

Novel alcohol-related genes suggest shared genetic mechanisms with neuropsychiatric disorders

Evangelos Evangelou^{1,2‡}, He Gao^{1,3‡}, Congying Chu^{4‡}, Georgios Ntritsos^{2‡}, Paul Blakeley^{1,5}, Andrew R Butts⁶, Raha Pazoki¹, Hideaki Suzuki^{7,8}, Fotios Koskeridis², Andrianos M Yiorkas^{9,10}, Ibrahim Karaman^{1,11}, Joshua Elliott¹, Qiang Luo^{12,13}, Stefanie Aeschbacher¹⁴, Traci M Bartz^{15,16}, Sebastian E Baumeister^{17,18}, Peter S Braund^{19,20}, Michael R Brown²¹, Jennifer A Brody¹⁵, Toni-Kim Clarke²², Niki Dimou², Jessica D Faul²³, Georg Homuth²⁴, Anne U Jackson²⁵, Katherine A Kentistou^{26,27}, Peter K Joshi²⁶, Rozenn N Lemaitre¹⁵, Penelope A Lind²⁸, Leo-Pekka Lyytikäinen²⁹⁻³¹, Massimo Mangino^{32,33}, Yuri Milaneschi³⁴, Christopher P Nelson^{19,20}, Ilja M Nolte³⁵, Mia-Maria Perälä^{36,37}, Ozren Polasek³⁸, David Porteous^{39,40}, Scott M Ratliff⁴¹, Jennifer A Smith^{23,41}, Alena Stančáková⁴², Alexander Teumer^{17,43}, Samuli Tuominen⁴⁴, Sébastien Thériault^{45,46}, Jagadish Vangipurapu⁴², John B Whitfield⁴⁷, Alexis Wood⁴⁸, Jie Yao⁴⁹, Bing Yu²¹, Wei Zhao⁴¹, Dan E Arking⁵⁰, Juha Auvinen^{51,52}, Chunyu Liu⁵³, Minna Männikkö⁵⁴, Lorenz Risch⁵⁵⁻⁵⁷, Jerome I Rotter⁵⁸, Harold Snieder³⁵, Juha Veijola⁵⁹⁻⁶¹, Alexandra I Blakemore^{9,10}, Michael Boehnke²⁵, Harry Campbell²⁶, David Conen⁴⁵, Johan G Eriksson⁶²⁻⁶⁴, Hans J Grabe^{65,66}, Xiuqing Guo⁴⁹, Pim van der Harst⁶⁷⁻⁶⁹, Catharina A Hartman⁷⁰, Caroline Hayward⁷¹, Andrew C Heath⁷², Marjo-Riitta Jarvelin⁷³⁻⁷⁷, Mika Kähönen^{78,79}, Sharon LR Kardina⁴¹, Michael Kühne¹⁴, Johanna Kuusisto⁸⁰, Markku Laakso⁸⁰, Jari Lahti⁴⁴, Terho Lehtimäki^{29,30}, Andrew M McIntosh^{22,40}, Karen L Mohlke⁸¹, Alanna C Morrison²¹, Nicholas G Martin⁴⁷, Albertine J Oldehinkel⁷⁰, Brenda WJH Penninx³⁴, Bruce M Psaty^{82,83}, Olli T Raitakari^{84,85}, Igor Rudan²⁶, Nilesh J Samani^{19,20}, Laura J Scott²⁵, Tim D Spector³², Niek Verweij⁶⁷, David R Weir²³, James F Wilson^{26,71}, Daniel Levy^{86,87}, Ioanna Tzoulaki¹⁻³, Jimmy D Bell^{88‡}, Paul M. Matthews^{7,11‡}, Adrian Rothenfluh^{6,89‡}, Sylvane Desrivières^{4‡}, Gunter Schumann^{4*‡}, Paul Elliott^{1,3,11,90,91*‡}

‡ Equal contribution

* Corresponding authors

1. Department of Epidemiology and Biostatistics, Imperial College London, London, UK.
2. Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece.
3. MRC-PHE Centre for Environment and Health, Imperial College London, London, UK.
4. Medical Research Council-Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London SE5 8AF, United Kingdom.
5. NIHR Imperial Biomedical Research Centre, ITMAT Data Science Group, Imperial College London, UK.
6. Molecular Medicine; School of Medicine; University of Utah, Salt Lake City, UT.

7. Centre for Restorative Neurosciences, Division of Brain Sciences, Department of Medicine, Hammersmith Campus, Imperial College London, Du Cane Road, London, W12 0NN, UK.
8. Tohoku Medical Megabank Organization, Tohoku University, 2-1 Seiryō, Aoba, Sendai, 980-8573, Japan.
9. Department of Life Sciences, Brunel University London, London, UK.
10. Section of Investigative Medicine, Imperial College London, London, UK.
11. UK Dementia Research Institute, Imperial College London, London, United Kingdom.
12. Institute of Science and Technology for Brain-Inspired Intelligence, MOE-Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence, Fudan University, Shanghai 200433, PR China.
13. Department of Psychology and the Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge CB2 3EB, United Kingdom.
14. Cardiology Division, University Hospital Basel, 4031 Basel, Switzerland.
15. Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, Washington, USA.
16. Department of Biostatistics, University of Washington, Seattle, WA, USA.
17. Institute for Community Medicine, University Medicine Greifswald, 17475 Greifswald, Germany.
18. Chair of Epidemiology, Ludwig-Maximilians-Universität München, UNIKA-T Augsburg, 86156 Augsburg, Germany.
19. Department of Cardiovascular Sciences, University of Leicester, Cardiovascular Research Centre, Glenfield Hospital, Leicester, LE3 9QP, UK.
20. NIHR Leicester Biomedical Research Centre, Glenfield Hospital, Groby Road, Leicester, LE3 9QP, UK.
21. Human Genetics Center, Department of Epidemiology, Human Genetics & Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX 77030 USA.
22. Department of Psychiatry, University of Edinburgh, Edinburgh, UK, EH10 5HF.
23. Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, USA.
24. Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, 17475 Greifswald, Germany.
25. Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan, USA.
26. Centre for Global Health Research, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, EH8 9AG, UK.
27. Centre for Cardiovascular Sciences, Queens Medical Research Institute, University of Edinburgh, Edinburgh, EH16 4TJ, Scotland.
28. Psychiatric Genetics, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia.

