
KCNH2	DB00176	fluvoxamine, SSRI for OCD
	DB00199	erythromycin, macrolide antibiotic
	DB00204	dofetilide, class 3 antiarrhythmic
	DB00276	amsacrine, cytotoxin for leukemia treatment
	DB00280	disopyramide, class 1A antiarrhythmic
	DB00308	ibutilide, class 3 antiarrythmic

	DB00342	terfenadine, antihistamine
	DB00346	alfuzosin, alpha-1 adrenergic antagonist
	DB00455	loratidine, 2nd generation antihistamine
	DB00457	prazosin, alpha-blocker for hypertension
	DB00458	imipramine, tricyclic antidepressant
	DB00472	fluoxetine, SSRI
	DB00477	chlorpromazine, phenothiazine antipsychotic
	DB00489	sotalol, methane sulfoanilide beta adrenergic antagonist for arrhythmia
	DB00537	ciprofloxacin, second generation fluoroquinolone
	DB00590	doxazosin, alpha-1 adrenergic receptor for hypertension
	DB00604	cisapride, GERD-associated heartburn medication
	DB00637	astemizole, second generation antihistamine
	DB00661	verapamil, non-dihydropyridine calcium channel blocker for angina, arrhythmia, hypertension
	DB00675	tamoxifen, selective estrogen receptor modulator used in certain breast cancers
	DB00679	thioridazine, phenothiazine antipsychotic for GAD and schizophrenia
	DB00908	quinidine, arrhythmia treatment
	DB01026	ketoconazole, broad spectrum antifungal
	DB01035	procainamide, arrhythmia treatment
	DB01074	perhexiline, coronary vasodilator
	DB01100	pimozide, antipsychotic used in Tourette's
	DB01110	miconazole, azole antifungal
	DB01118	amiodarone, class 3 antiarrhythmic
	DB01136	carvedilol, non-selective beta-adrenergic antagonist
	DB01142	doxepin, psychotropic agent with antidepressant and anxiolytic properties
	DB01149	nefazodone, antidepressant
	DB01162	terazosin, alpha-1 adrenergic antagonist
	DB01182	propafenone, class 1c antiarrhythmic
	DB01195	flecainide, class 1c antiarrhythmic
	DB01211	clarithromycin, macrolide antibiotic
	DB01218	halofantrine, antimalarial
	DB01244	bepridil, calcium channel blocker
	DB04855	dronedarone, antiarrhythmic
	DB04957	azimilide, investigational class 3 antiarrhythmic
	DB06144	sertindole, atypical antipsychotic
	DB06217	vernakalant, antiarrhythmic
	DB06457	tecastemizole, investigational small molecule
	DB11090	potassium nitrate, small wound cauterization
	DB11186	pentoxyverine, cough suppressant

	DB11386	chlorobutanol, alcohol-based perservative
	DB11633	Isavuconazole, triazole antifungal
	DB11642	pitolisant, antagonist of histamine H3 receptor, narcolepsy treatment
LATS1	DB12010	fostamatinib, spleen tyrosine kinase inhibitor (thrombocytopenia)
MST1R	DB12010	fostamatinib, spleen tyrosine kinase inhibitor (thrombocytopenia)
NMT1	DB03062	investigational small molecule
P4HTM	DB00126	vitamin C, nutritional supplement
PRKAR2A	DB05798	GEM-231, monoclonal antibody
PTK2	DB06423	endostatin, investigational small molecule
	DB07248	investigational small molecule
	DB07460	investigational small molecule
	DB12010	fostamatinib, spleen tyrosine kinase inhibitor (thrombocytopenia)
RAD51	DB04395	phosphoaminophosphonic acid-adenylate ester, investigational small molecule
	DB12742	amuvatinib, cancer treatment undergoing clinical trial
S100A1	DB00768	olopatadine, histamine H1 antagonist
SHMT1	DB00114	pyridoxal phosphate (B6), nutritional supplement
	DB00116	tetrahydrofolic acid, nutritional supplement
	DB00145	glycine, total parenteral nutrition component
	DB01055	mimosine, antineoplastic
	DB02067	triglu-5-formyl-tetrahydrofolate, investigational small molecule
	DB02800	5-hydroxymethyl-5,6-dihydrofolic acid, investigational small molecule
	DB02824	N-pyridoxyl-glycine-5-monophosphate, investigational small molecule
SLC25A13	DB00128	aspartic acid, total parenteral nutrition component
SLC38A3	DB00117	histidine, total parenteral nutrition component
	DB00174	asparagine, non-essential amino acid
SUOX	DB03983	investigational small molecule
TARS2	DB00156	threonine, total parenteral nutrition component

Table S9: DrugBank lookup results for 89 significant MCP-GREX genes. PD = Parkinson disease, AD = Alzheimer disease, CFS = chronic fatigue syndrome, CVD = cardiovascular disease, OCD = obsessive compulsive disorder. Genes in **bold** were also found to be significant in PheWAS analyses.

