Multivariate GWAS elucidates the genetic architecture of alcohol consumption and misuse, corrects biases, and reveals novel associations with disease

T.T. Mallard, J.E. Savage, E.C. Johnson, Y. Huang, A.C. Edwards, J.J. Hottenga, A.D. Grotzinger, D.E. Gustavson, M.V. Jennings, A. Anokhin, D.M. Dick, H.J. Edenberg, J.R. Kramer, D. Lai, J.L. Meyers, A.K. Pandey, K.P. Harden, M.G. Nivard, E.JC. de Geus, D.I. Boomsma, A. Agrawal, L.K. Davis, T.K. Clarke, A.A. Palmer, and S. Sanchez-Roige

Genome-wide association studies (GWASs) of the Alcohol Use Disorder Identification Test (AUDIT), a ten-item screener for alcohol use disorder (AUD), have elucidated novel loci for alcohol consumption and misuse. However, these studies also revealed that GWASs of alcohol consumption are influenced by numerous biases, exemplified by the paradoxically low genetic correlation between alcohol consumption and alcohol dependence, as well as the negative associations with other psychiatric disorders and medical conditions.

Methods: To address these persistent issues, we conducted the first item-level and largest GWAS of AUDIT (N=160,824), and applied a multivariate framework to mitigate previous biases.

Results: We identified novel patterns of similarity (and dissimilarity) among the AUDIT items, and found evidence of a correlated two-factor structure at the genetic level (Consumption and Problems factors, rg=.80). Moreover, by applying empirically-derived weights to each of the AUDIT items, we constructed an aggregate measure of alcohol consumption that is strongly associated with alcohol dependence (rg=.67) and several other psychiatric disorders, and no longer positively associated with health and socioeconomic factors. Lastly, by performing polygenic analyses in four independent cohorts that differed in their ascertainment and AUD risk, we identified novel genetic associations between alcohol consumption, alcohol misuse, and human health.

Discussion: Our work further emphasizes the value of AUDIT for both clinical and genetic studies of AUD, and the importance of using multivariate methods to study genetic associations that are more closely related to AUD.

This research was conducted using the UK Biobank Resource under Application Numbers 11425 and 16406. The AUDIT data collection in ALSPAC was funded by NIH AA018333; ACE was supported by a NIAAA grant (R01AA027522). The Collaborative Study on the Genetics of Alcoholism (COGA) is supported by NIH Grant U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute on Drug Abuse (NIDA). ECJ was supported by funding from NIAAA (F32AA027435); AA was supported by funding from K02 DA32573, MH109532, U10AA008401 grants; HJE and JRK were supported by NIAAA U10AA008401. For the Netherland Twin Register, funding was obtained from the Netherlands Organization for Scientific Research (NWO) and The Netherlands Organisation for Health Research and Development (ZonMW) grants 904-61-090, 985-10-002, 912-10-020, 904-61-193,480-04-004, 463-06-001, 451-04-034, 400-05-717, Addiction-31160008, 016-115-035, 481-08-011, 400-07-080, 056-32-010, Middelgroot-911-09-032, OCW_NWO Gravity program – 024.001.003, NWO-Groot 480-15-001/674, Center for Medical Systems Biology (CSMB, NWO

Genomics), NBIC/BioAssist/RK(2008.024), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI –NL, 184.021.007 and 184.033.111), X-Omics 184-034-019; Spinozapremie (NWO- 56-464-14192), KNAW Academy Professor Award (PAH/6635) and University Research Fellow grant (URF) to DIB; Amsterdam Public Health research institute (former EMGO+), Neuroscience Amsterdam research institute (former NCA); the European Community's Fifth and Seventh Framework Program (FP5-LIFE QUALITY-CT-2002-2006, FP7-HEALTH-F4-2007-2013, grant 01254: GenomEUtwin, grant 01413: ENGAGE and grant 602768: ACTION); the European Research Council (ERC Starting 284167, ERC Consolidator 771057, ERC Advanced 230374), Rutgers University Cell and DNA Repository (NIMH U24 MH068457-06), the National Institutes of Health (NIH, R01D0042157-01A1, R01MH58799-03, MH081802, DA018673, R01 DK092127-04, Grant Opportunity grants 1RC2 MH089951, and 1RC2 MH089995); the Avera Institute for Human Genetics, Sioux Falls, South Dakota (USA). Part of the genotyping and analyses were funded by the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health. Computing was supported by NWO through grant 2018/EW/00408559, BiG Grid, the Dutch e-Science Grid and SURFSARA. MGN is supported by R01MH120219, ZonMW grants 849200011 and 531003014 from The Netherlands Organisation for Health Research and Development, a VENI grant awarded by NWO (VI.Veni.191G.030) and is a Jacobs Foundation Fellow. PRS analyses using UK Biobank data were carried out on the Genetic Cluster Computer hosted by the Dutch National computing and Networking Services SurfSARA. DP was supported by NWO VICI 453-14-005. The dataset(s) used for the PheWAS analyses were obtained from Vanderbilt University Medical Center's BioVU which is supported by numerous sources: institutional funding, private agencies, and federal grants. These include the NIH funded Shared Instrumentation Grant S10RR025141; and CTSA grants UL1TR002243, UL1TR000445, and UL1RR024975. Genomic data are also supported by investigator-led projects that include U01HG004798, R01NS032830, RC2GM092618, P50GM115305, U01HG006378, U19HL065962, R01HD074711; and additional funding sources listed at https://victr.vumc.org/biovu-funding/. LKD obtained support from 1R01MH113362, 1R01MH118233 and 1R56MH120736. SSR was supported by a NARSAD Young Investigator Award from the Brain and Behavior Foundation (Grant Number 27676). YH, MVJ, SSR and AAP were supported by funds from the California Tobacco-Related Disease Research Program (TRDRP; Grant Number 28IR-0070 and T29KT0526).