29. Department of Clinical Chemistry, Fimlab Laboratories, Tampere 33520, Finland.
30. Department of Clinical Chemistry, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Life Sciences, University of Tampere, Tampere 33014, Finland.
31. Department of Cardiology, Heart Center, Tampere University Hospital, Tampere 33521, Finland.
32. Department of Twin Research and Genetic Epidemiology, Kings College London, London SE1 7EH, UK.
33. NIHR Biomedical Research Centre at Guys and St Thomas Foundation Trust, London SE1 9RT, UK.
34. Department of Psychiatry, Amsterdam Neuroscience and Amsterdam Public Health research institute, Amsterdam University Medical Center, 1081 HJ Amsterdam, the Netherlands.
35. Department of Epidemiology, University Medical Center Groningen, University of Groningen, The Netherlands.
36. Folkhälsan Research Center, 00290 Helsinki, Finland.
37. Department of Public Health Solutions, National Institute for Health and Welfare, 00271 Helsinki, Finland.
38. Faculty of Medicine, University of Split, Split, 21000, Croatia.
39. Generation Scotland, Medical Genetics Section, Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK, EH4 2XU.
40. Centre for Cognitive Ageing and Cognitive Epidemiology, Edinburgh, UK, EH8 9JZ.
41. Department of Epidemiology, University of Michigan, Ann Arbor, MI, USA.
42. Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland, 70210 Kuopio, Finland.
43. DZHK (German Centre for Cardiovascular Research), partner site Greifswald, 17475 Greifswald, Germany.
44. Department of Psychology and Logopedics, Faculty of Medicine, 00014 University of Helsinki, Helsinki, Finland.
45. Population Health Research Institute, McMaster University, L8L 2X2 Hamilton, Canada.
46. Department of Molecular Biology, Medical Biochemistry and Pathology, Laval University, G1V 0A6 Quebec City, Canada.
47. Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia.
48. Department of Pediatrics/Nutrition, Baylor College of Medicine, 1 Baylor Plaza, Houston, Texas 77030 USA.
49. The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, California 90502, USA.
50. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

51. Center for Life Course Health Research, Faculty of Medicine, University of Oulu, 90220, Oulu, Finland.
52. Oulunkaari Health Center, 91100 Ii, Finland.
53. Department of Biostatistics, Boston University School of Public Health.
54. Northern Finland Birth Cohorts, Faculty of Medicine, University of Oulu, 90220, Oulu, Finland.
55. Institute of Clinical Chemistry, Inselspital Bern, University Hospital, University of Bern, 3010 Bern, Switzerland.
56. Labormedizinisches Zentrum Dr. Risch, Vaduz, Liechtenstein.
57. Private University of the Principality of Liechtenstein, Triesen, Liechtenstein.
58. The Institute for Translational Genomics and Population Sciences, Departments of Pediatrics and Medicine, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, California 90502, USA.
59. Department of Psychiatry, Research Unit of Clinical Neuroscience, 90014 University of Oulu, Finland.
60. Department of Psychiatry, University Hospital of Oulu, 90220 Oulu, Finland.
61. Medical research Center Oulu, University and University Hospital of Oulu, 90220 Oulu, Finland.
62. Department of General Practice and Primary health Care, Tukholmankatu 8 B , 00014 University of Helsinki, Finland.
63. National Institute for Health and Welfare, 00271 Helsinki, Finland.
64. Helsinki University Central Hospital, Unit of General Practice, 00290 Helsinki, Finland.
65. Department of Psychiatry and Psychotherapy, University Medicine Greifswald, 17475 Greifswald, Germany.
66. German Center for Neurodegenerative Diseases (DZNE), Rostock/Greifswald, 17475 Greifswald, Germany.
67. University Medical Center Groningen, University of Groningen, Department of Cardiology, the Netherlands.
68. University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, The Netherlands.
69. Durrer Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, Utrecht, The Netherlands.
70. Department of Psychiatry, University Medical Center Groningen, University of Groningen, 9713 GZ Groningen, The Netherlands.
71. MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh, EH4 2XU, UK.
72. Department of Psychiatry, School of Medicine, Washington University in St. Louis, St. Louis, MO 63110, USA.

73. Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, W2 1PG, United Kingdom.
74. Center for Life Course Health Research, Faculty of Medicine, University of Oulu, PO Box 8000, FI-90014 Oulun yliopisto, Finland.
75. Biocenter Oulu, University of Oulu, Aapistie 5, 90220 Oulu, Finland.
76. Unit of Primary Health Care, Oulu University Hospital, OYS, Kajaanintie 50, 90220 Oulu, Finland.
77. Department of Life Sciences, College of Health and Life Sciences, Brunel University London, Kingston Lane, Uxbridge, Middlesex UB8 3PH, United Kingdom.
78. Department of Clinical Physiology, Tampere University Hospital, Tampere 33521, Finland.
79. Department of Clinical Physiology, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Life Sciences, University of Tampere, Tampere 33014, Finland.
80. Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland, and Kuopio University Hospital, 70210 Kuopio, Finland.
81. Department of Genetics, University of North Carolina, Chapel Hill, NC, USA.
82. Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology, and Health Services, University of Washington, Seattle, WA, 98101.
83. Kaiser Permanente Washington Health Research Institute, Seattle, WA, 98101.
84. Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku 20521, Finland.
85. Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku 20014, Finland.
86. Framingham Heart Study, Framingham, MA, USA.
87. Population Sciences Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD.
88. Research Centre for Optimal Health, Department of Life Sciences, University of Westminster, London, W1W 6UW.
89. Departments of Psychiatry, Neurobiology & Anatomy, Human Genetics; School of Medicine; University of Utah, Salt Lake City, UT.
90. National Institute for Health Research Imperial Biomedical Research Centre, Imperial College Healthcare NHS Trust and Imperial College London, London, UK.
91. Health Data Research-UK London substantive site, London, UK.

