doi 10.18699/bgrs2024-4.3-05

Efficiency of a new multi-trait approach applied to back pain-related phenotypes

Belonogova N.M.[1]*, Elgaeva E.E.[1, 2], Zorkoltseva I.V.[1], Kirichenko A.V.[1], Svishcheva G.R.[1, 3], Freidin M.B.[4], Williams F.M.K.[5], Suri P.[6, 7, 8, 9], Axenovich T.I.[1], Tsepilov Y.A.[1]

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

[2] *Department of Natural Sciences, Novosibirsk State University, Novosibirsk, Russia*

[3] *Vavilov Institute of General Genetics, RAS, Moscow, Russia*

[4] *Department of Biology, School of Biological and Behavioral Sciences, Queen Mary University of London, London, UK*

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

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

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

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

[9] *Department of Rehabilitation Medicine, University of Washington, Seattle, USA*

* *belon@bionet.nsc.ru*

Key words: chronic back pain; dorsalgia; intervertebral disc disorders; shared heritability; GWAS

Motivation and Aim: Back pain (BP) is a major contributor to disability worldwide. We conducted a cross-sectional study analyzing three BP-related phenotypes: chronic BP (CBP), dorsalgia and intervertebral disc disorders (IDD), with heritability estimated at 40–60 %. Less than half of the heritability is explained by common genetic variants identified by GWAS. More powerful methods of statistical analysis may offer additional insights. In this study, we utilized a novel multi-trait approach to merge three BP-related phenotypes: CBP, dorsalgia, and IDD, which have been previously used in BP studies. Dorsalgia is a phenotype based on electronic health records (EHR) and is defined by clinical diagnostic codes for back and neck pain, reflecting mainly the former. IDD is another phenotype related to BP and defined by EHR-based codes for intervertebral disc degeneration. These codes are typically used when BP is present, but not necessarily. There are high genetic correlations between BP-related phenotypes, estimated to be in the range of 87 to 92 % [1, 2], allowing for an approach that maximizes multi-trait heritability. The aim of this approach is to reduce the genetic heterogeneity of the traits and to identify a common genetic background for a number of traits that are genetically correlated [3].

Methods and Algorithms: The SHAHER framework aims to create a multi-trait phenotype from a set of genetically correlated traits and identify the genetic variants that control this trait [3]. The method assumes that the genetic background of each of the genetically correlated traits can be decomposed into two components: one common to all traits (the shared genetic impact or SGI), and the other specific to each trait. A new multi-trait phenotype, SGIT (the shared genetic impact trait), was created by linearly combining the original traits. The coefficients in this combination were defined by maximizing the heritability of SGIT. A GWAS of SGIT was performed using an

algorithm based on the summary statistics calculated for the original traits. The SHAHER method was previously described in our studies [2, 3].

Using imputed genotypes from the UK Biobank, we conducted a multi-trait gene-based association analysis of three BP-related phenotypes: CBP, dorsalgia and IDD. The first step of the analysis was performed using a discovery sample from the UK Biobank. Initially, the GWAS summary statistics for the three BP-related phenotypes were obtained. They were used for the calculation of the multi-trait (SGIT) summary statistics. Then we performed a gene-based association analysis (GBA) of the individual and multi-trait phenotypes. For all identified genes, conditional analysis (COJO) was applied. The second step of the analysis used a replication sample from the FinnGen database. We conducted a replication study by applying GBA to the 32 genes identified in the first step and using GWAS summary statistics for the two BP-related traits available on the FinnGen sample.

Results: We identified and replicated 16 genes associated with BP-related traits (Fig. 1). Seven of the detected genes, namely, *MIPOL1*, *PTPRC*, *RHOA*, *MAML3*, *JADE2*, *MLLT10* and *RERG*, have not been previously reported. The rest nine genes have been previously detected as associated with traits genetically correlated with BP or included in pathways associated with BP.

