doi 10.18699/bgrs2024-4.3-13

Uncovering genetic factors of alcohol use disorder through GWAS and polygenic risk scores in East Slavs

Trofimov M.[1, 2*, #], Kudryavskiy V.[1, #], Shaheen L.[2, 3], Kovalenko E.[2], Vergasova K.[2], Kamelin A.[2], Rubinova V.[2], Kharitonov D.[2], Kim A.[2], Plotnikov N.[2], Elmuratov A.[2], Gainetdinov R.[4], Krupitsky E.[5, 6], Kibitov A.[5, 6], Rakitko A.[2, 7]

[1] *Skolkovo Institute of Science and Technology, Moscow, Russia*

[2] *Genotek Ltd., Moscow, Russia*

[3] *Moscow Institute of Physics and Technology, Dolgoprudny, Moscow region, Russia*

[4] *Institute of Translational Biomedicine, Saint Petersburg State University, St. Petersburg, Russia*

[5] *Saint Petersburg First Pavlov State Medical University, St. Petersburg, Russia*

[6] *V.M. Bekhterev National Medical Research Center for Psychiatry and Neurology, St. Petersburg, Russia*

[7] *Higher School of Economics, Faculty of Computer Science, Moscow, Russia*

* *m.trofimov@genotek.ru*

[#] *These authors contributed equally*

Key words: alcohol use disorder; polygenic risk score (PRS); East Slavic population; genome-wide association study (GWAS); genetic correlations; logistic regression

Motivation and Aim: Alcohol use disorder (AUD) significantly impacts over 200 health conditions, causing approximately 3 million annual deaths globally. Twin and adoption studies estimate that AUD is approximately 50 % heritable [1]. Multiple GWAS of alcohol-related traits, such as AUD [2] and alcohol dependence [3], have identified genetic associations primarily highlighting variants in genes directly related to alcohol metabolism (e. g., ADH1B and ALDH2). Our aims were to perform a GWAS of alcohol use disorder in the East Slavic population and develop a polygenic risk score (PRS). Additionally, we aimed to validate our findings in patients clinically diagnosed with AUD and in different ethnic groups represented in Russia.

Methods and Algorithms: We analyzed the genetic data of 41,575 individuals from Genotek Ltd. DNA was extracted and genotyped using Illumina Infinium Global Screening Array v.1-v.3 microarrays (650,000 SNPs). The Alcohol Use Disorder Identification Test (AUDIT), a ten-item questionnaire, was used to estimate alcohol drinking behavior and dependence. GWAS analyses were performed using PLINK v2.0 with the AUDIT total score as the dependent variable, accounting for age, sex, body mass index, and the first 20 principal components from genotype data as covariates. To calculate heritability and genetic correlations, we used LDSC tool v1.0.1 and summary statistics from UKB. PRS were generated using external GWAS summary statistics from P. Barr et al. [4], utilizing LDpred2 with default parameters. For comparison, a logistic regression model was built, classifying participants into ‘healthy’ and ‘case’ categories. Independent genotype data of patients diagnosed with alcohol dependence (AD) were acquired for PRS validation [5].

Results: From Genotek’s cohort, 41,575 individuals completed the AUDIT questionnaire. After filtering for the East Slavic population and removing close relatives, the sample size was 41,575 participants. In GWAS we detected one genome-wide significant SNP rs1229984 (odds ratio (OR) = 0.75, 95 % CI = 0.6–0.9) and p-value = 2.5×10^{-18} associated with AUD, located in the ADH1B gene. This SNP showed

significant associations in several ethnic groups, including Russians, Ukrainians, and Tatars, with consistent positive effect sizes. The estimated genetic heritability calculated from summary statistics of our GWAS was 6.4 %. Genetic correlations were assessed between AUD and 36 phenotypes, showing significant correlations with alcohol-related phenotypes, smoking status, and psychiatric disorders. PRS analysis using different AUDIT-T thresholds showed that a threshold of more than 14 performed slightly better, though the difference was not statistically significant. The OR for AUD were similar for both thresholds (Fig. 1). Our PRS demonstrated comparable performance to the external PRS, and it was validated in the independent clinical cohort of people diagnosed with alcohol dependence (AUC = 0.6, 95 % CI = 0.56–0.64). In addition, the PRS increased the predictive power of the multivariate model with nongenetic predictors by 1.33 % to AUC = 0.762 (p-value = 8×10^{-5}).

