Research Paper

PAIN®

OPEN

Rare variant analyses in large-scale cohorts identified SLC13A1 associated with chronic pain

Xiang Ao[a], Marc Parisien[a], Maha Zidan[a], Audrey V. Grant[a], Amy E. Martinsen[b,c,d], Bendik S. Winsvold[b,c,e], Luda Diatchenko[a,*]

Abstract

Chronic pain is a prevalent disease with increasing clinical challenges. Genome-wide association studies in chronic pain patients have identified hundreds of common pathogenic variants, yet they only explained a portion of individual variance of chronic pain. With the advances in next-generation sequencing technologies, it is now feasible to conduct rarer variants studies in large-scale databases. Here, we performed gene-based rare variant analyses in 200,000 human subjects in the UK biobank whole-exome sequencing database for investigating 9 different chronic pain states and validated our findings in 3 other large-scale databases. Our analyses identified the *SLC13A1* gene coding for sodium/sulfate symporter associated with chronic back pain and multisite pain at the genome-wide level and with chronic headache, knee, and neck and shoulder pain at the nominal level. Seven loss-of-function rare variants were identified within the gene locus potentially contributing to the development of chronic pain, with 2 of them individually associated with back pain and multisite pain. These 2 rare variants were then tested for replication in 3 other biobanks, and the strongest evidence was found for rs28364172 as an individual contributor. Transcriptional analyses of *Slc13a1* in rodents showed substantial regulation of its expression in the dorsal root ganglia and the sciatic nerve in neuropathic pain assays. Our results stress the importance of the *SLC13A1* gene in sulfate homeostasis in the nervous system and its critical role in preventing pain states, thus suggesting new therapeutic approaches for treating chronic pain in a personalized manner, especially in people with mutations in the *SLC13A1* gene.

Keywords: Rare variants association test, Whole-exome sequencing (WES), UK biobank, HUNT, CLSA, Chronic pain, SLC13A1

1. Introduction

Whole-exome sequencing (WES)[7] technologies have proven to be efficient and economic tools in detecting disease-causing and drug targets.[17,38] It captures and sequences the coding regions of human genome, where functional proteins are produced. Whole-exome sequencing presents unique advantages in contrast to whole-genome sequencing technologies, including reduced costs, greater flexibility, and easier data storage and analysis.[26] Compared with genome-wide association study (GWAS) approaches that require genetic imputation inevitably yielding some missing and/or unreliable rare variants from imputation, WES produces sequenced nucleotides directly and credibly.

Chronic pain is a common medical condition that affects around 1 in 5 individuals worldwide.[48] It incurs appreciable health care and socioeconomic costs and leads to suffering by those affected. A comprehensive understanding of the molecular mechanisms underlying the development of chronic pain would benefit pain management and drug development. In the past decades, genetic studies in pain have made significant advances toward identifying crucial molecular factors contributing to the onset and maintenance of chronic pain.[35,37,42]

Most genetic association studies of chronic pain have focused on common genetic variations,[2,6,8,13,15,34] whereas studies based on WES point to a contribution from rare variants.[14,25] Although rare variants may make a smaller contribution to the overall risk for common disorders in comparison with common variants, they can be of a very substantial effect size. The identification of associated, rare, nonsynonymous variants offers advantages for translating genetic findings into biological understanding by, for example, implicating susceptibility genes more precisely than common variant analysis, or by providing causal mutations that may be more readily modeled in cellular and animal systems. Analysis of rare variants also helps to explain

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

[a] Faculty of Dental Medicine and Oral Health Sciences, Department of Anesthesia, Faculty of Medicine and Health Sciences, Alan Edwards Centre for Research on Pain, McGill University, Montreal, QC, Canada, [b] Department of Research, Innovation and Education, Oslo University Hospital, Oslo, Norway, [c] K. G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, Norway, [d] Faculty of Medicine, Institute of Clinical Medicine, University of Oslo, Oslo, Norway, [e] Department of Neurology and Department of Research, Innovation and Education, Oslo University Hospital, Oslo, Norway

**Corresponding author. Address: Faculty of Dental Medicine and Oral Health Sciences, Department of Anesthesia, Faculty of Medicine and Health Sciences, Alan Edwards Centre for Research on Pain, McGill University, 3775 Rue University, Room 207, Montreal, QC H3A 2B4, Canada. Tel.: 514 398-2833; 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 164 (2023) 1841–1851

http://dx.doi.org/10.1097/j.pain.0000000000002882

missing heritability that is not accountable by common variant studies.[16,32,45] Research of human pain genetics has pinpointed many common variants contributing to pain[2,6,8,13,15,34]; however, only 25% to 50% of the heritability for chronic pain conditions has been explained, for example, 25% for irritable bowel syndrome, 35% for back and neck pain, and 50% for chronic widespread pain and migraine.[36] Because of the low frequency of rare variants (minor allele frequency, MAF < 0.01) and the limited sample size of available cohorts, previous studies were unable to detect reliable rare variant associations. In the recent decade, large-scale databases have become increasingly available,[5,20,22,43] empowering the studies of rare variants of complex traits. For example, the UK biobank (UKBB) recruited and sequenced 500,000 participants with an abundant record of various phenotypes.[5,43] The Trøndelag Health Study (HUNT), Norway's largest collection of health data from a population, started to collect clinical measurements and samples since 1984 and has recruited 230,000 participants.[20,22] The Canadian Longitudinal Study on Aging (CLSA) is a long-term study that collected more than 50,000 participants.[39]

Here, we investigated the association of rare variants with several chronic pain conditions in 4 distinct large-scale data sets: UKBB 200k WES, UKBB 300k GWAS database, HUNT, and CLSA. We used a gene-based rare variant analysis method to identify candidate genes and rare variants that lead to chronic pain.

2. Methods

2.1. Cohorts

We conducted our studies in the large-scale UKBB cohorts, under Application Number 20802. The genetic data consisted of 2 subcohorts: one was a whole-exome sequencing data set cohort containing more than 200,000 participants (UKBB 200k),[43] whereas the other one was a genome-wide genotyping data set cohort, with more than 500,000 participants (UKBB 500k).[5] We assigned individuals not in the UKBB 200k cohort to the henceforth named UKBB 300k cohort. We set the UKBB 200k cohort as our discovery cohort and replicated our results in the UKBB 300k cohort. Although the UKBB 200k cohort was smaller than the UKBB 300k, access to high-accuracy rare variant information was only available in the former, enabling true rare variants analyses. The demographic description of the 2 cohorts was listed in **Table S1** (available as supplemental digital content at http://links.lww.com/PAIN/B790). This study was done within individuals of white British ancestry, of both sexes, regardless of their age. Individuals with failed genetic quality controls (QC), sex mismatches or aneuploidy, and those who opted out of the study were not considered.

The Nord-Trøndelag Health Study is an ongoing population-based cohort study from the county of Nord-Trøndelag in Norway.[20,22] All inhabitants aged 20 years or older were invited to participate in the HUNT1 survey (1984-1986), the HUNT2 survey (1995-1997), and the HUNT3 survey (2006-2008). Participation rates in HUNT1, HUNT2, and HUNT3 were 89.4% (n = 77,212), 69.5% (n = 65,237), and 54.1% (n = 50,807), respectively.[22] Taken together, the study included more than 120,000 different individuals from Nord-Trøndelag County. For this study, we included participants from HUNT2 and HUNT3. All participants have provided questionnaires, interviews, and measurement data, which can be found at the HUNT databank [https://hunt-db.medisin.ntnu.no/hunt-db]. In addition, approximately 80,000 participants have provided biological samples for storage at the HUNT biobank [https://www.ntnu.edu/hunt/hunt-biobank]. Written informed consent was obtained from all participants, and The Regional Committee for Medical and Health Research Ethics approved the study (ref. 2015/573).

The CLSA is a longitudinal cohort study of 51,338 participants, aged 45 to 85 years at enrollment, from 10 Canadian provinces.[39,40] Recruitment and baseline data collection took place from 2010 to 2015, and follow-up visits were planned to take place every 3 years with data collection for the first follow-up visit completed in 2018 and available for analysis. Individuals from the CLSA cohort who underwent full onsite evaluation were considered for this analysis. From this set, individuals who were genotyped and who were identified as the European descent majority subset were included.[39] In total, the total study population of European-descent individuals with genotyping and imputed data with back pain status at least at baseline or at visit 1 was 25,239 subjects. This included 8271 cases (54% females), and 16,968 controls (48% females).

2.2. Variant annotation

Variant annotation was performed with Ensembl Variant Effect Predictor (VEP, release 99).[33] Coding regions were defined by a bed file (UKBB resource 3803). A variant was deemed rare if its minor allele frequency (MAF) was below 1% (MAF < 0.01) in all gnomAD populations and passed the MAF filter in white British population of cohorts.[21] A variant was deemed loss of function (LoF) if it satisfied the following conditions: (1) not a benign in PolyPhen[1] or Sorting Intolerant From Tolerant (SIFT)[24] (PolyPhen benign is < 0.15 or SIFT benign is > 0.05) and (2) one of stop_lost, start_lost, splice_donor_variant, frameshift_variant, splice_acceptor_variant, or stop_gained type of variant. The study analyzed rare LoF variants.

2.3. Sample selection

In the UKBB cohorts, we first selected participants who had answered the question "In the last month have you experienced any of the following that interfered with your usual activites? (You can select more than one answer)" (Data field 6159). The options contained the following: headache, facial pain, neck or shoulder

Table 1

Gene-based loss-of-function rare variants association tests on SLC13A1 for chronic pain conditions in the UK biobank 200k whole-exome sequencing cohort.