Gene	Phecode Description	Tissue	Zscore	P (FDR)	P (Raw)	Ncase	Nctrl
UFL1	Pain and other symptoms associated with female genital organs	Brain Cerebellum	-4.44	0.0425	9.02E-06	14	7715
SLC38A3	Joint/ ligament sprain	Brain Caudate basal ganglia	5.83	0.00001	5.69E-09	20	7810
ERICH2	Disc disorders	Brain Amygdala	-5.05	0.0015	4.38E-07	738	17112
Novel Transcript	Spondylosis with myeolopathy	Brain Anterior cingulate cortex BA24	4.55	0.0315	5.34E-06	140	17813
Novel Transcript	Spondylosis with myeolopathy	Brain Cortex	4.62	0.0315	3.87E-06	140	17813

Table S10: PheWas results: mean painscore-adjusted. P (Raw) = unadjusted p value, P (FDR) = FDR-adjusted p value, Ncase = N cases for phecode, Nctrl = N controls for phecode.

Gene	Phecode Description	Tissue	Zscore	P (FDR)	P (Raw)	Ncase	Nctrl
PTK2	Excessive or frequent menstruation	Brain_Hypothalamus	4.81	0.018	1.49E-06	393	17300
SLC38A3	Joint/ ligament sprain	Brain_Caudate_basal_ganglia	5.86	1.10E-05	4.66E-09	20	7815
ERICH2	disc disorders	Brain_Amygdala	-5.18	0.00078	2.21E-07	738	17123
ERICH2	displacement of intervertebral disc	Brain_Amygdala	-4.22	0.043	2.41E-05	426	17311
Novel Transcript	spondylosis with myelopathy	Brain_Anterior_cingulate_cortex_BA24	4.54	0.033	5.53E-06	140	17824
Novel Transcript	spondylosis with myelopathy	Brain_Cortex	4.61	0.033	4.05E-06	140	17824

Table S11: PheWas results: matched sample for mean painscore-adjusted, without adjustment for mean painscore. P (Raw) = unadjusted p value, P (FDR) = FDR-adjusted p value, Ncase = N cases for phecode, Nctrl = N controls for phecode.

Supplementary References

1. Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM, et al. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. Nat Commun. 2018 May 8;9(1):1825.

2. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, et al. A gene-based association method for mapping traits using reference transcriptome data. Nat Genet. 2015 Sep;47(9):1091–8.

3. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017 Dec;8(1):1826.

4. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences. 2005 Oct 25;102(43):15545–50.

5. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015 Dec 23;1(6):417–25.

6. Martens M, Ammar A, Riutta A, Waagmeester A, Slenter DN, Hanspers K, et al. WikiPathways: connecting communities. Nucleic Acids Research. 2021 Jan 8;49(D1):D613–21.

7. Sollis E, Mosaku A, Abid A, Buniello A, Cerezo M, Gil L, et al. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. Nucleic Acids Res. 2023 Jan 6;51(D1):D977–85.

8. Toikumo S, Vickers-Smith R, Jinwala Z, Xu H, Saini D, Hartwell E, et al. The genetic architecture of pain intensity in a sample of 598,339 U.S. veterans [Internet]. medRxiv; 2023 [cited 2023 Jul 19]. p. 2023.03.09.23286958. Available from: https://www.medrxiv.org/content/10.1101/2023.03.09.23286958v1

9. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. 2018 Jan 4;46(Database issue):D1074–82.

10. Bastarache L. Using Phecodes for Research with the Electronic Health Record: From PheWAS to PheRS. Annu Rev Biomed Data Sci. 2021 Jul 20;4(1):1–19.

11. Wang L, Zhang X, Meng X, Koskeridis F, Georgiou A, Yu L, et al. Methodology in phenome-wide association studies: a systematic review. 2021;58:720–8.

12. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BWJH, et al. Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet. 2016 Mar;48(3):245–52.

13. Melzack R. From the gate to the neuromatrix. PAIN. 1999 Aug;82:S121.

14. De Ridder D, Adhia D, Vanneste S. The anatomy of pain and suffering in the brain and its clinical implications. Neuroscience & Biobehavioral Reviews. 2021 Nov 1;130:125–46.

15. Yang S, Chang MC. Chronic Pain: Structural and Functional Changes in Brain Structures and Associated Negative Affective States. Int J Mol Sci. 2019 Jun 26;20(13):3130.