Corresponding authors: Gunter Schumann (gunter.schumann@kcl.ac.uk) and Paul Elliott (p.elliott@imperial.ac.uk)

ABSTRACT

Excessive alcohol consumption is one of the main causes of death and disability worldwide. Alcohol consumption is a heritable complex trait. We conducted a meta-analysis of genome-wide association studies (GWAS) of gram/day (g/d) alcohol consumption in UK-Biobank, AlcGen and CHARGE+ consortia accumulating 480,842 people of European descent to decipher the genetic architecture of alcohol intake. We identified 46 novel, common loci, and investigated their potential functional significance using magnetic resonance imaging data and gene expression studies. Our results identify genetic pathways associated with alcohol consumption and suggest shared genetic mechanisms with neuropsychiatric disorders including schizophrenia.

1 Excessive alcohol consumption is a major public health problem that is responsible
2 for 2.2% and 6.8% age-standardized deaths for women and men respectively¹. Most
3 genetic studies of alcohol use focus on alcohol dependency, although the population
4 burden of alcohol-related disease mainly reflects a broader range of alcohol
5 consumption behaviors². Small reductions in alcohol consumption could have major
6 public health benefits; even moderate amounts of alcohol/day may have significant
7 impact on mortality³.

8 Alcohol consumption is a heritable complex trait⁴, but genetic studies to date have
9 robustly identified only a small number of associated genetic variants⁵⁻⁸. These
10 include variants in the aldehyde dehydrogenase (ADH) gene family, a group of
11 enzymes that catalyze the oxidation of aldehydes⁹, including a cluster of genes on
12 chromosome 4q23 (*ADH1B*, *ADH1C*, *ADH5*, *ADH6*, *ADH7*)⁶.

13 Here, we report a GWAS meta-analysis of alcohol intake (log transformed g/day)
14 among people of European ancestry drawn from UK Biobank (UKB)¹⁰, the Alcohol
15 Genome-Wide Consortium (AlcGen) and the Cohorts for Heart and Aging Research in
16 Genomic Epidemiology Plus (CHARGE+) consortia. Briefly, UKB is a prospective
17 cohort study of ~500,000 individuals recruited between the ages of 40 and 69 years.
18 Participants were asked to report their average weekly and monthly alcohol
19 consumption through a self-completed touchscreen questionnaire¹⁰. Based on these
20 reports, we calculated the g/d alcohol intake (**Methods**). Participants were
21 genotyped using a customized array with imputation from the Haplotype Reference
22 Consortium (HRC) panel¹¹, yielding ~7 million common single nucleotide
23 polymorphisms (SNPs) with minor allele frequency (MAF) $\geq 1\%$ and imputation
24 quality score [INFO] ≥ 0.1 . After quality control (QC) and exclusions (**Methods**) we
25 performed GWAS of alcohol consumption using data from 404,731 UKB participants
26 of European descent under an additive genetic model (**Methods and Supplementary**
27 **Table 1**). We found that genomic inflation in the UKB analysis was $\lambda_{GC}=1.45$, but did
28 not adjust for inflation as the LD score regression intercept was 1.05, indicating that
29 this was due to polygenicity rather than to population stratification¹². The estimated
30 SNP-wide heritability of alcohol consumption in the UKB data was 0.09.

31 We also carried out GWAS in 25 independent studies from the AlcGen and CHARGE+
32 consortia including 76,111 participants of European descent for which alcohol g/d
33 could be calculated (**Supplementary Table 2**). Various arrays were used for
34 genotyping, with imputations performed using either the 1,000 Genomes Reference
35 Panel or the HRC platforms (**Supplementary Table 3**). After QC, we applied genomic

36 control at the individual study level and obtained summary results for ~7 million
37 SNPs with imputation quality score ≥ 0.3 (**Methods**).

38 We combined the UKB, AlcGen and CHARGE+ results using a fixed effects inverse
39 variance weighted approach for a total of 480,842 individuals¹³. To maximize power,
40 we performed a single-stage analysis to test common SNPs with MAF $\geq 1\%$. We set a
41 stringent P -value threshold of $P < 5 \times 10^{-9}$ to denote significance in the combined
42 meta-analysis¹⁴, and required signals to be at $P < 5 \times 10^{-7}$ in UKB, with same direction
43 of effect in UKB and AlcGen plus CHARGE+, to minimize false positive findings. We
44 excluded SNPs within 500kb of variants reported as genome-wide significant in
45 previous GWAS of alcohol consumption^{5,6}, identified novel loci by requiring SNPs to
46 be independent of each other (LD $r^2 < 0.1$), and selected the sentinel SNP within each
47 locus according to lowest P -value (**Methods**).

48 We then tested for correlations of alcohol-associated SNPs with Magnetic Resonance
49 Imaging (MRI) phenotypes of brain, heart and liver, and gene expression. We tested
50 the sentinel SNPs for association with other traits/diseases and *Drosophila* mutant
51 models were used to investigate functional effects on ethanol-induced behavior.

52 **RESULTS**

53 Our meta-analysis identified 46 novel loci associated with alcohol consumption (log
54 transformed g/day) (**Fig. 1 and Table 1**). All inferential statistics for the novel loci are
55 reported in Table 1 whereas heterogeneity metrics are presented in **Supplementary**
56 **Table 4**. In addition, we discovered a further eight variants in the combined analysis
57 at nominal genome-wide significance ($P < 1 \times 10^{-8}$) that may also be associated with
58 alcohol intake (**Supplementary Table 5**). The most significantly associated variant,
59 rs1991556 ($P = 4.5 \times 10^{-23}$), is an intronic variant in *MAPT* gene that encodes the
60 microtubule-associated protein tau, and was found through Phenoscanner not only
61 to be associated with dementia¹⁵ and Parkinson's disease^{16,17}, but also with
62 neuroticism, schizophrenia¹⁸ and other traits¹⁹⁻²¹ (**Methods, Fig. 2 and**
63 **Supplementary Table 6**). The second most significantly associated variant is
64 rs1004787 ($P = 6.7 \times 10^{-17}$), near *SIX3* gene, which encodes a member of the sine
65 oculis homeobox transcription factor family involved in eye development²². The third
66 SNP is rs13107325 ($P = 1.3 \times 10^{-15}$), a missense SNP in *SLC39A8*
67 (<https://www.ncbi.nlm.nih.gov/gene/64116>), a gene that encodes a member of the
68 SLC39 family of metal ion transporters, which has been associated with

69 schizophrenia²³ as well as inflammatory bowel disease, cardiovascular and metabolic
70 phenotypes^{24 25-27} in previous GWAS (**Fig. 2 and Supplementary Table 6**).