Pain	#LoF	*P*			Adj-*P*
		Burden	SKAT	SKAT-O	
Back	7	2.0E-07	2.3E-05	2.9E-07	—
Facial	4	5.6E-02	1.6E-01	8.2E-02	5.7E-01
General	4	8.3E-01	6.7E-01	7.9E-01	1
Headache	5	9.4E-04	9.2E-03	1.2E-03	8.5E-03
Hip	6	9.1E-01	2.9E-01	1.3E-01	9.3E-01
Knee	7	1.7E-05	4.8E-05	1.7E-05	1.2E-04
Neck or shoulder	7	1.5E-05	6.1E-05	1.7E-05	1.4E-04
Stomach or abdominal	4	4.2E-02	8.8E-02	6.2E-02	4.3E-01
Multisite	7	1.0E-07	4.6E-06	1.4E-07	-

P values were calculated with 3 different models: burden test (Burden), SKAT test (SKAT), and SKAT-O test (SKAT-O). SKAT-O *P* values were adjusted (in column adj-*P*) with Bonferroni correction among all chronic pain conditions except chronic back pain and multisite pain (N = 7). Significant results from primary analysis passed a genome-wide statistical threshold of 2.6E-06.
LoF, loss of function.

pain, back pain, stomach or abdominal pain, hip pain, knee pain, pain all over the body, none of the above, and prefer not to answer. Among the participants who answered "yes," we further collected information on those who experienced the pain for more than 3 months. This information was accessed for chronic back pain (Data field: 3571), chronic facial pain (Data field: 4067), chronic general pain/pain all over the body (Data field: 2956), chronic headaches (Data field: 3799), chronic hip pain (Data field: 3414), chronic knee pain (Data field: 3773), chronic neck or shoulder pain (Data field: 3404), and chronic stomach or abdominal pain (Data field: 3741). We selected the participants who reported any chronic pain (Data field 3571, 4067, 2956, 3799, 3414, 3773, 3404, 3741) as our study groups. Participants who reported pain at more than one body site except the general pain were also bound together in a new group, chronic multisite pain, to explore the impact of the number of pain sites. Participants who did not report any pain (Data field 6159, answered "none of the above") in this study were assigned to a control group. See Table S1, available as supplemental digital content at http://links.lww.com/PAIN/B790 for the demographics and sample size of each pain group. We also retrieved intervertebral disc disorders (IDD) samples by their International Classification of Diseases (ICD-10) codes, ie, M51, from UKBB (Data field: 42,040, 3522 cases and 163,230 controls) to control for structural alteration of muscular–skeletal tissues.

In HUNT2 and HUNT3, one question stated, "During the last year, have you had pain and/or stiffness in your muscles and limbs that has lasted for at least 3 consecutive months?" If a respondent answered in the affirmative, the following question was asked: "Where did you have pain and/or stiffness?" The lower back was one site listed among several possibilities. Respondents answering "yes" to the first question and indicating the lower back as a site of pain were regarded as experiencing chronic low back pain (n = 19,760 with valid genotypes). Respondents answering "no" to the first question comprised the controls (n = 28,674 with valid genotypes).[19]

In the CLSA cohort, we defined back pain cases as individuals who had ever experienced chronic back pain on the basis of a series of questions in the general questionnaire, Data Collection Site Questionnaires (Comprehensive). In the analysis, we included all individuals who answered the question, "Have you ever had pain in your back on most days for at least one month?" with either a yes or a no, at either baseline or at the first follow-up visit. We excluded individuals who answered, "Don't know," who did not answer or refused to answer. Chronic back pain cases were defined as those individuals who answered the follow-up question, "For how long?" with 3 or more months, again, at either baseline or the follow-up visit. All other individuals were classified as controls.

2.4. Statistical analysis

After collecting the sequencing data from UKBB 200k cohort, we used SAIGE-GENE[50] for the gene-based rare-variant association test and single-variant association test. SAIGE-GENE contains 3 steps (step 0, step 1, and step 2) in analyzing rare variants and reports results of burden test, Sequence Kernel Association Test (SKAT), and Optimal SKAT (SKAT-O) test. We used variants in both coding regions and high-confidence Genome in a Bottle (GIAB) regions[51] to construct a genetic relationship matrix and fit a null model in step 0 and step 1. Sex, age, squared age, first 40 principal components of the population's genetics, and types of the Illumina NovaSeq 6000 sequencing flow cells were used as covariates in the SAIGE-GENE step 2 association test. Genome-wide significance for primary analyses was defined as $P < 2.6E\text{-}6$ for results from gene-based analyses and $P < 5E\text{-}8$ for single-variant test results, which correspond to Bonferroni corrections for 19,000 genes and for 1 million haploblocks, respectively. Significance for secondary analyses established from adjusted P value of <0.05. We adjusted P values with Bonferroni when tests were independent or with false discovery rate (FDR) when independence cannot be guaranteed (eg, same gene in multiple tissues, one single nucleotide polymorphism (SNP) in multiple chronic pain conditions). We also performed a secondary analysis in the UKBB 200k cohort where we tested an association between *SLC13A1* with each reported pain site at the gene level and single-variant level. Linkage disequilibrium (LD) between variants was measured by Lewontin's D'.[27]

As a complementary analysis, we further investigated the common genetic characteristics between self-reported chronic pain conditions and clinically diagnosed IDD and dorsalgia. We used common variants from whole-genome analysis from the study of Bjornsdottir et al.,[4] which presented association with IDD and dorsalgia, and tested their association with multisite pain. We retrieved the common variants in both UKBB 200k cohort and UKBB 500k cohort; for those variants that were not presented in the cohort, we used their proxy variants instead. The linkage disequilibrium between these variants and their proxy was measured in R^2. We searched the proxy variants in a window of 500,000 base pairs of the target common variants in the Great British Population of the 1000 Genomes Project using the module LDproxy from LDlink.[30] We selected the top proxy variant with $R^2 > 0.8$ for each common variant. The common variants or their proxy variants were tested in association with multisite pain in UKBB 200k and UKBB 500k by SAIGE-GENE.

2.5. Study replications

To verify the reproducibility of our findings, we replicated our analyses in the UKBB 300k cohort, CLSA cohorts, and HUNT cohorts.

In HUNT, DNA from 71,860 samples was genotyped using 1 of 3 different Illumina HumanCoreExome arrays (HumanCoreExome12 v1.0, HumanCoreExome12 v1.1 and UM HUNT Biobank v1.0). For details on genotyping, quality control and imputation, refer Hautakangas et al.[18] The association analysis was conducted using SAIGE,[49] a generalized mixed-effects model approach, to account for cryptic population structure and relatedness when modelling the association between genotype probabilities (dosages) and chronic low back pain. The model was adjusted for sex, genotyping batch, and the first 5 principal components (PCs).

In CLSA, 26,622 participants were genotyped using Affymetrix Axiom array, tested 794,409 genetic variants, and imputed approximately 308 million genetic variants through TOPmed.[44] Individual-level and SNP-level QC was performed by a CLSA team to result in this data set, as described previously [https://www.clsa-elcv.ca/doc/2748]. The genome-wide case–control scan on any back pain was conducted using SAIGE.[49] The model was adjusted for age, sex, and the first 50 principal components.

We applied inverse variance based method from METAL[46] to do meta-analysis for variants of interest across replication cohorts UKBB300k, HUNT, and CLSA.

2.6. Differential gene expression analysis and pathway analysis

Using data from the Gene Expression Omnibus (GEO) database,[3] we performed analyses of differential gene expression in data sets designed from animal pain assays. A total of 181 contrasts were

surveyed in 4 different tissues that featured at least 10 pain contrasts. The tissues were as follows: dorsal root ganglia, sciatic nerve, spinal cord, and whole blood. Detection of differential gene expression was performed using either DESeq2[29] for high-throughput sequencing data or GEOquery[9] plus limma[41] for cDNA microarray data.

3. Results

3.1. Study overview

We designed this study to explore the contribution of rare genetic variants to chronic pain conditions, by investigating the WES genetic data from the UKBB with gene-based rare variants association models. We focused our analysis on LoF rare variants because these are most likely to produce functional consequences of substantial effect size. **Figure 1** shows the schematic flow chart of this study, which contained 7 main steps. In step I, we annotated LoF rare variants, and we screened case and control samples for analyses of 8 distinctive chronic pain reports in step II. We then applied gene-based analytical models in step III for discovering the association between LoF variants and chronic pain. After the analysis of the primary findings at the gene level in step IV, we tested for association of single variant in step V. Replication analyses were performed in step VI. Finally, characterization of the regulatory role of Slc13a1 on pain development was done through the differential gene expression analyses in mouse pain models, as well as the pathway analyses, and were completed in step VII.

3.2. Gene-based rare-variant analysis in chronic pain conditions

We first tested whether there were genes associated with chronic pain conditions at a genome-wide significance level in the UKBB 200k WES cohort. Pain phenotypes included in the analyses were chronic back pain, chronic facial pain, chronic general pain, chronic headache, chronic hip pain, chronic knee pain, chronic neck or shoulder pain, chronic stomach or abdominal pain, and chronic multisite pain. Table S1, available as supplemental digital content at http://links.lww.com/PAIN/B790 listed the demographics and number of cases for each chronic pain phenotype.

We used SAIGE-GENE in the UKBB 200k WES database to group LoF rare variants in each gene and investigated their integrative association with each of the chronic pain conditions (**Table 1**). The gene-based analysis reported one gene, Solute Carrier Family 13 (sodium/sulfate symporters), member 1 (*SLC13A1*), to be associated at the genome-wide significant level ($P < 2.6E-6$) with the report of chronic back pain and multisite pain (Fig. S1, **Fig. 2A**, available as supplemental digital content at http://links.lww.com/PAIN/B790). Both the Burden and SKAT-O tests showed significant results suggesting most rare LoF variants detected from *SLC13A1* functioned in the same direction in regulating chronic back pain and multisite pain. Although analysis for the other pain phenotypes did not produce genome-wide significant results, the same gene *SLC13A1* was associated with chronic headache, knee, and neck and shoulder pain at an adjusted significance level (adjusted $P < 0.05$). Gene-based tests with rare coding variants were also conducted for the same chronic pain conditions; however, these tests did not reveal any genes that had a genome-wide significant association with the report of chronic pain body sites (Table S2, available as supplemental digital content at http://links.lww.com/PAIN/B790).