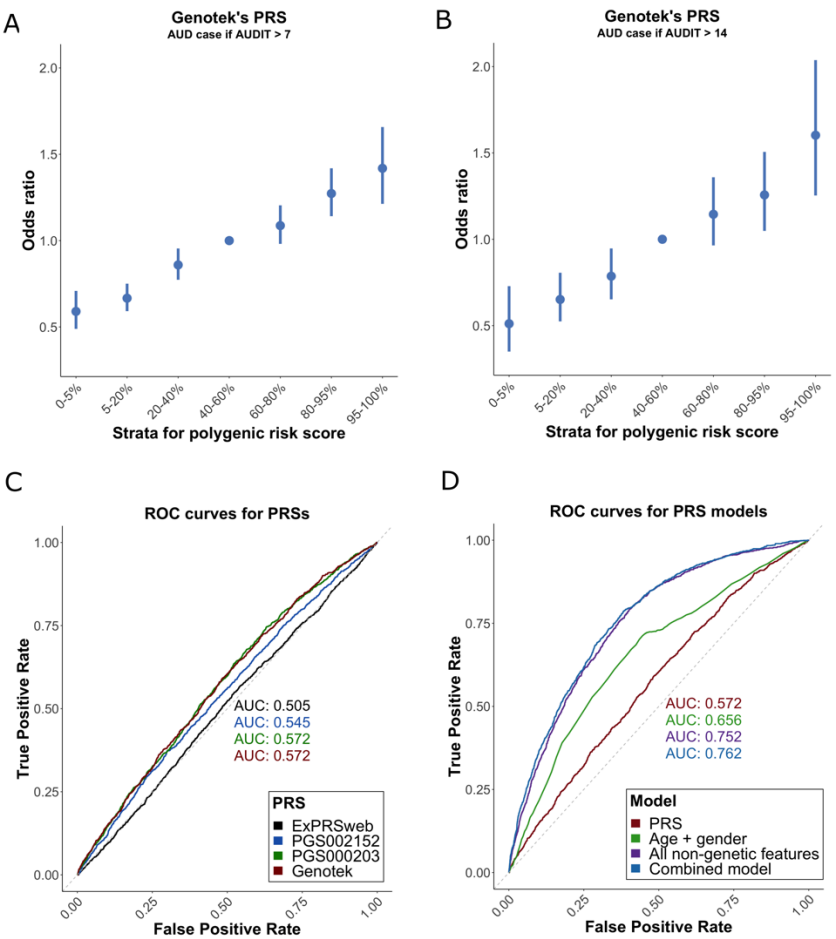

Fig. 1. PRS results. **A, B**: quantile plots of alcohol use disorder PRS built by LDpred-2. Based on PRS Individuals can be stratified into quantiles according to their risk score. It allows identifying groups of people with increased risk. Odds ratio represents comparison of PRS odds from different quantiles to the reference quantile (40–60 %). Bars represent confidence intervals. A, Individuals with AUDIT-T greater than 7 were defined as 'case'. B, Individuals with AUDIT-T greater than 14 were defined as 'case'. **C**, ROC curves plot for external PRSs and Genotek's PRS. **D**, ROC curves plot for models utilizing genetic and non-genetic features

Conclusion: We identified a well-known SNP in the ADH1B gene associated with AUD in the East Slavic population and found consistent associations across several ethnic groups. Our genetic correlation analysis confirmed expected relationships with alcohol- and smoking-related phenotypes. The performance of our PRS indicates that it can be applied across different European populations. The PRS was validated in an independent clinically established cohort, indicating that AUDIT-T is applicable for generating PRS to predict genetic predisposition to AUD. Further research may develop better scores for describing alcohol use disorder more comprehensively.

References

1. Hildebrand Karlén M. et al. Prevalence and heritability of alcohol use disorders in 18-year old Swedish twins. *Nordisk Alkohol Nark.* 2023;40:391-405
2. Kranzler H.R. et al. Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. *Nat Commun.* 2019;10:1499
3. Walters R.K. et al. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat Neurosci.* 2018;21:1656-1669
4. Barr P.B. et al. Using polygenic scores for identifying individuals at increased risk of substance use disorders in clinical and population samples. *Transl Psychiatry.* 2020;10:196
5. Levchenko A. et al. A Genome-Wide Association Study Reveals a BDNF-Centered Molecular Network Associated with Alcohol Dependence and Related Clinical Measures. *Biomedicines.* 2022;10:3007