71 Another of our most significant variants, an intronic SNP rs7121986 ($P = 6.2 \times 10^{-14}$)
72 in *DRD2* (<https://www.ncbi.nlm.nih.gov/gene/1813>), encodes the dopamine
73 receptor D2 that has been associated with cocaine addiction, neuroticism and
74 schizophrenia¹⁸. We also found significant associations with SNP rs988748 ($P = 4.4 \times$
75 10^{-9}) in the *BDNF* gene (<https://www.ncbi.nlm.nih.gov/gene/627>, that encodes a
76 member of the nerve growth factor family of proteins and rs7517344, which is near
77 *ELAVL4* (<https://www.ncbi.nlm.nih.gov/gene/1996>) ($P = 2.0 \times 10^{-10}$), the gene
78 product of which is involved in BDNF regulation²⁸. Previous studies have suggested
79 that a variant in *BDNF* is associated with alcohol consumption and that alcohol
80 consumption modulates BDNF expression²⁹.

81
82 Additionally, we found association of alcohol consumption with SNP rs838145 ($P =$
83 3.2×10^{-15}), which has been associated with macronutrient intake in a previous
84 GWAS³⁰. This variant is nearest *IZUMO* (<https://www.ncbi.nlm.nih.gov/gene/284359>)
85 in a locus of around 50kb that spans a number of genes including *FGF21*
86 (<https://www.ncbi.nlm.nih.gov/gene/26291>), whose gene product FGF21 is a liver
87 hormone involved in the regulation of alcohol preference, glucose and lipid
88 metabolism³¹. We previously reported significant association of alcohol intake with
89 SNP rs11940694 in *KLB* (<https://www.ncbi.nlm.nih.gov/gene/152831>), an obligate
90 receptor of FGF21 in the brain⁵, and we strongly replicated that finding here ($P = 3.3$
91 $\times 10^{-68}$).

92
93 As well as variants in *KLB* and in the alcohol dehydrogenase locus (smallest $P = 1.2 \times$
94 10^{-125}), we found support ($P = 1 \times 10^{-5}$) for association of common variants in the
95 three other alcohol intake-related loci previously reported in GWAS (**Supplementary**
96 **Table 7**), including SNP rs6943555 in *AUTS2*
97 (<https://www.ncbi.nlm.nih.gov/gene/26053>) ($P = 2.9 \times 10^{-6}$). In addition, we found a
98 novel alcohol intake-related SNP rs1421085 in *FTO*
99 (<https://www.ncbi.nlm.nih.gov/gene/79068>) in high LD ($r^2 = 0.92$) with a variant
100 reported previously as genome-wide significant for association with alcohol
101 dependence³².

102
103 Conditional analysis using Genome-wide Complex Trait Analysis (GCTA) did not
104 reveal any independent secondary signals related to alcohol consumption. Among
105 ~14,000 individuals in the independent Airwave cohort³³ (**Methods**), 7% of the
106 variance in alcohol consumption was explained by the novel and known common

107 variants. Using weights from our analysis, we constructed an unbiased weighted
108 genetic risk score (GRS) in Airwave (**Methods**) and found a strong association of the
109 novel and known variants on alcohol consumption levels ($P = 2.75 \times 10^{-14}$), with mean
110 difference in sex-adjusted alcohol intake of 2.6 g/d comparing the top vs the bottom
111 quintile of the GRS (**Supplementary Table 8**).
112

113 **Associations with MRI imaging phenotypes**

114 We functionally characterized novel variants by carrying out single-SNP analyses of
115 the imaging phenotypes in UKB (**Methods**), focusing on brain (N=9,702), heart
116 (N=10,706) and liver (N=8,479).

117 With Bonferroni correction (corrected P -value 6.6×10^{-6} , corresponding to 0.05/46
118 SNPs*164 imaging phenotypes), we found significant positive associations between
119 SNP rs13107325 in *SLC39A8* and the volumes of multiple brain regions; All inferential
120 statistics for these associations are reported in **Supplementary Table 9**. The
121 strongest associations were with putamen (left: $P = 2.5 \times 10^{-45}$, right: $P = 2.8 \times 10^{-47}$),
122 ventral striatum (left: $P = 9.5 \times 10^{-53}$, right: $P = 9.6 \times 10^{-51}$) and cerebellum (strongest
123 association for left I-IV volume; $P = 1.2 \times 10^{-9}$) (**Supplementary Table 9**); similar
124 findings were recently reported in a GWAS on brain imaging in UKB³⁴. The other
125 significant association was for rs1991556 with the parahippocampal gyrus ($P = 1.2 \times$
126 10^{-6}).

127 We then tested these brain regions for association with alcohol consumption and
128 found a significant effect for the left ($t_{8601} = -3.7$; $\beta \pm SE = -0.0019 \pm 0.0005$; $P =$
129 2.0×10^{-4}) and right ($t_{8601} = -3.65$; $\beta \pm SE = -0.0070 \pm 0.0005$; $P = 2.6 \times 10^{-4}$)
130 putamen. Finally, we used data from N= 8,610 individuals and performed a
131 mediation analysis using a standard three-variable path model, bootstrapping 10,000
132 times to calculate the significance of the mediation effect of putamen volume for
133 genetic influences on alcohol consumption (**Methods**). We found evidence that the
134 effect of SNP rs13107325 in *SLC39A8* on alcohol intake is partially mediated via its
135 association with left ($t_{8601} = -3.03$; $\beta \pm SE = -0.27 \pm 0.09$; $P = 1.9 \times 10^{-3}$) and right
136 ($t_{8601} = -2.82$; $\beta \pm SE = -0.27 \pm 0.09$; $P = 1.7 \times 10^{-3}$) putamen volume (**Fig. 3 and**
137 **Supplementary Table 10**). To exclude the possibility of an inverse causal pathway we
138 performed additional analyses in UKB non-drinkers (N =589). With 10,000 random
139 permutations, associations of rs13107325 with both left and right putamen
140 remained significant (left putamen: $t_{541}=1.06$; $P = 0.02$; right putamen: $t_{541}=0.38$; $P =$

141 0.04) indicating that the association between rs13107325 and putamen regions is
142 not mediated by alcohol intake.