3.3. Single-variant analysis in chronic pain conditions

We further investigated all of the LoF rare variant(s) within *SLC13A1* available from exome-sequencing data in the UKBB 200K cohort to explore their individual contribution to chronic pain. In total, we found 7 LoF rare variants from the analysis of chronic back pain and multisite pain samples (Fig. S2, **Fig. 2B**, available as supplemental digital content at http://links.lww.com/PAIN/B790; **Table 2**). Five of them were stop_gained variants (5 of 7) and 2 were frameshift variants. The variants were found to be independently distributed in the human population, not displaying any detectable pairwise LD (all pairwised Lewontin's D' = 0, **Fig. 2C**). All of them have a

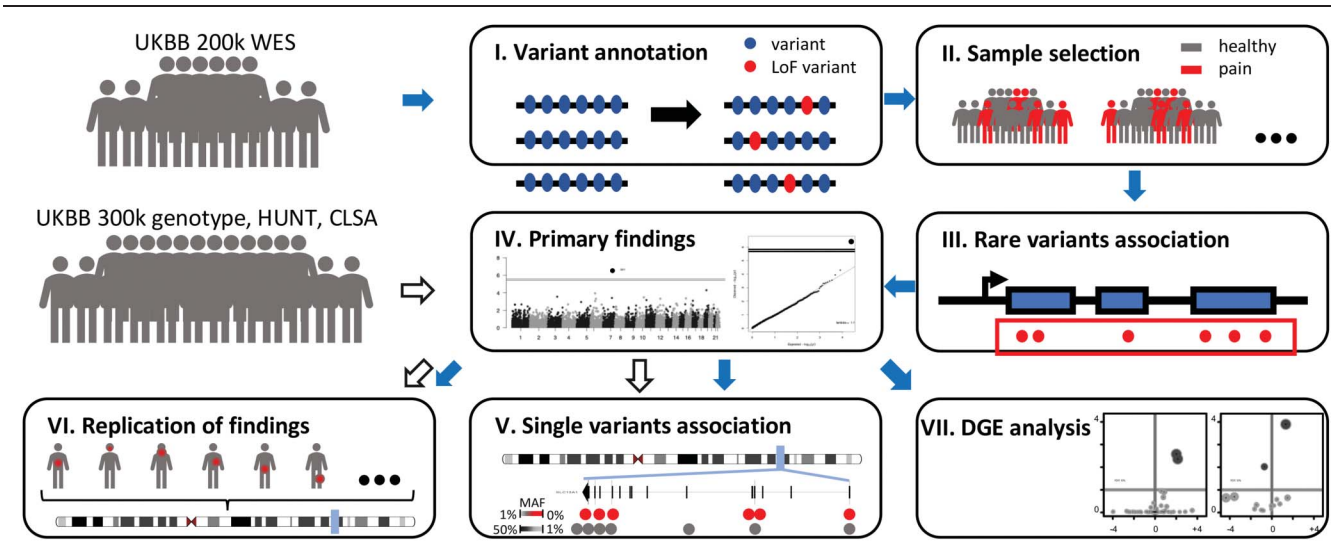

Figure 1. Schematic flow chart of the study design. Analyses were done on cohorts UKBB 200k, UKBB 300k, HUNT, and CLSA. Step I. Variant annotation. Step II. Sample selection. Step III. Rare variants association test. Step IV. Primary results investigation. Step V. Single-variant association tests. Step VI. Replication analyses. Step VII. Differential gene expression (DGE) analysis in animal pain models. Blue arrows indicate primary analyses and white arrows indicate replication analyses. CLSA, Canadian longitudinal study on aging; HUNT, The Trøndelag Health Study; MAF, minor allele frequency; UKBB, UK biobank; WES, whole-exome sequencing.

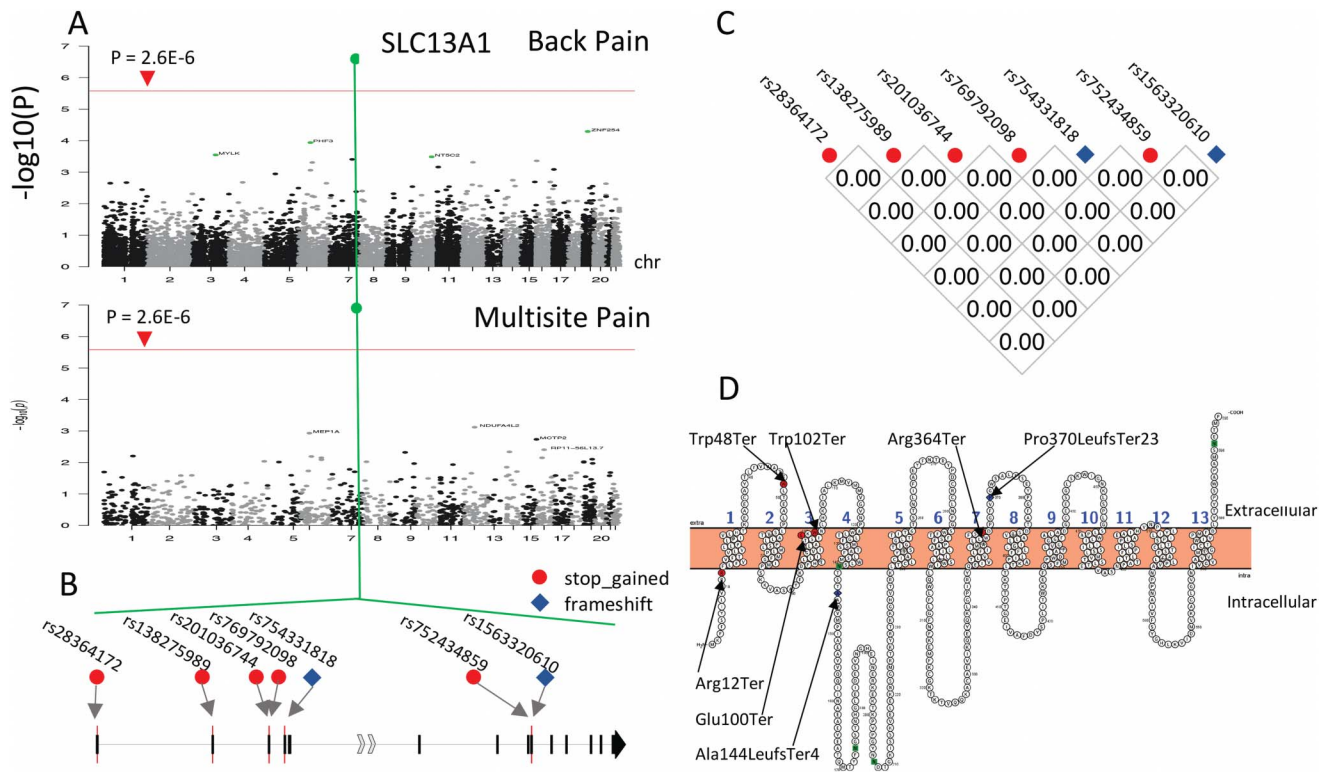

Figure 2. Primary finding from chronic pain samples in the UKBB 200k WES data set. (A) Manhattan plot of the gene-based rare LoF variants test in chronic back pain (up) and multisite pain (bottom) in the UKBB 200k WES cohort. The red line represents the genome-wide significance threshold. Gene SLC13A1, ie, Solute Carrier Family 13 (sodium/sulfate symporters) member 1, was identified to be associated with the development of chronic back pain and multisite pain. (B) Genomic coordinates of the rare LoF variants from the primary findings, red circles represent stop gained variants, whereas blue diamonds represent frameshift variants. (C) A linkage disequilibrium plot of the primary LoF variants, the order of variants is the same as in panel (B). (D) Mutations corresponding to the identified variants in protein SLC13A1. LoF, loss-of-function; UKBB, UK biobank; WES, whole-exome sequencing.

deleterious Combined Annotation Dependent Depletion (CADD) score higher than 15, and 6 of them have a CADD score of at least 30 (**Table 2**). Of importance, such large CADD scores (CADD > 15) indicate a high risk of pathogenicity. The 7 variants were all located closer to the N-terminus, thus likely completely eliminating the functional protein (**Fig. 2D**). Single-variant association analysis showed that 2 variants were associated with back pain and 3 with multisite pain with the adjusted $P < 0.05$: rs28364172, rs752434859, and rs754331818 (**Table 3**). The variants rs28364172 and rs754331818 were found associated with both chronic back pain (odds ratios [OR] = 1.5, adjusted $P = 9.5E-05$ and OR = 5.8, adjusted $P = 4.1E-05$) and multisite pain (OR = 1.1, adjusted $P = 3.0E-05$ and OR = 1.8, adjusted $P = 1.8E-05$). Of importance, in all cases, these LoF variants were risk factors for the development of chronic pain.

We then examined whether 7 *SLC13A1* LoF rare variants could be found in the UKBB 300k cohort where we had access to genotyping data (Table S3, available as supplemental digital content at http://links.lww.com/PAIN/B790). We found that 2 variants rs28364172 and rs138275989, out of the 7 rare variants were also genotyped in the UKBB 300k cohort. The variant rs28364172 was found associated with chronic back pain (adjusted $P = 2.0E-02$). These 2 variants were further reevaluated in the entire UKBB 500k GWAS cohort (Table S4, available as supplemental digital content at http://links.lww.com/PAIN/B790). Indeed, the variant rs28364172 was significantly

Table 2

Loss-of-function rare variants found in chronic back pain and multisite pain samples from UK biobank 200k whole-exome sequencing cohort.

rsid	Chr:Pos	Mutation	Ref	Alt	Consequence	CADD
rs28364172	7:123199913	Arg12Ter	G	A	stop_gained	34
rs138275989	7:123181057	Trp48Ter	C	T	stop_gained	36
rs201036744	7:123171835	Glu100Ter	C	A	stop_gained	38
rs769792098	7:123171827	Trp102Ter	C	T	stop_gained	38
rs754331818	7:123169270	Ala144LeufsTer4	GC	G	frameshift	16
rs752434859	7:123128888	Arg364Ter	G	A	stop_gained	34
rs1563320610	7:123128868	Pro370LeufsTer23	AG	A	frameshift	30

Columns show the rsid, chromosome number and position (Chr:Pos), mutation (Mutation), reference allele (Ref), alternative allele (Alt), functional consequence of alternative allele (Consequence), CADD score (CADD, rounded to nearest integer).