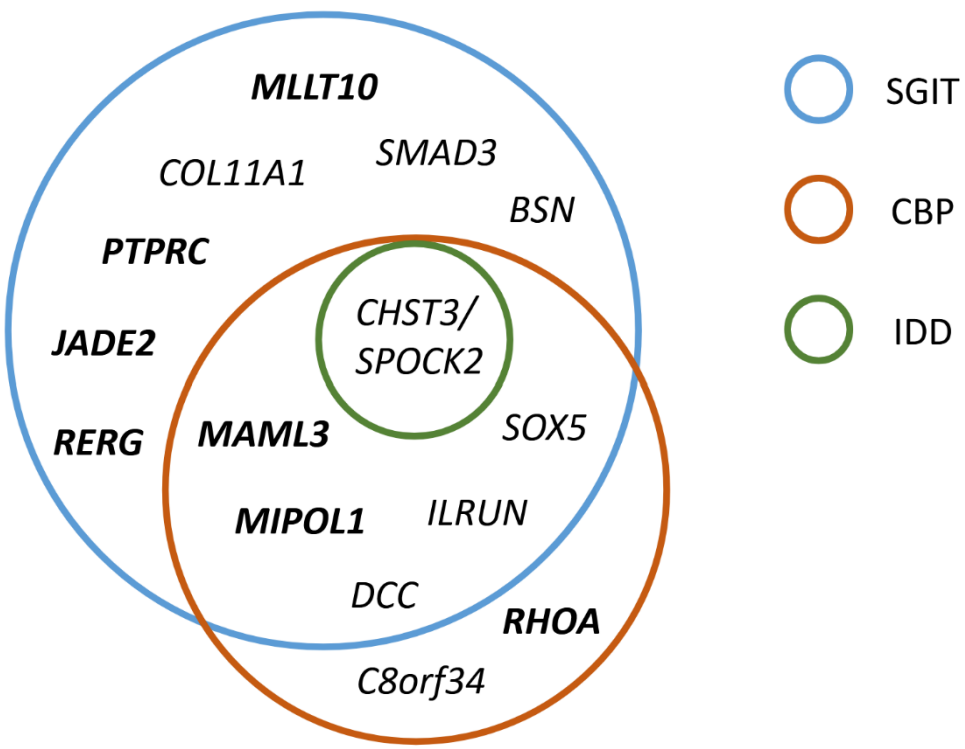

Fig. 1. Genes replicated on FinnGen data. Names of previously unreported genes are shown in bold

Conclusion: We conducted a genome-wide gene-based association analysis of BP-related phenotypes using imputed genotypes from a UK Biobank project and applying the SHAHER technique of multi-trait analysis [3].

SHAHER constructs a multi-trait phenotype as a linear combination of original traits, where the coefficients are estimated to maximize the heritability of the multi-trait phenotype. As a result, we identified seven genes associated with SGIT, which did not show a significant association with the original traits. Additional six genes were associated with both SGIT and CBP. These results confirm the effectiveness of the chosen strategy.

In total, we identified and replicated 16 genes significantly associated with BP-related traits. Thirteen genes were detected on the multi-trait phenotype that is in accordance with high genetic correlation between BP-related traits. Some genes have been previously described as associated with BP or with the genetically correlated traits, or as included in pathways associated with BP. Our results verify the role of these genes in BP-related traits and provide new insights into the genetics of back pain.

Funding: The work was supported by the Russian Science Foundation (RSF) grant No. 22-15-20037 and the Government of the Novosibirsk region.

Acknowledgments: This research has been conducted using the UK Biobank Resource under Applications No. 18219 and No. 59345. We want to acknowledge the participants and investigators of the FinnGen study.

References

1. Bjornsdottir G., Stefansdottir L., Thorleifsson G. et al. Rare *SLC13A1* variants associate with intervertebral disc disorder highlighting role of sulfate in disc pathology. *Nat Commun.* 2022;13(1):634. doi 10.1038/s41467-022-28167-1
2. Zorkoltseva I.V., Elgaeva E.E., Belonogova N.M., Kirichenko A.V. et al. Multi-trait exome-wide association study of back pain-related phenotypes. *Genes (Basel)*. 2023;14(10):1962. doi 10.3390/genes14101962
3. Svishcheva G.R., Tiys E.S., Elgaeva E.E. et al. A novel framework for analysis of the shared genetic background of correlated traits. *Genes (Basel).* 2022;13:1694. doi 10.3390/genes13101694