143 We did not find any significant associations of novel SNPs with either cardiac (left
144 ventricular mass or end diastolic volume or right ventricular end diastolic volume)
145 (**Supplementary Table 11**) or liver fat measures on MRI (**Supplementary Table 12**),
146 after adjustment for multiple testing.

147 **Effects of SNPs on gene expression**

148 We carried out expression quantitative trait loci eQTL analyses using the Genotype-
149 Tissue Expression (GTEx) and the UK Brain Expression Consortium (UKBEC) datasets;
150 34 of the 53 novel and known SNPs associated with alcohol consumption have a
151 significant effect on gene expression in at least one tissue, including 33 SNPs that
152 affect gene expression in the brain (**Supplementary Tables 13 and 14, and**
153 **Supplementary Figures 1-3**). We found that the most significant eQTLs often do not
154 involve the nearest gene and that several of the SNPs affect expression of different
155 genes in different tissues. For example, SNP rs1991556 in the *MAPT* gene
156 (<https://www.ncbi.nlm.nih.gov/gene/4137>) affects expression of 33 genes overall,
157 with most significant effects on the expression of the non-protein coding genes
158 *CRHR1-IT1* (also known as *C17orf69* or *LINC02210*)
159 (<https://www.ncbi.nlm.nih.gov/gene/147081>) and *LRRC37A4P*
160 (<https://www.ncbi.nlm.nih.gov/gene/?term=LRRC37A4P>), near *MAPT*, across a wide
161 range of tissues including brain, adipose tissue and skin ($P = 7.2 \times 10^{-126}$ to $P = 2.5 \times$
162 10^{-6}) (**Supplementary Figure 2**). Similarly, the A-allele at SNP rs2071305 within
163 *MYBPC3* (<https://www.ncbi.nlm.nih.gov/gene/4607>) affects the expression of
164 several genes and is most significantly associated with increased expression of
165 *C1QTNF4* (<https://www.ncbi.nlm.nih.gov/gene/114900>) across several tissues ($P =$
166 1.9×10^{-25} to $P = 8.4 \times 10^{-5}$).

167 Several of these eQTLs were found to affect expression of genes known to be
168 involved in reward and addiction. SNP rs1053651 in the *TCAP-PNMT-STAR3* gene
169 cluster affects expression of the *PPP1R1B* gene (also known as *DARPP-32*)
170 (<https://www.ncbi.nlm.nih.gov/gene/84152>) which encodes a protein that mediates
171 the effects of dopamine in the mesolimbic reward pathway³⁵. Other known
172 addiction-related genes include
173 *ANKK1* (<https://www.ncbi.nlm.nih.gov/gene/255239>) and *DRD2* (expression affected
174 by SNP rs7121986) implicated in alcohol and nicotine dependence^{36,37}, *CRHR1*
175 (<https://www.ncbi.nlm.nih.gov/gene/1394>) (affected by SNP rs1991556) involved in

176 stress-mediated alcohol dependence^{38,39} and *PPM1G* (SNP rs1260326)
177 (<https://www.ncbi.nlm.nih.gov/gene/5496>) whose epigenetic modification was
178 reported to be associated with alcohol abuse⁴⁰.

179 Over-representation enrichment analyses based on functional annotations and
180 disease-related terms indicated that genes whose expressions are affected by the
181 identified eQTLs are most significantly enriched for terms related to abdominal
182 (n=91) and other malignant cancers, motor function (n= 5) and cellular homeostasis
183 (n= 22) (**Supplementary Figure 4**). We performed a gene-based analysis and
184 repeated the over-representation enrichment analysis adding the new set of
185 identified genes (**Supplementary Table 15**). The results were similar supporting an
186 enrichment for abdominal (n=100) and other cancers, as well as motor function
187 (n=5) and cellular homeostasis (n=24) (**Supplementary Figure 5**).

188 **Other traits and diseases**

189

190 Using LD score regression¹², we assessed genetic correlations between alcohol
191 consumption and 235 complex traits and diseases from publicly available summary
192 GWAS statistics (**Methods**). All results including their statistics (i.e. r_g , standard
193 errors, z value and P value) are included in **Supplementary Table 16**. The strongest
194 positive genetic correlations based on false discovery rate $P < 0.02$ were found for
195 smoking ($r_g = 0.42$, $P = 1.0 \times 10^{-23}$) and HDL cholesterol levels ($r_g = 0.26$, $P = 5.1 \times 10^{-13}$).
196 We also found negative correlations for sleep duration ($r_g = -0.14$, $P = 3.8 \times 10^{-7}$) and
197 fasting insulin levels ($r_g = -0.25$, $P = 4.5 \times 10^{-6}$). A significant genetic correlation was
198 also found with schizophrenia ($r_g = 0.07$, $P = 3.9 \times 10^{-3}$) and bipolar disorder ($r_g = 0.15$,
199 $P = 5.0 \times 10^{-4}$) (**Supplementary Table 16**). Over-representation enrichment analysis
200 using WebGestalt⁴¹ (<http://www.webgestalt.org>) showed that our list of novel and
201 known variants is significantly enriched for several diseases and traits including
202 developmental disorder in children ($P = 7.3 \times 10^{-5}$), epilepsy ($P = 1.4 \times 10^{-4}$), heroin
203 dependence ($P = 5.7 \times 10^{-4}$) and schizophrenia ($P = 8.4 \times 10^{-4}$) (**Supplementary Figure**
204 **6**). The result of the Mendelian randomization analysis (**Methods**) to assess a
205 potential causal effect of alcohol on schizophrenia risk, using the inverse variance
206 weighted approach, was not significant ($P = 0.089$), with large heterogeneity of the
207 estimates of the tested variants.