Table 3

Loss-of-function rare variants found in chronic pain samples from UK biobank 200k whole-exome sequencing cohort.

rsid	Pain	AF	N	BETA	SE	*P*	FDR	OR (95% CI)
rs28364172	Back	2.8E-03	96,058	0.41	0.10	2.7E-05	9.5E-05	1.5 (1.2-1.8)
rs138275989	Back	1.0E-03	96,058	0.18	0.16	2.6E-01	3.3E-01	1.2 (0.9-1.6)
rs201036744	Back	2.1E-05	96,058	0.98	1.10	3.8E-01	3.8E-01	2.7 (0.3-23.0)
rs769792098	Back	1.6E-05	96,058	−1.41	1.31	2.8E-01	3.3E-01	0.2 (0.02-3.2)
rs754331818	Back	1.8E-04	96,058	1.75	0.39	5.8E-06	4.1E-05	5.8 (2.7-12.4)
rs752434859	Back	2.1E-05	96,058	2.06	1.09	5.9E-02	1.0E-01	7.9 (0.9-66.4)
rs1563320610	Back	3.1E-05	96,058	1.85	0.85	2.9E-02	6.8E-02	6.4 (1.2-33.6)
rs28364172	Multisite	2.9E-03	137,168	0.13	0.03	8.6E-06	3.0E-05	1.1 (1.1-1.2)
rs138275989	Multisite	1.0E-03	137,168	0.04	0.05	0.4E-00	4.7E-01	1.0 (0.9-1.1)
rs201036744	Multisite	1.8E-05	137,168	−0.01	0.37	9.8E-01	9.8E-01	1.0 (0.5-2.0)
rs769792098	Multisite	1.1E-05	137,168	−0.74	0.48	1.2E-01	1.7E-01	0.5 (0.2-1.2)
rs754331818	Multisite	1.6E-04	137,168	0.58	0.12	2.5E-06	1.8E-05	1.8 (1.4-2.3)
rs752434859	Multisite	1.8E-05	137,168	1.02	0.37	5.6E-03	1.3E-02	2.8 (1.3-5.7)
rs1563320610	Multisite	2.2E-05	137,168	0.63	0.33	6.0E-02	1.1E-01	1.9 (0.9-3.6)

Columns show the rsid, pain site (Pain), allele frequency (AF) of the alternative allele, sample size (N), beta value from single variant analysis, and the corresponding standard error (SE), *P* values (*P*), FDR adjusted *P* values (FDR), and odds ratios (OR) with their 95% confidence intervals (95% CI).
FDR, false discovery rate.

associated with chronic back pain and multisite pain (adjusted $P = 9.9E\text{-}06$ and adjusted $P = 4.4E\text{-}03$, respectively). Furthermore, in all but 2 associations, the variants presented odd ratios of ≥1, hence, indicating that these LoF variants confer risk for the development of different types of chronic pain conditions.

Next, we sought to further replicate our findings in independent cohorts. We found that 2 rare variants that were genotyped in the UKBB cohorts, rs28364172 and rs138275989, were also genotyped in the HUNT and CLSA cohort (Table S5, available as supplemental digital content at http://links.lww.com/PAIN/B790). Only the chronic back pain but not multisite pain phenotype was available in these cohorts for replication. The LoF rare variant rs28364172 reported a significant *P* value in the CLSA cohort (OR = 1.6, adjusted $P = 5.6E\text{-}03$). In the HUNT cohort, association between SNP rs138275989 and chronic back pain was approaching the nominal significance (OR = 1.5, $P = 0.07$). Meta-analysis of rs28364172 and rs138275989 in chronic back pain in 3 replication cohorts (UKBB300k, HUNT, and CLSA) showed that rs28364172 was significantly associated with chronic back (Fig. S3, available as supplemental digital content at http://links.lww.com/PAIN/B790). We were unable to do meta-analysis of the 2 variants in multisite pain because we do not have the same multisite pain phenotype in both the HUNT and the CLSA cohorts. Together, our results indicated that both variants, rs28364172 and rs138275989, show some evidence for association with chronic pain, but the strongest evidence was found for rs28364172 in relation to the development of chronic back pain.

We then asked whether *SLC13A1* contributed to chronic pain through the alteration of muscular–skeletal tissues. We repeated our analysis using intervertebral disc disorders (IDD) status as a binary covariate in our analysis of chronic back pain and by excluding IDD from chronic back pain samples. The gene-based results demonstrated that *P* values of both analyses remained genome-wide significant (Table S6, available as supplemental digital content at http://links.lww.com/PAIN/B790).However, we did not obtain a genome-wide significant signal from the 3 gene-based tests for IDD status itself. The individual variant association test was also repeated with IDD as a binary covariate (Table S7, available as supplemental digital content at http://links.lww.com/PAIN/B790). Together, our results suggest that *SLC13A1* contributes to the experience of pain.

We further investigated the common variants additional to *SLC13A1* that were reported to be associated with IDD and dorsalgia recently[4] and tested their association with multisite pain in UKBB 200k and UKBB 500k (Table S8, available as supplemental digital content at http://links.lww.com/PAIN/B790). In total, 13 of 44 common variants or their proxy variants were sequenced in the UKBB 200k WES cohort, whereas 31 of 44 common variants or their proxies were tested in the UKBB 500k cohort. For common variants that detected significant association with IDD ,[4] we discovered that 4 of 9 variants and 7 of 21 variants were associated with multisite pain in the UKBB 200k and the UKBB 500k, respectively. As to the common variants reported contribution to dorsalgia in Ref. 4, we found absolute majority of them associated with multisite pain—5 of 5 variants and 10 of 12 variants in the UKBB 200k and the UKBB 500k, respectively.

3.4. Differential gene expression in pain assays

Next, we asked whether the *Slc13a1* gene is regulated or differentially expressed in animal pain assays. To do so, we used the GEO database, which contains full transcriptomics data on hundreds of animal pain assays conducted mostly in mice and rats. Considered tissues were those from the central nervous system (spinal cord), peripheral nervous system (dorsal root ganglia and sciatic nerve), and whole blood. After correcting all tests across all available data sets using FDR, we found evidence for significant differential transcription of the *Slc13a1* gene in rat pain assays: in dorsal root ganglia 14 days after transection of the sciatic nerve (Fig. S4A-I; Fig. S4E, available as supplemental digital content at http://links.lww.com/PAIN/B790) and spinal nerve ligation (Fig. S4A-II; Fig. S4E, available as supplemental digital content at http://links.lww.com/PAIN/B790), and in the sciatic nerve 7 and 1 days after transection of the sciatic nerve

(Fig.S4B-III, Fig.S4E; and Fig.S4B-IV, Fig. S4E, available as supplemental digital content at http://links.lww.com/PAIN/B790). We did not find evidence for differential transcription in whole blood (Fig. S4C, available as supplemental digital content at http://links.lww.com/PAIN/B790) or in the spinal cord (Fig. S4D, available as supplemental digital content at http://links.lww.com/PAIN/B790). Consequently, we performed prospective differential gene expression analysis in the available rat sciatic nerve and dorsal root ganglia data set to determine the trajectory of *Slc13a1* expression across time. We used a gene expression data set from rat models of the study of Li et al.,[28] which provided gene expression profiles in L4-6 dorsal root ganglia tissues and proximal sciatic nerve tissues at day 0, day 1, day 4, day 7, and day 14 following sciatic nerve resection. In the sciatic nerve, we found a significant change in the expression of *Slc13a1* at day 1 (down-regulated) and day 7 (up-regulated), compared with day 0 (**Fig. 3**). In the dorsal root ganglia, *Slc13a1* was significantly up-regulated at day 4 and day 14 when compared with day 0 (**Fig. 3**). Overall, directions of fold changes suggest gradual elevation expression of *Slc13a1* in pain assays compared with sham after initial downregulation of the gene on day 1 (**Fig. 3**; Fig. S4E, available as supplemental digital content at http://links.lww.com/PAIN/B790).

4. Discussion

Chronic pain conditions are complex disorders that require a systematic search for the underlying genetic risk factors to understand their molecular pathophysiology. Although a fair number of chronic pain GWASs have been conducted and have identified many common SNPs associated with a genome-wide significance, their contribution to the development of pain is limited and the identified variants explain a limited portion of the heritability.[47] Meanwhile, large-scale databases enable the study of rare variants in many scientific fields. A part of the missing heritability of diseases such as chronic pain could hence be addressed from rare variants association studies, although such analyses have not been engaged yet in pain research.

In this study, we adopted the large-scale UKBB whole-exome sequencing database and examined the contribution of rare variants to chronic pain conditions. We concentrated on LoF variants because these have the strongest effect size and easily interpretable functional consequences. Our study revealed a novel association between *SLC13A1* and chronic pain. Although at the genome-wide level only, back pain and multisite pain were associated with a gene, at an adjusted significance level, we also found an association with chronic headache, knee, and neck and shoulder pain. Thus, it is likely that *SLC13A1* systemically contributes to pain states at large, and the genome-wide significant association of *SLC13A1* with back pain is mostly a reflection of the largest sample size of back pain subjects (Table S1, available as supplemental digital content at http://links.lww.com/PAIN/B790) across body sites. In fact, the rank of association evidence through *P* values matched the sample size rank across body sites considering back, knee, and neck and shoulder pain (Table S2, available as supplemental digital content at http://links.lww.com/PAIN/B790).

When we tested all LoF rare variants within *SLC13A1* in the single-variant analysis, we found 4 variants associated with the back pain or multisite pain. All of them substantially increase the risk of chronic pain, up to 6.4 fold for rs1563320610 (**Table 3**). Because all of them were LoF mutations, we can conclude that *SLC13A1* expression is protective against chronic pain. Importantly, we were able to replicate the findings in the other 3 large-scale databases, ie, the UKBB 300k genotyping database, the HUNT database, and the CLSA database.

GWASs on rare variants used to lack of power to detect accurate associations subject to the low frequency of rare variants and the limited sample size. A simulation study from Rivas and Moutsianas showed that approximately 80,000 samples were desired to achieve 80% power in the association detection between rare variants and dichotomous traits when setting the maximum effect size to 4.[23] Consistent with this estimation, we observed the strongest contribution at the individual variant level at OR of 5.8 (**Table 3**).

Finally, we studied the expression pattern of *Slc13a1* in animal pain assays in the wide range of publicly available transcriptomics-wide data sets to further validate the potential role of Slc13a1 in pain regulation. Our analyses identified a biphasic regulation of *Slc13a1* expression in the rat sciatic nerve and dorsal root ganglia in the posttransection of the sciatic nerve and spinal nerve ligation assays. We found that after the initial downregulation in the sciatic nerve, *Slc13a1* was overexpressed in both the sciatic nerve and dorsal root ganglia with advancement of the pain in the model. Because the pain is resolved over time in these models, our results suggest the protective regulatory role of *Slc13a1* in pain resolution, consistent with our genetic findings from genome-wide association studies.

