DOI 10.18699/SBB-2022-259

Association of *PANX3* with chronic back pain discovered using gene-based analysis

Belonogova N.M.[1]*, Kirichenko A.V.[1], Freidin M.B.[2], Williams F.M.K.[2], Suri P.[3, 4, 5], Aulchenko Y.S.[1, 6], Axenovich T.I.[1], Tsepilov Y.A.[1, 6, 7]

[1] *Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia*

[2] *Department of Twin Research and Genetic Epidemiology, King's College London, London, UK*

[3] *Seattle Epidemiologic Research and Information Center, VA Puget Sound Health Care System, Seattle, USA*

[4] *Division of Rehabilitation Care Services, Seattle, USA*

[5] *Clinical Learning, Evidence, and Research Center, University of Washington, Seattle, USA*

[6] *Laboratory of Theoretical and Applied Functional Genomics, Novosibirsk State University, Novosibirsk, Russia*

[7] *Kurchatov Genomic Center of the Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia*

** belon@bionet.nsc.ru*

Key words: back pain, gene-based analysis, GWAS summary statistics, intervertebral disc disorder

Motivation and Aim: Back pain is the leading cause of years lived with disability worldwide, yet surprisingly little is known regarding the biology underlying this condition. The impact of genetics is known for the chronic back pain: its heritability is estimated to be at least 40 % [1]. Large genome-wide association studies (GWAS) have shown that common variation may account for up to 35 % of chronic back pain heritability; rare variants may explain a portion of the heritability not explained by common variants [2]. In this study, we aimed to analyze the genes whose variants are directly responsible for differences in chronic back pain risk between individuals using data from UK Biobank.

Methods and Algorithms: For gene-based association analysis we used GWAS summary statistics for chronic back pain obtained on the UK Biobank imputed data including rare variants with moderate imputation quality, N = 439,831 participants of European ancestry [3]. A sample of 265,000 randomly selected participants was used as the discovery sample. The remaining participants comprised a replication cohort. We used the sumFREGAT R-package [4] to perform three gene-based methods: SKAT-O, PCA, and ACAT-V. The results from these methods were combined using the aggregated Cauchy omnibus test, ACAT-O [5]. Genes having a combined ACAT-O p-value $\leq 2.5 \times 10^{-6}$ (standard gene-based significance threshold) were considered as genome-wide significant genes. Genes with p-value $\leq 2.5 \times 10^{-5}$ were considered as suggestively significant. For genes with genome-wide significant p-values, replication was carried out on the sample of remaining UK Biobank participants (N = 174,831). An additional validation step was performed using FinnGen Biobank GWAS results version 6 (Finngen-r6, https://r6.finngen.fi/).

Results: We obtained 39 suggestively significant genes on discovery stage. Two of them, *SOX5* and *PANX3*, reached the genome-wide significance threshold (2.5×10^{-6}) and were further replicated with p-value < 0.025 (0.05/2). The *SOX5* gene is well known back pain gene. The *PANX3* gene has not previously been described as having a role in chronic back pain. We showed that the association of *PANX3* with chronic back pain is

driven by rare non-coding intronic polymorphisms. Association of *PANX3* has been replicated on independent sample from UK Biobank and validated using similar phenotype, dorsalgia, from FinnGen Biobank. We also found that the *PANX3* gene is associated with intervertebral disc disorders on FinnGen sample. We can speculate that a possible mechanism of action of *PANX3* on the back pain is due to its effect on the intervertebral discs.

Conclusion: We found that *SOX5* and *PANX3* genes influenced chronic back pain. The *PANX3* gene has not previously been described as associated with chronic back pain. Its association was shown to be due to non-coding rare intron polymorphisms.

Acknowledgements: The study was conducted using the UK Biobank resource under application No. 59345. The work was supported by the Russian Foundation for Basic Research No. 20-04-00464, the budget project No. FWNR-2022-0020, the Russian Ministry of Education and Science under the 5-100 Excellence Programme, the Russian Science Foundation No. 22-15-20037, and NIH/NIAMS P30AR072572.

References

1. Hartvigsen J. et al. Heritability of spinal pain and consequences of spinal pain: a comprehensive genetic epidemiologic analysis using a population-based sample of 15,328 twins ages 20-71 years. *Arthritis Rheum*. 2009;61(10):1343-1351.
2. Wainschtein P. et al. Assessing the contribution of rare variants to complex trait heritability from whole-genome sequence data. *Nat Genet*. 2022;54(3):263-273.
3. Tsepilov Y.A. et al. Analysis of genetically independent phenotypes identifies shared genetic factors associated with chronic musculoskeletal pain conditions. *Commun Biol.* 2020;3(1):329.
4. Svishcheva G.R. et al. Gene-based association tests using GWAS summary statistics. *Bioinformatics.* 2019;35(19):3701-3708.
5. Liu Y. et al. ACAT: A fast and powerful p-value combination method for rare-variant analysis in sequencing studies. *Am J Hum Genet.* 2019;104(3):410-421.