208 **Functional studies in *Drosophila***

209 Based on our GWAS and brain imaging findings we took forward SNP rs13107325 in
210 *SLC39A8* (alias *Zip8* gene) for additional testing in *Drosophila*, which employ

211 conserved mechanisms to modulate ethanol-induced behaviors^{42,43}. First, we
212 overexpressed human *Zip8* using a Gal4-driver that included expression in neurons
213 involved in multiple ethanol-induced behaviors⁴³. Flies carrying *ics^{Gal4}/+ UAS-*
214 *hZip8/+* showed a slight, but significant, resistance to ethanol-induced sedation
215 compared to control flies ($t_{30} = 2.3$; Hedge's $g = 0.80$; 95% CI: 0.08 – 1.53; $P = 0.026$;
216 $N = 16$ per genotype). Ethanol tolerance, induced with repeat exposures spaced by a
217 4-hour recovery, was unchanged in these flies ($t = 1.0$; $P = 0.33$; **Fig. 4a**). We next
218 used the same Gal4-driver to knock down the endogenous *Drosophila* ortholog
219 of *hZip8*, namely *dZip71B*. This caused the flies to display naïve sensitivity to ethanol-
220 induced sedation ($t_{14} = 3.98$; Hedge's $g = -1.84$; 95% CI: -0.67 – -3.01; $P = 0.0014$; $N =$
221 8 per genotype), and in addition, these flies developed greater tolerance to ethanol
222 upon repeat exposure ($t_{14} = 4.80$; Hedge's $g = 2.29$; 95% CI: 1.03 – 3.55; $P = 0.0003$;
223 **Fig. 4b**). To corroborate this phenotype, we then tested flies transheterozygous for
224 two independent transposon-insertions in the middle of the *dZip71B* gene
225 (**Supplementary Figure 7**) and found that these *dZip71B^{Mi/MB}* flies also displayed
226 naïve sensitivity ($t_{14} = 3.23$; Hedge's $g = -1.54$; 95% CI: -0.42 – -2.65; $P = 0.006$) and
227 increased ethanol-induced tolerance ($t_{14} = 2.39$; Hedge's $g = 1.13$; 95% CI: 0.07 -
228 2.18; $P = 0.032$) compared to controls ($N = 8$ each) (**Fig. 4c**).

229

230 DISCUSSION

231 Our discovery utilizing data on common variants from over 480,000 people of
232 European descent extends our knowledge of the genetic architecture of alcohol
233 intake, increasing the number of identified loci to 46. We found loci involved in
234 neuropsychiatric conditions such as schizophrenia, Parkinson's disease and
235 dementia, as well as *BDNF* where gene expression is affected by alcohol abuse. Our
236 findings illustrate that large-scale studies of genetic associations with alcohol intake
237 in the general population, rather than on alcohol dependency alone, can provide
238 additional insights into genetic mechanisms regulating alcohol consumption.

239 We highlight the role of the highly pleiotropic *MAPT* and *SLC39A8* genes in the
240 genetics of alcohol consumption. *MAPT* plays a key role in tau-associated dementia⁴⁴
241 and both genes are also implicated in other neuropsychiatric conditions including
242 neuroticism, schizophrenia and Parkinson's disease¹⁶⁻¹⁸. The *SLC39A8* gene encodes a
243 member of the SLC39 family of metal ion transporters. The encoded protein is
244 glycosylated and found in plasma membrane and mitochondria, and is involved in
245 the cellular transport of zinc, modulation of which could affect microglial
246 inflammatory responses⁴⁵. Our gain- and loss-of-function studies in *Drosophila*

247 indicate a potential causal role of *SLC39A8* in alcohol drinking behavior, even though
248 results should be interpreted with caution due to small sample size in our
249 experiment. The MRI brain imaging demonstrates a significant association of SNP
250 rs13107325 in the *SLC39A8* gene and putamen volume differences, and these
251 structural differences appear to partially mediate associations of rs13107325 with
252 alcohol consumption. The putamen has been associated with alcohol consumption
253 and the withdrawal syndrome after chronic administration to rodents and non-
254 human primates⁴⁶. Our mediation analysis is suggestive of a plausible causal pathway
255 linking rs13107325 in *SLC39A8* with alcohol intake via an effect on putamen volume,
256 but follow-up work is needed to conclusively demonstrate causal links. Putamen
257 volume differences have also been associated with both schizophrenia and
258 psychosis^{47,48} and robust association between SNP rs13107325 in *SLC39A8* and
259 schizophrenia was reported in a previous GWAS²³.

260 We also report SNP rs7121986 near *DRD2* as a novel alcohol intake variant in GWAS.
261 The gene product of *DRD2*, D2 dopamine receptor, is a G protein-coupled receptor
262 on post-synaptic dopaminergic neurons that has long been implicated in
263 alcoholism⁴⁹. In addition, we identify SNP rs988748 in *BDNF* as a novel alcohol intake
264 variant; BDNF expression is differentially affected by alcohol exposure in animal
265 models^{50,51}. Both genes (along with *PPP1R1P*) are centrally involved in reward-
266 mediating mesocortico-limbic pathways and both are implicated in the development
267 of schizophrenia. For example, there is a robust GWAS association between
268 schizophrenia and SNP rs4938021 in *DRD2* (in perfect LD with our novel alcohol
269 intake-related variant rs7121986) and *DRD2* appears to be pivotal in network
270 analyses of genes involved in schizophrenia⁵². Taken together, our results suggest
271 that there are shared genetic mechanisms between the regulation of alcohol intake
272 and susceptibility to schizophrenia, as well as other neuropsychiatric disorders. In
273 this regard, large prospective epidemiological studies report a three-fold risk of
274 schizophrenia in relation to alcohol abuse⁵³.

275 We previously reported genome-wide significant associations of alcohol intake with
276 *KLB*, and identified a liver-brain axis linking the liver hormone FGF21 with central
277 regulation of alcohol intake involving β -Klotho receptor (the gene product of *KLB*) in
278 the brain⁵. Here, we identify a significant variant near *FGF21* gene and strongly
279 replicate the previously reported *KLB* gene variant, strengthening the genetic
280 evidence for the importance of this pathway in regulating alcohol consumption.

281 The LD score regression analysis showed a positive genetic correlation between
282 alcohol consumption, smoking and HDL cholesterol levels. This confirms previous
283 findings that reported an almost identical genetic correlation of alcohol consumption
284 with number of cigarettes per day⁵⁴. Furthermore, the observed genetic correlation
285 with HDL levels is consistent with previous observations of an association between
286 alcohol consumption and HDL^{55,56}, including results of a Mendelian randomization
287 study that suggested a possible causal role linking alcohol intake with increased HDL
288 levels⁵⁷. Furthermore, we found a genetic correlation (inverse) between sleep
289 duration and alcohol consumption, an association previously reported only in a few
290 small epidemiological studies⁵⁸. We also found a significant genetic correlation with
291 schizophrenia and bipolar disorder, a result that is supported by a recently published
292 trans-ethnic meta-analysis of case-control studies on alcohol dependence⁵⁹. We
293 could not test for a genetic association between alcohol and risk of alcohol-related
294 cancers⁶⁰ because of limited availability of summary data. However, our gene-set
295 enrichment analysis showed a significant enrichment for genes related to abdominal
296 as well as other cancers.