SLC13A1, also known as *NAS1* and *NaSi-1*, encodes sodium/sulfate cotransporter that is responsible for the regulation of inorganic sulfate levels in serum. Sulfate is one of the most important macronutrients in the human body. It is a hydrophilic anion that requires active transport across the lipid bilayer of cell membranes. There are 2 classes of sulfate transporters: the sodium-dependent sulfate symporters that mediate the intake of sulfate with the concomitant uptake of sodium ions and the sodium-independent transporters that mediate sulfate uptake independent of sodium. *SLC13A1* is a sodium-dependent sulfate symporter that is primarily expressed in the kidney and intestine, where it mediates reabsorption from the proximal renal tubule of the kidney and mediates absorption of dietary sulfate in the intestine, respectively. Mice that lack *SLC13A1* exhibit hyposulfatemia and hypersulfaturia leading to a multiple clinically relevant phenotypes, including enhanced tumor growth,[12] reduced body weights,[10] impaired memory,[11] clonic seizure,[10] and

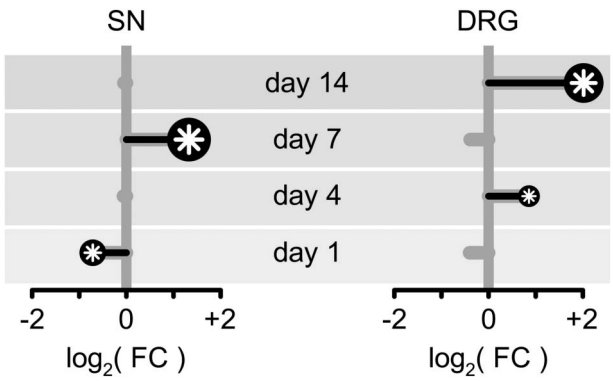

Figure 3. Differential expression of Slc13a1 in a rat model of nerve injury. Plots track fold change (FC) intensities and direction (up-regulation or down-regulation), in sciatic nerve (SN; left) and dorsal root ganglia (DRG; right) tissues, at multiple time points compared with baseline (day 0). Significant fold changes after Bonferroni correction (n = 4) are highlighted using a white star on top of a black circle. Circle size proportional to $-\log_{10}(P)$ of the *P* of the fold change. Vertical grey bars indicate the null fold change. Data from GEO accession GSE30165. GEO, gene expression omnibus.

the like. Biochemical measurement of neurotransmitter levels in *Slc13a1*-knockout (KO) mice displayed a reduced brain serotonin level, suggesting that circulating sulfate concentrations, which are regulated by *Slc13a1*, modulate serotonin neurotransmitter and receptor levels and hence impaired mice's cognitive abilities. Overall, KO mice experiments indicate that *SLC13A1* is essential for maintaining sulfate homeostasis and its disruption leads to detrimental effects on many functions including the nervous system. Because our analyses either with IDD as covariates or excluding IDD samples did not change much of the results of the association tests (Table S6, available as supplemental digital content at http://links.lww.com/PAIN/B790), we concluded that an effect on the nervous system is driving its association with the report of pain, including back pain. Furthermore, other pain sites, including headaches, were also nominally significant in our analysis (Table S2, available as supplemental digital content at http://links.lww.com/PAIN/B790), reinforcing the effect of the gene on overall pain states. Finally, regulation of *SLC13A1* expression in animal neuropathic pain models also suggests at least a partial neurogenic origin of SLC13A1-depended pain.

Recently, rare variants of *SLC13A1* have been associated with intervertebral disc disorder, and its role in disc pathology has been proposed.[4] These results were obtained from the meta-analysis of 4 large-scale cohorts, including UKBB. Our results add to the body of our knowledge of the phenotypes associated with *SLC13A1.* This study by Bjornsdottir et al.[4] focused on 2 back pain pathologies diagnosed by the International Statistical Classification of Diseases (ICD-10): IDD and dorsalgia, with a focus on IDD. The authors reported that *SLC13A1* associated with IDD in a LoF burden test, with risk effects from the rare LoF variants rs28364172 (OR = 1.4, P = 2.5E-08) and rs138275989 (OR = 1.4, P = 1.2E-4). We found an even stronger contribution by variants rs754331818 (OR = 6.3, P = 2.7E-06) and rs1563320610 (OR = 4.6, P = 9.7E-02) for the report of back pain. Interestingly, both of these variants are frameshift rather than stop_gained mutations but were not identified in the study of Bjornsdottir et al.[4] More importantly, the evidence from our study supports the notion that *SLC13A1* contributes to the larger body of pain phenotypes, including headache and overlapping pain conditions. Furthermore, the strength of association with a report of back pain was much stronger than that with IDD if directly compared in UKBB (Table S6, available as supplemental digital content at http://links.lww.com/PAIN/B790), and when IDD was used as a covariate or IDD samples were removed from analysis, the associations with back pain report remain significant; in fact, it became stronger (Table S6, available as supplemental digital content at http://links.lww.com/PAIN/B790). A further investigation between IDD common variants reported by Bjornsdottir et al.[4] and their association with multisite pain showed that majority of IDD common variants did not contribute to multisite pain, indicating distinct development processes contributing to IDD and other chronic pain conditions (Table S8, available as supplemental digital content at http://links.lww.com/PAIN/B790). As a contrast, association between dorsalgia common variants and multisite pain in the study by Bjornsdottir et al. revealed that most of the variants associated with multisite pain, suggesting dorsalgia and other chronic pain conditions, share some common characteristics (Table S8, available as supplemental digital content at http://links.lww.com/PAIN/B790). Finally, when we tested the differential expression of *SLC13A1* in rat assays of neuropathic pain, we found that it was differentially expressed in sciatic nerve and DRGs following disease progression, consistent with the peripheral nervous system route of *SLC13A1* contribution to back pain. However, because the pleiotropic effect of *SLC13A1* gene on multiple systems is evident from the KO mice phenotypes,[10,11,31] it is likely that *SLC13A1* contributes to back pain through more than one route, both neurological and musculoskeletal pathways, yielding the strongest contribution of the gene to back pain vs other pain phenotypes.

In summary, our study identified *SLC13A1* as a candidate gene contributing to chronic pain. Further studies will be needed to detect the entire multiplicity of the biological pathways, thereby SLC13A1 contributes to the pain experience, as well as the multiplicity and impact of the metabolic, musculoskeletal, and neurological phenotypes that could be affected by functional variants in *SLC13A1* gene. Nevertheless, our results together with the results of others[4] stress the critical role of the *SLC13A1* gene in sulfate homeostasis and the importance of the optimal sulfate levels in preventing pain states. These findings open the door for further exploration of the impact of the *SLC13A1* gene on chronic pain.

Conflict of interest statement

L. Diatchenko was/is a consultant for Duke University, ONO PHARMA USA Inc, Releviate Inc, and Orthogen AG. All other authors declare that they have no competing interests.

Acknowledgements

The authors thank all participants of this study. This study was conducted under UK biobank application no. 20802. The Trøndelag Health Study (HUNT) is a collaboration between HUNT Research Centre (Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology NTNU), Trøndelag County Council, Central Norway Regional Health Authority, and the Norwegian Institute of Public Health. The genotyping was financed by the National Institute of Health (NIH), University of Michigan, The Norwegian Research council, and Central Norway Regional Health Authority and the Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU). The genotype quality control and imputation has been conducted by the K.G. Jebsen center for genetic epidemiology, Department of public health and nursing, Faculty of medicine and health sciences, Norwegian University of Science and Technology (NTNU). The "HUNT All-In Pain" working group: Sigrid Børte, Ben M. Brumpton, Egil A. Fors, Lars G. Fritsche, Maiken E. Gabrielsen, Knut Hagen, Ingrid Heuch, Oddgeir L. Holmen, Kristian Hveem, Espen S. Kristoffersen, Marie U. Lie, Amy E. Martinsen, Ingunn Mundal, Jonas B. Nielsen, Kristian Bernhard Nilsen, Linda M. Pedersen, Anne Heidi Skogholt, Synne Ø. Stensland, Kjersti Storheim, Laurent F. Thomas, Cristen J. Willer, Wei Zhou, John-Anker Zwart. The authors acknowledge contributions in data processing, analysis, post-analysis quality control and student supervision for the CLSA back pain genomewide association scan by Nikolay Dimitrov, Sahir Bhatnagar, Peter Her and Lovéni Hanumunthadu. Supported by Pfizer Canada Professorship in Pain Research, the Canadian Excellence Research Chairs grant CERC09, and the Canadian Institutes of Health Research (CIHR) Grant # 81528 to L. Diatchenko.

Author contributions: X. Ao performed bioinformatics analyses and data visualisation, interpreted the results, and drafted and revised the manuscript. M. Parisien conducted the analysis, visualisation, and interpretation of pain assays. A. V. Grant oversaw CLSA phenotype classification, QC of CLSA genotyping data and association testing and participated in interpreting of findings. M. Zidan participated in CLSA data processing and

analysis. A. E. Martinsen and B. S. Winsvold analyzed and interpreted the results from HUNT cohort. L. Diatchenko contributed to the study conceptualization, funding acquisition, study design analysis, result interpretation, and manuscript writing. All authors read and approved the final version of the manuscript.

Data and materials availability: The sequencing and/or imputation data were downloaded from UK biobank, HUNT and CLSA. Expression data was retrieved from GSE30165 and GSE24982.

The "HUNT All-In Pain" working group: Sigrid Børte[1,2,3], Ben M. Brumpton[2], Egil A. Fors[4], Lars G. Fritsche[5], Maiken E. Gabrielsen[2], Knut Hagen[6], Ingrid Heuch[7], Oddgeir L. Holmen[8], Kristian Hveem[2,8,9], Espen S. Kristoffersen[10,11,7], Marie U. Lie[1,3], Amy E. Martinsen[7,1,2], Ingunn Mundal[12], Jonas B. Nielsen[2,13], Kristian Bernhard Nilsen[14,6], Linda M. Pedersen[7], Anne Heidi Skogholt[2], Synne Ø. Stensland[3,15], Kjersti Storheim[3,16], Laurent F. Thomas[2,17,18,19], Cristen J. Willer[13], Wei Zhou[20,21], John-Anker Zwart[7,1,2]

[1]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway. [2]K. G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. [3]Research and Communication Unit for Musculoskeletal Health (FORMI), Department of Research and Innovation, Division of Clinical Neuroscience, Oslo University Hospital, Oslo, Norway. [4]Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. [5]Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, MI, United States. [6]Department of Neuromedicine and Movement Science, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. [7]Department of Research and Innovation, Division of Clinical Neuroscience, Oslo University Hospital, Oslo, Norway. [8]HUNT Research Center, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. [9]Department of Research, Innovation and Education, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway. [10]Department of General Practice, University of Oslo, Oslo, Norway. [11]Department of Neurology, Akershus University Hospital, Lørenskog, Norway. [12]Department of Health Science, Molde University College, Molde, Norway. [13]Department of Internal Medicine, Division of Cardiovascular Medicine, University of Michigan, Ann Arbor, MI, United States. [14]Department of Neurology, Oslo University Hospital, Oslo, Norway. [15]Norwegian Centre for Violence and Traumatic Stress Studies, Oslo, Norway. [16]Department of physiotherapy, Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway. [17]Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. [18]BioCore—Bioinformatics Core Facility, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. [19]Clinic of Laboratory Medicine, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway. [20]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, United States. [21]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, United States.

Appendix A. Supplemental digital content

Supplemental digital content associated with this article can be found online at http://links.lww.com/PAIN/B790.