16. Farmer MA, Baliki MN, Apkarian AV. A dynamic network perspective of chronic pain. Neuroscience Letters. 2012 Jun 29;520(2):197–203.

17. Barroso J, Branco P, Apkarian AV. Brain mechanisms of chronic pain: critical role of translational approach. Translational Research. 2021 Dec;238:76–89.

18. Su Q, Song Y, Zhao R, Liang M. A review on the ongoing quest for a pain signature in the human brain. Brain Science Advances. 2019 Dec 1;5(4):274–87.

19. IASP Announces Revised Definition of Pain [Internet]. International Association for the Study of Pain (IASP). [cited 2022 Sep 20]. Available from: https://www.iasp-pain.org/publications/iasp-news/iasp-announces-revised-definition-of-pain/

20. Dubin AE, Patapoutian A. Nociceptors: the sensors of the pain pathway. J Clin Invest. 2010 Nov 1;120(11):3760–72.

21. Thompson JM, Neugebauer V. Cortico-limbic pain mechanisms. Neurosci Lett. 2019 May 29;702:15–23.

22. Fakhoury M, Salman I, Najjar W, Merhej G, Lawand N. The Lateral Hypothalamus: An Uncharted Territory for Processing Peripheral Neurogenic Inflammation. Frontiers in Neuroscience [Internet]. 2020 [cited 2023 Mar 14];14. Available from: https://www.frontiersin.org/articles/10.3389/fnins.2020.00101

23. Fuchs PN, Peng YB, Boyette-Davis JA, Uhelski ML. The anterior cingulate cortex and pain processing. Front Integr Neurosci. 2014 May 5;8:35.

24. Ong WY, Stohler CS, Herr DR. Role of the Prefrontal Cortex in Pain Processing. Mol Neurobiol. 2019 Feb 1;56(2):1137–66.

25. Harris HN, Peng YB. Evidence and explanation for the involvement of the nucleus accumbens in pain processing. Neural Regen Res. 2019 Oct 18;15(4):597–605.

26. Lu C, Yang T, Zhao H, Zhang M, Meng F, Fu H, et al. Insular Cortex is Critical for the Perception, Modulation, and Chronification of Pain. Neurosci Bull. 2016 Feb 22;32(2):191–201.

27. Moulton EA, Schmahmann JD, Becerra L, Borsook D. The Cerebellum and Pain: Passive Integrator or Active Participator? Brain Res Rev. 2010 Oct 5;65(1):14–27.

28. Markovic T, Pedersen C, Massaly N, Vachez YM, Ruyle B, Murphy CA, et al. Pain induces adaptations in ventral tegmental area dopamine neurons to drive anhedonia-like behavior. Nat Neurosci. 2021 Nov;24(11):1601–13.

29. Elman I, Borsook D. Common Brain Mechanisms of Chronic Pain and Addiction. Neuron. 2016 Jan 6;89(1):11–36.

30. Wang M, Tutt JO, Dorricott NO, Parker KL, Russo AF, Sowers LP. Involvement of the cerebellum in migraine. Front Syst Neurosci. 2022;16:984406.

31. Kong J, Huang Y, Liu J, Yu S, Ming C, Chen H, et al. Altered functional connectivity between hypothalamus and limbic system in fibromyalgia. Molecular Brain. 2021 Jan 20;14(1):17.

32. Farrell SF, Campos AI, Kho PF, de Zoete RMJ, Sterling M, Rentería ME, et al. Genetic basis to structural grey matter associations with chronic pain. Brain. 2021 Dec 31;144(12):3611–22.

33. Hashmi JA, Baliki MN, Huang L, Baria AT, Torbey S, Hermann KM, et al. Shape shifting pain: chronification of back pain shifts brain representation from nociceptive to emotional circuits. Brain. 2013 Sep;136(9):2751–68.

34. Li T, Mamillapalli R, Ding S, Chang H, Liu ZW, Gao XB, et al. Endometriosis alters brain electrophysiology, gene expression and increases pain sensitization, anxiety, and depression in female mice. Biol Reprod. 2018 Aug;99(2):349–59.

35. Zangrandi A, Allen Demers F, Schneider C. Complex Regional Pain Syndrome. A Comprehensive Review on Neuroplastic Changes Supporting the Use of Non-invasive Neurostimulation in Clinical Settings. Front Pain Res (Lausanne). 2021 Sep 21;2:732343.

36. Baliki MN, Mansour AR, Baria AT, Apkarian AV. Functional Reorganization of the Default Mode Network across Chronic Pain Conditions. Zang YF, editor. PLoS ONE. 2014 Sep 2;9(9):e106133.