297 Strengths of our study include its size, detailed attention to the alcohol phenotype,
298 dense coverage of the genome through imputation, and incorporation of brain and
299 other imaging data to explore potential mechanisms. Over 80% of the data came
300 from UKB, which combines high-quality phenotypic data and imputed genome-wide
301 genetic data with strict attention to quality control⁶¹. We adopted a stringent
302 approach to claim novel variants involving a conservative *P*-value threshold, internal
303 replication in UKB and consistent direction of effect with the other studies, to
304 minimize the reporting of false positive signals.

305 However, since alcohol intake is socio-culturally as well as genetically determined, it
306 is influenced by other lifestyle and environmental factors which may modify or dilute
307 the genetic signal. A key limitation is that assessment of alcohol intake relies on self-
308 report, which is prone to errors and biases including recall bias and systematic
309 under-reporting by heavy drinkers^{62,63}. Furthermore, questionnaires on alcohol
310 intake covered a short duration (e.g. day or week) at a single period, which may not
311 be representative of broader drinking patterns of cohort participants. We
312 harmonized data across cohorts by converting alcohol intake into a common metric
313 of g/d, with imputation as necessary in UKB for participants reporting consumption
314 of small amounts of alcohol. Taking this approach, we were able to detect strong
315 genetic associations with alcohol intake that explained 7% of the variance in alcohol
316 in an independent cohort, while our GRS analysis indicates that individuals in the

317 lower fifth of the GRS distribution were consuming daily approximately one third of a
318 standard drink (2.6 g/d alcohol) less compared with those in the upper fifth.

319 We should also point out that our eQTL analyses are a first step in the identification
320 of causal genes. Yet, as the most significant eQTLs affected expression of many
321 genes, not necessarily the nearest, there is a need to further prioritize potential
322 causal genes. Unbiased strategies that leverage information from multiple data sets
323 including extensive genomic annotations and high-throughput functional screening
324 in a broad range of tissues will be essential for effective prioritization of genes and
325 uncovering of underlying causal mechanisms⁶⁴. Establishing confidence in the
326 prioritized genes in such a way is a prerequisite for performing functional follow-up
327 studies in appropriate model systems, as demonstrated by the identification of the
328 causal genes and potential disease mechanisms at the obesity- associated *FTO*
329 locus⁶⁵.

330
331 In summary, in this large study of genetic associations with alcohol consumption, we
332 identified common variants in 46 novel loci, with several of the genes expressed in
333 the brain as well as other tissues. Our findings suggest that there may be shared
334 genetic mechanisms underpinning regulation of alcohol intake and development of a
335 neuropsychiatric disorders including schizophrenia. This may form the basis for
336 greater understanding of observed associations between alcohol consumption,
337 schizophrenia⁶⁶ and other disorders.

338 **METHODS**

339

340 **UK Biobank data**

341 We conducted a Genome Wide Association Study (GWAS) analysis among 458,577
342 UKB participants of European descent, identified from a combination of self-
343 reported and genetic data. The details of the selection of the participants has been
344 described elsewhere¹⁴. These comprise 408,951 individuals from UKB genotyped at
345 825,927 variants with a custom Affymetrix UK Biobank Axiom Array chip and 49,626
346 individuals genotyped at 807,411 variants with a custom Affymetrix UK BiLEVE Axiom
347 Array chip from the UK BiLEVE study, which is a subset of UKB. For our analyses, we
348 used SNPs imputed centrally by UKB using the Haplotype Reference Consortium
349 (HRC) panel.

350

351 *Alcohol intake*

352 We calculated the alcohol intake as grams of alcohol per day (g/d) based on self-
353 reported alcohol drinking from the touch-screen questionnaire. The quantity of each

354 type of drink (red wine, white wine, beer/cider, fortified wine, spirits) was multiplied
355 by its standard drink size and reference alcohol content. Drink-specific intake during
356 the reported drinking period (a week for frequent drinkers defined as: daily or
357 almost daily/once or twice a week/three or four times a week; or a month for
358 occasional drinkers defined as: one to three times a month/special occasions only)
359 was summed up and converted to g/d alcohol intake for all participants with
360 complete response to the quantitative drinking questions. The alcohol intake for
361 participants with incomplete response was imputed by bootstrap resampling from
362 the complete responses, stratified by drinking frequency (occasional or frequent)
363 and sex.

364

365 Participants were defined as life-time non-drinkers if they reported 'never' on the
366 question on alcohol drinking frequency (UKB field 1558) and 'no' for the question on
367 former drinker (UKB field 3731); they were excluded from further analysis. We
368 considered participants with alcohol consumption > 500 g/d as outliers and they
369 were dropped from the analyses. We also excluded participants with missing
370 covariates, leaving data on 404,732 individuals. We \log_{10} transformed g/d alcohol
371 and sex-specific residuals were derived from the regression of \log_{10} transformed g/d
372 alcohol on age, age², genotyping chip and weight.

373

374 **UKB genetic analysis**

375 We performed linear mixed modeling using BOLT-LMM software⁶⁷, under an additive
376 genetic model, for associations of measured and imputed SNPs with alcohol
377 consumption (sex-specific residuals of the \log_{10} transformed g/d variable). Model
378 building was based on SNPs with MAF > 5%, call rate > 98.5% and HWE $P > 1 \times 10^{-6}$.
379 SNPs were imputed using the HRC panel with imputation quality INFO score > 0.1.
380 We estimated the LD score regression (LDSR) intercept to assess the degree of
381 genomic inflation beyond polygenicity as well as the lambda inflation factor λ_{GC} ⁶⁸.

382 **The Alcohol Genome-Wide Consortium (AlcGen) and the Cohorts for Heart and** 383 **Aging Research in Genomic Epidemiology Plus (CHARGE+) consortia**

384 We analyzed available GWAS data from 25 independent studies (N=76,111) from the
385 AlcGen and the CHARGE+ consortia. All study participants were of reported
386 European ancestry and data were imputed to either the 1000 Genome Project or the
387 HRC panel. Alcohol intake in g/d was computed and the \log_{10} transformed residuals
388 were analyzed as described above. Study names, cohort information and general
389 study methods are included in **Supplementary Table 2 and 3**.