Article history:
Received 20 July 2022
Received in revised form 14 November 2022
Accepted 16 December 2022
Available online 22 March 2023

References

[1] Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet 2013;76:7.20.1–41.

[2] Anttila V, Stefansson H, Kallela M, Todt U, Terwindt GM, Calafato MS, Nyholt DR, Dimas AS, Freilinger T, Müller-Myhsok B, Artto V, Inouye M, Alakurtti K, Kaunisto MA, Hämäläinen E, de Vries B, Stam AH, Weller CM, Heinze A, Heinze-Kuhn K, Goebel I, Borck G, Göbel H, Steinberg S, Wolf C, Björnsson A, Gudmundsson G, Kirchmann M, Hauge A, Werge T, Schoenen J, Eriksson JG, Hagen K, Stovner L, Wichmann H-E, Meitinger T, Alexander M, Moebus S, Schreiber S, Aulchenko YS, Breteler MMB, Uitterlinden AG, Hofman A, van Duijn CM, Tikka-Kleemola P, Vepsäläinen S, Lucae S, Tozzi F, Muglia P, Barrett J, Kaprio J, Färkkilä M, Peltonen L, Stefansson K, Zwart J-A, Ferrari MD, Olesen J, Daly M, Wessman M, van den Maagdenberg AMJM, Dichgans M, Kubisch C, Dermitzakis ET, Frants RR, Palotie A; International Headache Genetics Consortium. Genome-wide association study of migraine implicates a common susceptibility variant on 8q22.1. Nat Genet 2010;42:869–73.

[3] Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A. NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res 2012;41:D991–5.

[4] Bjornsdottir G, Stefansdottir L, Thorleifsson G, Sulem P, Norland K, Ferkingstad E, Oddsson A, Zink F, Lund SH, Nawaz MS, Bragi Walters G, Skuladottir AT, Gudjonsson SA, Einarsson G, Halldorsson GH, Bjarnadottir V, Sveinbjornsson G, Helgadottir A, Styrkarsdottir U, Gudmundsson LJ, Pedersen OB, Hansen TF, Werge T, Banasik K, Troelsen A, Skou ST, Thorner LW, Erikstrup C, Nielsen KR, Mikkelsen S, Andersen S, Brunak S, Burgdorf K, Hjalgrim H, Jemec G, Jennum P, Johansson PI, Nielsen KR, Nyegaard M, Bruun MT, Pedersen OB, Dinh KM, Sorensen E, Ostrowski S, Johansson PI, Gudbjartsson D, Stefansson H, Þorsteinsdottir U, Larsen MAH, Didriksen M, Sækmose S, Zeggini E, Hatzikotoulas K, Southam L, Gilly A, Barysenka A, van Meurs JBJ, Boer CG, Uitterlinden AG, Styrkarsdottir U, Stefansdottir L, Jonsson H, Ingvarsson T, Esko T, Magi R, Teder-Laving M, Ikegawa S, Terao C, Takuwa H, Meulenbelt I, Coutinho de Almeida R, Kloppenburg M, Tuerlings M, Slagboom PE, Nelissen RRGHH, Valdes AM, Mangino M, Tsezou A, Zengini E, Alexiadis G, Babis GC, Cheah KSE, Wu TT, Samartzis D, Cheung JPY, Sham PC, Kraft P, Kang JH, Hveem K, Zwart JA, Luetge A, Skogholt AH, Johnsen MB, Thomas LF, Winsvold B, Gabrielsen ME, Lee MTM, Zhang Y, Lietman SA, Shivakumar M, Smith GD, Tobias JH, Hartley A, Gaunt TR, Zheng J, Wilkinson JM, Steinberg J, Morris AP, Jonsdottir I, Bjornsson A, Olafsson IH, Ulfarsson E, Blondal J, Vikingsson A, Brunak S, Ostrowski SR, Ullum H, Thorsteinsdottir U, Stefansson H, Gudbjartsson DF, Thorgeirsson TE, Stefansson K. Rare SLC13A1 variants associate with intervertebral disc disorder highlighting role of sulfate in disc pathology. Nat Commun 2022;13:634.

[5] Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P, Marchini J. The UK Biobank resource with deep phenotyping and genomic data. Nature 2018;562:203–9.

[6] Chasman DI, Schürks M, Anttila V, de Vries B, Schminke U, Launer LJ, Terwindt GM, van den Maagdenberg AMJM, Fendrich K, Volzke H, Ernst F, Griffiths LR, Buring JE, Kallela M, Freilinger T, Kubisch C, Ridker PM, Palotie A, Ferrari MD, Hoffmann W, Zee RYL, Kurth T. Genome-wide association study reveals three susceptibility loci for common migraine in the general population. Nat Genet 2011;43:695–8.

[7] Clark MJ, Chen R, Lam HYK, Karczewski KJ, Chen R, Euskirchen G, Butte AJ, Snyder M. Performance comparison of exome DNA sequencing technologies. Nat Biotechnol 2011;29:908–14.

[8] Cox HC, Lea RA, Bellis C, Carless M, Dyer TD, Curran J, Charlesworth J, Macgregor S, Nyholt D, Chasman D, Ridker PM, Schurks M, Blangero J, Griffiths LR. A genome-wide analysis of "Bounty" descendants implicates several novel variants in migraine susceptibility. Neurogenetics 2012;13:261–6.

[9] Davis S, Meltzer PS. GEOquery: a bridge between the gene expression omnibus (GEO) and BioConductor. Bioinformatics 2007;23:1846–7.

[10] Dawson PA, Beck L, Markovich D. Hyposulfatemia, growth retardation, reduced fertility, and seizures in mice lacking a functional NaS $_i$ -1 gene. Proc Natl Acad Sci 2003;100:13704–9.
[11] Dawson PA, Steane SE, Markovich D. Impaired memory and olfactory performance in NaSi-1 sulphate transporter deficient mice. Behav Brain Res 2005;159:15–20.
[12] Dawson PA, Choyce A, Chuang C, Whitelock J, Markovich D, Leggatt GR. Enhanced tumor growth in the NaS1 sulfate transporter null mouse. Cancer Sci 2010;101:369–73.
[13] Diatchenko L, Slade GD, Nackley AG, Bhalang K, Sigurdsson A, Belfer I, Goldman D, Xu K, Shabalina SA, Shagin D, Max MB, Makarov SS, Maixner W. Genetic basis for individual variations in pain perception and the development of a chronic pain condition. Hum Mol Genet 2005;14: 135–43.
[14] Dong W, Jin SC, Allocco A, Zeng X, Sheth AH, Panchagnula S, Castonguay A, Lorenzo LE, Islam B, Brindle G, Bachand K, Hu J, Sularz A, Gaillard J, Choi J, Dunbar A, Nelson-Williams C, Kiziltug E, Furey CG, Conine S, Duy PQ, Kundishora AJ, Loring E, Li B, Lu Q, Zhou G, Liu W, Li X, Sierant MC, Mane S, Castaldi C, Lopez-Giraldez F, Knight JR, Sekula RF, Simard JM, Eskandar EN, Gottschalk C, Moliterno J, Gunel M, Gerrard JL, Dib-Hajj S, Waxman SG, Barker FG, Alper SL, Chahine M, Haider S, De Koninck Y, Lifton RP, Kahle KT. Exome sequencing implicates impaired GABA signaling and neuronal ion transport in trigeminal neuralgia. iScience 2020;23:101552.
[15] Freilinger T, Anttila V, de Vries B, Malik R, Kallela M, Terwindt GM, Pozo-Rosich P, Winsvold B, Nyholt DR, van Oosterhout WPJ, Artto V, Todt U, Hamalainen E, Fernandez-Morales J, Louter MA, Kaunisto MA, Schoenen J, Raitakari O, Lehtimaki T, Vila-Pueyo M, Gobel H, Wichmann E, Sintas C, Uitterlinden AG, Hofman A, Rivadeneira F, Heinze A, Tronvik E, van Duijn CM, Kaprio J, Cormand B, Wessman M, Frants RR, Meitinger T, Muller-Myhsok B, Zwart JA, Farkkila M, Macaya A, Ferrari MD, Kubisch C, Palotie A, Dichgans M, van den Maagdenberg AMJM. Genome-wide association analysis identifies susceptibility loci for migraine without aura. Nat Genet 2012;44: 777–82.
[16] Fritsche LG, Igl W, Bailey JNC, Grassmann F, Sengupta S, Bragg-Gresham JL, Burdon KP, Hebbring SJ, Wen C, Gorski M, Kim IK, Cho D, Zack D, Souied E, Scholl HPN, Bala E, Lee KE, Hunter DJ, Sardell RJ, Mitchell P, Merriam JE, Cipriani V, Hoffman JD, Schick T, Lechanteur YTE, Guymer RH, Johnson MP, Jiang Y, Stanton CM, Buitendijk GHS, Zhan X, Kwong AM, Boleda A, Brooks M, Gieser L, Ratnapriya R, Branham KE, Foerster JR, Heckenlively JR, Othman MI, Vote BJ, Liang HH, Souzeau E, McAllister IL, Isaacs T, Hall J, Lake S, Mackey DA, Constable IJ, Craig JE, Kitchner TE, Yang Z, Su Z, Luo H, Chen D, Ouyang H, Flagg K, Lin D, Mao G, Ferreyra H, Stark K, von Strachwitz CN, Wolf A, Brandl C, Rudolph G, Olden M, Morrison MA, Morgan DJ, Schu M, Ahn J, Silvestri G, Tsironi EE, Park KH, Farrer LA, Orlin A, Brucker A, Li M, Curcio CA, Mohand-Said S, Sahel JA, Audo I, Benchaboune M, Cree AJ, Rennie CA, Goverdhan SV, Grunin M, Hagbi-Levi S, Campochiaro P, Katsanis N, Holz FG, Blond F, Blanche H, Deleuze JF, Igo RP Jr, Truitt B, Peachey NS, Meuer SM, Myers CE, Moore EL, Klein R, Hauser MA, Postel EA, Courtenay MD, Schwartz SG, Kovach JL, Scott WK, Liew G, Tan AG, Gopinath B, Merriam JC, Smith RT, Khan JC, Shahid H, Moore AT, McGrath JA, Laux R, Brantley MA Jr, Agarwal A, Ersoy L, Caramoy A, Langmann T, Saksens NTM, de Jong EK, Hoyng CB, Cain MS, Richardson AJ, Martin TM, Blangero J, Weeks DE, Dhillon B, van Duijn CM, Doheny KF, Romm J, Klaver CCW, Hayward C, Gorin MB, Klein ML, Baird PN, den Hollander AI, Fauser S, Yates JRW, Allikmets R, Wang JJ, Schaumberg DA, Klein BEK, Hagstrom SA, Chowers I, Lotery AJ, Leveillard T, Zhang K, Brilliant MH, Hewitt AW, Swaroop A, Chew EY, Pericak-Vance MA, DeAngelis M, Stambolian D, Haines JL, Iyengar SK, Weber BHF, Abecasis GR, Heid IM. A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. Nat Genet 2016;48:134–43.
[17] Goh G, Choi M. Application of whole exome sequencing to identify disease-causing variants in inherited human diseases. Genomics Inform 2012;10:214–9.
[18] Hautakangas H, Winsvold BS, Ruotsalainen SE, Bjornsdottir G, Harder AVE, Kogelman LJA, Thomas LF, Noordam R, Benner C, Gormley P, Artto V, Banasik K, Bjornsdottir A, Boomsma DI, Brumpton BM, Burgdorf KS, Buring JE, Chalmer MA, de Boer I, Dichgans M, Erikstrup C, Farkkila M, Garbrielsen ME, Ghanbari M, Hagen K, Happola P, Hottenga JJ, Hrafnsdottir MG, Hveem K, Johnsen MB, Kahonen M, Kristoffersen ES, Kurth T, Lehtimaki T, Lighart L, Magnusson SH, Malik R, Pedersen OB, Pelzer N, Penninx BWJH, Ran C, Ridker PM, Rosendaal FR, Sigurdardottir GR, Skogholt AH, Sveinsson OA, Thorgeirsson TE, Ullum H, Vijfhuizen LS, Widen E, van Dijk KW, de Boer I, van den Maagdenberg AMJM, Aromaa A, Belin AC, Freilinger T, Ikram MA, Jarvelin MR, Raitakari OT, Terwindt GM, Kallela M, Wessman M, Olesen J, Chasman DI, Nyholt DR, Stefansson H, Stefansson K, van den Maagdenberg AMJM, Hansen TF, Ripatti S, Zwart JA, Palotie A, Pirinen M. Genome-wide analysis of 102,084 migraine cases identifies 123 risk loci and subtype-specific risk alleles. Nat Genet 2022;54:152–60.
[19] Heuch I, Heuch I, Hagen K, Zwart JA. A comparison of anthropometric measures for assessing the association between body size and risk of chronic low back pain: the HUNT study. PLoS One 2015;10:e0141268.
[20] Holmen TL, Bratberg G, Krokstad S, Langhammer A, Hveem K, Midthjell K, Heggland J, Holmen J. Cohort profile of the Young-HUNT Study, Norway: a population-based study of adolescents. Int J Epidemiol 2014; 43:536–44.
[21] Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer-Berk M, England EM, Seaby EG, Kosmicki JA, Walters RK, Tashman K, Farjoun Y, Banks E, Poterba T, Wang A, Seed C, Whiffin N, Chong JX, Samocha KE, Pierce-Hoffman E, Zappala Z, O'Donnell-Luria AH, Minikel EV, Weisburd B, Lek M, Ware JS, Vittal C, Armean IM, Bergelson L, Cibulskis K, Connolly KM, Covarrubias M, Donnelly S, Ferriera S, Gabriel S, Gentry J, Gupta N, Jeandet T, Kaplan D, Llanwarne C, Munshi R, Novod S, Petrillo N, Roazen D, Ruano-Rubio V, Saltzman A, Schleicher M, Soto J, Tibbetts K, Tolonen C, Wade G, Talkowski ME, Genome Aggregation Database Consortium, Neale BM, Daly MJ, MacArthur DG. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020;581:434–43.
[22] Krokstad S, Langhammer A, Hveem K, Holmen T, Midthjell K, Stene T, Bratberg G, Heggland J, Holmen J. Cohort profile: the HUNT study, Norway. Int J Epidemiol 2013;42:968–77.
[23] Kuchenbaecker K, Appel EVR. Assessing rare variation in complex traits. Methods Mol Biol 2018;1793:51–71.
[24] Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009;4:1073–81.
[25] Kurth I, Baumgartner M, Schabhüttl M, Tomni C, Windhager R, Strom TM, Wieland T, Gremel K, Auer-Grumbach M. Whole exome sequencing in congenital pain insensitivity identifies a novel causative intronic *NTRK1*-mutation due to uniparental disomy. Am J Med Genet B Neuropsychiatr Genet 2016;171:875–8.
[26] Lelieveld SH, Spielmann M, Mundlos S, Veltman JA, Gilissen C. Comparison of exome and genome sequencing technologies for the complete capture of protein-coding regions. Hum Mutat 2015;36:815–22.
[27] Lewontin RC. The interaction of selection and linkage. I. General considerations; heterotic models. Genetics 1964;49:49–67.
[28] Li S, Liu Q, Wang Y, Gu Y, Liu D, Wang C, Ding G, Chen J, Liu J, Gu X. Differential gene expression profiling and biological process analysis in proximal nerve segments after sciatic nerve transection. PLoS One 2013; 8:e57000.
[29] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.
[30] Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics 2015;31:3555–7.
[31] Markovich D. Na+–sulfate cotransporter SLC13A1. Pflügers Arch 2014; 466:131–7.
[32] Marouli E, Graff M, Medina-Gomez C, Lo KS, Wood AR, Kjaer TR, Fine RS, Lu Y, Schurmann C, Highland HM, Rueger S, Thorleifsson G, Justice AE, Lamparter D, Stirrups KE, Turcot V, Young KL, Winkler TW, Esko T, Karaderi T, Locke AE, Masca NGD, Ng MCY, Mudgal P, Rivas MA, Vedantam S, Mahajan A, Guo X, Abecasis G, Aben KK, Adair LS, Alam DS, Albrecht E, Allin KH, Allison M, Amouyel P, Appel EV, Arveiler D, Asselbergs FW, Auer PL, Balkau B, Banas B, Bang LE, Benn M, Bergmann S, Bielak LF, Bluher M, Boeing H, Boerwinkle E, Boger CA, Bonnycastle LL, Bork-Jensen J, Bots ML, Bottinger EP, Bowden DW, Brandslund I, Breen G, Brilliant MH, Broer L, Burt AA, Butterworth AS, Carey DJ, Caulfield MJ, Chambers JC, Chasman DI, Chen YDI, Chowdhury R, Christensen C, Chu AY, Cocca M, Collins FS, Cook JP, Corley J, Galbany JC, Cox AJ, Cuellar-Partida G, Danesh J, Davies G, de Bakker PIW, de Borst GJ, de Denus S, de Groot MCH, de Mutsert R, Deary IJ, Dedoussis G, Demerath EW, den Hollander AI, Dennis JG, Di Angelantonio E, Drenos F, Du M, Dunning AM, Easton DF, Ebeling T, Edwards TL, Ellinor PT, Elliott P, Evangelou E, Farmaki AE, Faul JD, Feitosa MF, Feng S, Ferrannini E, Ferrario MM, Ferrieres J, Florez JC, Ford I, Fornage M, Franks PW, Frikke-Schmidt R, Galesloot TE, Gan W, Gandin I, Gasparini P, Giedraitis V, Giri A, Girotto G, Gordon SD, Gordon-Larsen P, Gorski M, Grarup N, Grove ML, Gudnason V, Gustafsson S, Hansen T, Harris KM, Harris TB, Hattersley AT, Hayward C, He L, Heid IM, Heikkila K, Helgeland O, Hernesniemi J, Hewitt AW, Hocking LJ, Hollensted M, Holmen OL, Hovingh GK, Howson JMM, Hoyng CB,

Huang PL, Hveem K, Ikram MA, Ingelsson E, Jackson AU, Jansson JH, Jarvik GP, Jensen GB, Jhun MA, Jia Y, Jiang X, Johansson S, Jorgensen ME, Jorgensen T, Jousilahti P, Jukema JW, Kahali B, Kahn RS, Kahonen M, Kamstrup PR, Kanoni S, Kaprio J, Karaleftheri M, Kardia SLR, Karpe F, Kee F, Keeman R, Kiemeney LA, Kitajima H, Kluivers KB, Kocher T, Komulainen P, Kontto J, Kooner JS, Kooperberg C, Kovacs P, Kriebel J, Kuivaniemi H, Kury S, Kuusisto J, La Bianca M, Laakso M, Lakka TA, Lange EM, Lange LA, Langefeld CD, Langenberg C, Larson EB, Lee IT, Lehtimaki T, Lewis CE, Li H, Li J, Li-Gao R, Lin H, Lin LA, Lin X, Lind L, Lindstrom J, Linneberg A, Liu Y, Liu Y, Lophatananon A, Luan J, Lubitz SA, Lyytikainen LP, Mackey DA, Madden PAF, Manning AK, Mannisto S, Marenne G, Marten J, Martin NG, Mazul AL, Meidtner K, Metspalu A, Mitchell P, Mohlke KL, Mook-Kanamori DO, Morgan A, Morris AD, Morris AP, Muller-Nurasyid M, Munroe PB, Nalls MA, Nauck M, Nelson CP, Neville M, Nielsen SF, Nikus K, Njolstad PR, Nordestgaard BG, Ntalla I, O'Connel JR, Oksa H, Loohuis LMO, Ophoff RA, Owen KR, Packard CJ, Padmanabhan S, Palmer CNA, Pasterkamp G, Patel AP, Pattie A, Pedersen O, Peissig PL, Peloso GM, Pennell CE, Perola M, Perry JA, Perry JRB, Person TN, Pirie A, Polasek O, Posthuma D, Raitakari OT, Rasheed A, Rauramaa R, Reilly DF, Reiner AP, Renstrom F, Ridker PM, Rioux JD, Robertson N, Robino A, Rolandsson O, Rudan I, Ruth KS, Saleheen D, Salomaa V, Samani NJ, Sandow K, Sapkota Y, Sattar N, Schmidt MK, Schreiner PJ, Schulze MB, Scott RA, Segura-Lepe MP, Shah S, Sim X, Sivapalaratnam S, Small KS, Smith AV, Smith JA, Southam L, Spector TD, Speliotes EK, Starr JM, Steinthorsdottir V, Stringham HM, Stumvoll M, Surendran P, 't Hart LM, Tansey KE, Tardif JC, Taylor KD, Teumer A, Thompson DJ, Thorsteinsdottir U, Thuesen BH, Tonjes A, Tromp G, Trompet S, Tsafantakis E, Tuomilehto J, Tybjaerg-Hansen A, Tyrer JP, Uher R, Uitterlinden AG, Ulivi S, van der Laan SW, Van Der Leij AR, van Duijn CM, van Schoor NM, van Setten J, Varbo A, Varga TV, Varma R, Edwards DRV, Vermeulen SH, Vestergaard H, Vitart V, Vogt TF, Vozzi D, Walker M, Wang F, Wang CA, Wang S, Wang Y, Wareham NJ, Warren HR, Wessel J, Willems SM, Wilson JG, Witte DR, Woods MO, Wu Y, Yaghootkar H, Yao J, Yao P, Yerges-Armstrong LM, Young R, Zeggini E, Zhan X, Zhang W, Zhao JH, Zhao W, Zhao W, Zheng H, Zhou W, EPIC-InterAct Consortium, CHD Exome+ Consortium, ExomeBP Consortium, T2D-Genes Consortium, GoT2D Genes Consortium, Global Lipids Genetics Consortium, ReproGen Consortium, MAGIC Investigators, Rotter JI, Boehnke M, Kathiresan S, McCarthy MI, Willer CJ, Stefansson K, Borecki IB, Liu DJ, North KE, Heard-Costa NL, Pers TH, Lindgren CM, Oxvig C, Kutalik Z, Rivadeneira F, Loos RJF, Frayling TM, Hirschhorn JN, Deloukas P, Lettre G. Rare and low-frequency coding variants alter human adult height. Nature 2017; 542:186–90.

[33] McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P, Cunningham F. The Ensembl variant effect predictor. Genome Biol 2016; 17:122.

[34] Meloto CB, Segall SK, Smith S, Parisien M, Shabalina SA, Rizzatti-Barbosa CM, Gauthier J, Tsao D, Convertino M, Piltonen MH, Slade GD, Fillingim RB, Greenspan JD, Ohrbach R, Knott C, Maixner W, Zaykin D, Dokholyan NV, Reenila I, Mannisto PT, Diatchenko L. COMT gene locus: new functional variants. PAIN 2015;156:2072–83.

[35] Mogil JS. Pain genetics: past, present and future. Trends Genet 2012;28: 258–66.

[36] Nielsen C, Knudsen G, Steingrímsdóttir Ó. Twin studies of pain. Clin Genet 2012;82:331–40.

[37] Palada V, Kaunisto MA, Kalso E. Genetics and genomics in postoperative pain and analgesia. Curr Opin Anaesthesiol 2018;31:569–74.

[38] Rabbani B, Tekin M, Mahdieh N. The promise of whole-exome sequencing in medical genetics. J Hum Genet 2014;59:5–15.

[39] Raina PS, Wolfson C, Kirkland SA, Griffith LE, Oremus M, Patterson C, Tuokko H, Penning M, Balion CM, Hogan D, Wister A, Payette H, Shannon H, Brazil K. The Canadian longitudinal study on aging (CLSA). Can J Aging 2009;28:221–9.

[40] Raina P, Wolfson C, Kirkland S, Griffith LE, Balion C, Cossette B, Dionne I, Hofer S, Hogan D, van den Heuvel ER, Liu-Ambrose T, Menec V, Mugford G, Patterson C, Payette H, Richards B, Shannon H, Sheets D, Taler V, Thompson M, Tuokko H, Wister A, Wu C, Young L. Cohort profile: the Canadian longitudinal study on aging (CLSA). Int J Epidemiol 2019;48: 1752–3j.

[41] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.

[42] Sexton JE, Cox JJ, Zhao J, Wood JN. The genetics of pain: implications for therapeutics. Annu Rev Pharmacol Toxicol 2018;58:123–42.

[43] Szustakowski JD, Balasubramanian S, Kvikstad E, Khalid S, Bronson PG, Sasson A, Wong E, Liu D, Wade Davis J, Haefliger C, Katrina Loomis A, Mikkilineni R, Noh HJ, Wadhawan S, Bai X, Hawes A, Krasheninina O, Ulloa R, Lopez AE, Smith EN, Waring JF, Whelan CD, Tsai EA, Overton JD, Salerno WJ, Jacob H, Szalma S, Runz H, Hinkle G, Nioi P, Petrovski S, Miller MR, Baras A, Mitnaul LJ, Reid JG, Moiseyenko O, Rios C, Saha S, Abecasis G, Banerjee N, Beechert C, Boutkov B, Cantor M, Coppola G, Economides A, Eom G, Forsythe C, Fuller ED, Gu Z, Habegger L, Jones MB, Lanche R, Lattari M, LeBlanc M, Li D, Lotta LA, Manoochehri K, Mansfield AJ, Maxwell EK, Mighty J, Nafde M, O'Keeffe S, Orelus M, Padilla MS, Panea R, Polanco T, Pradhan M, Rasool A, Schleicher TD, Sharma D, Shuldiner A, Staples JC, Van Hout CV, Widom L, Wolf SE, John S, Chen CY, Sexton D, Kupelian V, Marshall E, Swan T, Eaton S, Liu JZ, Loomis S, Jensen M, Duraisamy S, Tetrault J, Merberg D, Badola S, Reppell M, Grundstad J, Zheng X, Deaton AM, Parker MM, Ward LD, Flynn-Carroll AO, Austin C, March R, Pangalos MN, Platt A, Snowden M, Matakidou A, Wasilewski S, Wang Q, Deevi S, Carss K, Smith K, Sogaard M, Hu X, Chen X, Ye Z. Advancing human genetics research and drug discovery through exome sequencing of the UK Biobank. Nat Genet 2021;53:942–8.

[44] Taliun D, Harris DN, Kessler MD, Carlson J, Szpiech ZA, Torres R, Taliun SAG, Corvelo A, Gogarten SM, Kang HM, Pitsillides AN, LeFaive J, Lee SB, Tian X, Browning BL, Das S, Emde AK, Clarke WE, Loesch DP, Shetty AC, Blackwell TW, Smith AV, Wong Q, Liu X, Conomos MP, Bobo DM, Aguet F, Albert C, Alonso A, Ardlie KG, Arking DE, Aslibekyan S, Auer PL, Barnard J, Barr RG, Barwick L, Becker LC, Beer RL, Benjamin EJ, Bielak LF, Blangero J, Boehnke M, Bowden DW, Brody JA, Burchard EG, Cade BE, Casella JF, Chalazan B, Chasman DI, Chen YI, Cho MH, Choi SH, Chung MK, Clish CB, Correa A, Curran JE, Custer B, Darbar D, Daya M, de Andrade M, DeMeo DL, Dutcher SK, Ellinor PT, Emery LS, Eng C, Fatkin D, Fingerlin T, Forer L, Fornage M, Franceschini N, Fuchsberger C, Fullerton SM, Germer S, Gladwin MT, Gottlieb DJ, Guo X, Hall ME, He J, Heard-Costa NL, Heckbert SR, Irvin MR, Johnsen JM, Johnson AD, Kaplan R, Kardia SLR, Kelly T, Kelly S, Kenny EE, Kiel DP, Klemmer R, Konkle BA, Kooperberg C, Köttgen A, Lange LA, Lasky-Su J, Levy D, Lin X, Lin KH, Liu C, Loos RJF, Garman L, Gerszten R, Lubitz SA, Lunetta KL, Mak ACY, Manichaikul A, Manning AK, Mathias RA, McManus DD, McGarvey ST, Meigs JB, Meyers DA, Mikulla JL, Minear MA, Mitchell BD, Mohanty S, Montasser ME, Montgomery C, Morrison AC, Murabito JM, Natale A, Natarajan P, Nelson SC, North KE, O'Connell JR, Palmer ND, Pankratz N, Peloso GM, Peyser PA, Pleiness J, Post WS, Psaty BM, Rao DC, Redline S, Reiner AP, Roden D, Rotter JI, Ruczinski I, Sarnowski C, Schoenherr S, Schwartz DA, Seo JS, Seshadri S, Sheehan VA, Sheu WH, Shoemaker MB, Smith NL, Smith JA, Sotoodehnia N, Stilp AM, Tang W, Taylor KD, Telen M, Thornton TA, Tracy RP, Van Den Berg DJ, Vasan RS, Viaud-Martinez KA, Vrieze S, Weeks DE, Weir BS, Weiss ST, Weng LC, Willer CJ, Zhang Y, Zhao X, Arnett DK, Ashley-Koch AE, Barnes KC, Boerwinkle E, Gabriel S, Gibbs R, Rice KM, Rich SS, Silverman EK, Qasba P, Gan W; NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium; Papanicolaou GJ, Nickerson DA, Browning SR, Zody MC, Zöllner S, Wilson JG, Cupples LA, Laurie CC, Jaquish CE, Hernandez RD, O'Connor TD, Abecasis GR. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. Nature 2021;590:290–9.

[45] UK10K Consortium, Walter K, Min JL, Huang J, Crooks L, Memari Y, McCarthy S, Perry JRB, Xu C, Futema M, Lawson D, Iotchkova V, Schiffels S, Hendricks AE, Danecek P, Li R, Floyd J, Wain LV, Barroso I, Humphries SE, Hurles ME, Zeggini E, Barrett JC, Plagnol V, Richards JB, Greenwood CMT, Timpson NJ, Durbin R, Soranzo N. The UK10K project identifies rare variants in health and disease. Nature 2015;526:82–90.

[46] Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190–1.

[47] Williams FMK, Spector TD, MacGregor AJ. Pain reporting at different body sites is explained by a single underlying genetic factor. Rheumatology 2010;49:1753–5.

[48] Zelaya CE, Dahlhamer JM, Lucas JW, Connor EM. Chronic pain and high-impact chronic pain among U.S. adults, 2019. NCHS Data Brief, no 390. Hyattsville, MD: National Center for Health Statistics. 2020.

[49] Zhou W, Nielsen JB, Fritsche LG, Dey R, Gabrielsen ME, Wolford BN, LeFaive J, VandeHaar P, Gagliano SA, Gifford A, Bastarache LA, Wei WQ, Denny JC, Lin M, Hveem K, Kang HM, Abecasis GR, Willer CJ, Lee S. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. Nat Genet 2018;50:1335–41.

[50] Zhou W, Zhao Z, Nielsen JB, Fritsche LG, LeFaive J, Gagliano Taliun SA, Bi W, Gabrielsen ME, Daly MJ, Neale BM, Hveem K, Abecasis GR, Willer CJ, Lee S. Scalable generalized linear mixed model for region-based association tests in large biobanks and cohorts. Nat Genet 2020;52: 634–9.

[51] Zook JM, McDaniel J, Olson ND, Wagner J, Parikh H, Heaton H, Irvine SA, Trigg L, Truty R, McLean CY, De La Vega FM, Xiao C, Sherry S, Salit M. An open resource for accurately benchmarking small variant and reference calls. Nat Biotechnol 2019;37:561–6.