390 All studies were centrally quality-controlled using easyQC⁶⁹ including filtering for
391 MAF. Finally, we analyzed data on ~7.1 M SNPs at MAF >1% and imputation quality
392 score (Impute [Info score] or Mach [r^2]) > 0.3. Genomic control (GC) was applied at
393 study level. We synthesized the available GWAS using a fixed effects inverse variance
394 weighted meta-analysis and summary estimates were derived for AlcGen and
395 CHARGE+.

396 **One-stage meta-analysis**

397 We performed a one-stage meta-analysis applying a fixed-effects inverse variance
398 weighted meta-analysis using METAL⁷⁰ to obtain summary results from the UKB and
399 and the AlcGen plus CHARGE+ GWAS, for up to N=480,842 participants and ~7.1 M
400 SNPs with MAF \geq 1% for variants present in both the UKB data and AlcGen and
401 CHARGE+ meta-analysis. We assessed the observed heterogeneity using Cochran's Q
402 and we quantified this using the I^2 metric. We considered a Cochran's Q $P < 1 \times 10^{-4}$
403 as significant. The LDSR intercept (standard error), in the discovery meta-analysis
404 was 1.05 and no further correction was applied. QQ plots of the combined meta-
405 analysis summary results, UK Biobank only as well as AlcGen and CHARGE+ only, are
406 presented in **Supplementary Figure 8**.

407

408 **Previously reported (known) SNPs**

409 We looked up in the GWAS catalog (<http://www.ebi.ac.uk/gwas/>) and identified 17
410 SNPs associated with alcohol consumption at genome-wide significance level ($P < 5$
411 $\times 10^{-8}$). We enhanced the list by reference to a recent GWAS by Clarke et al⁶ that
412 was not covered by the GWAS catalog at the time of the analysis, reporting 14
413 additional rare and common SNPs. Together with a SNP in *RASGRF2* shown to be
414 associated with alcohol-induced reinforcement⁷¹, we found 31 previously reported
415 alcohol consumption related SNPs.

416

417 **Novel loci**

418 According to locus definition of i) SNPs within ± 500 kb distance of each other; ii) SNPs
419 in linkage disequilibrium LD ($r^2 > 0.1$) calculated with PLINK, we augmented the list of
420 known SNPs with all SNPs present within our data, not contained within the
421 previously published loci. We further excluded SNPs in the HLA region (chromosome
422 6, 25-34Mb) due to its complex LD structure. We performed LD clumping in PLINK on
423 4,515 unknown SNPs with $P < 1 \times 10^{-8}$ using an $r^2 > 0.1$ and distance threshold of
424 500kb. We further grouped the lead SNPs within 500kb from each other into the
425 same loci and selected the SNP with smallest P -value from the locus as sentinel SNP.

426 To report a SNP as novel signal of association with alcohol consumption:

- 427 i) the sentinel SNP has $P < 5 \times 10^{-9}$ in the one-stage meta-analysis;
428 ii) the sentinel SNP is strongly associated ($P < 5 \times 10^{-7}$) in the UKB GWAS
429 alone;
430 iii) the sentinel SNP has concordant direction of effect between UKB and
431 AlcGen and CHARGE+ datasets;
432 iv) The sentinel SNP is not located within any of the previously reported loci

433 We selected the above criteria i) to iii) to minimize false positive findings including
434 use of a conservative one-stage P -value threshold that is an order of magnitude
435 more stringent than a genome-wide significance P -value. (The threshold of $P < 5 \times$
436 10^{-9} has been proposed e.g. for whole-genome sequencing-based studies.) This
437 approach led us to the identification of 46 sentinel SNPs in total. Regional plots for
438 all 46 sentinel SNPs are presented in **Supplementary Figure 9**.

439

440 **Conditional analysis**

441 We conducted locus-specific conditional analysis using the GCTA (Genome-wide
442 Complex Trait Analysis) software (<http://cnsgenomics.com/software/gcta>). For each
443 of the 46 novel sentinel SNPs, we obtained conditional analysis results for the SNPs
444 with $MAF > 1\%$ and within 500kb from the sentinel SNP after conditioning on the
445 sentinel SNP. The meta-analysis results of the GWAS in UKB, AlcGen and CHARGE+
446 were used as input summary statistics and the individual-level genetic data from UKB
447 were used as the reference sample. Results for a SNP were considered conditionally
448 significant if the difference between the conditional P -value and the original P -value
449 is greater than 1.5-fold ($-\log_{10}P / -\log_{10}(P_{\text{conditional}}) > 1.5$) and the conditional P -
450 value is smaller than 5×10^{-8} .

451

452 **Gene-based analysis**

453 We performed a gene-based analysis using fastBAT, a method that performs a set-
454 based association analysis using summary-level data from GWAS. We used the UKB
455 dataset as a reference set for the LD calculation⁷². Gene-based associations with $P <$
456 5×10^{-9} were considered significant.

457

458 **Gene expression analyses**

459 To analyze the impact of genetic variants on expression of neighboring genes and
460 identify expression quantitative trait loci (*cis*-eQTLs; i.e., SNPs associated with
461 differences in local gene expression), we used two publicly available databases, the
462 Genotype-Tissue Expression (GTEx) database⁷³ (www.gtexportal.org) and the UK
463 Brain Expression Consortium (UKBEC) dataset⁷⁴ (<http://www.braineac.org>). We

464 searched these databases for significant variant-transcripts pairs for genes within
465 1Mb of each input SNP.

466 With the GTEx database, we tested for *cis*-eQTL effects in 48 tissues from 620
467 donors. The data described herein were obtained from the GTEx Portal, Release: V7
468 and used FastQTL⁷⁵, to map SNPs to gene-level expression data and calculate q-
469 values based on beta distribution-adjusted empirical *P*-values⁷⁶. A false discovery
470 rate (FDR) threshold of ≤ 0.05 was applied to identify genes with a significant eQTL.
471 The effect size, defined as the slope of the linear regression, was computed in a
472 normalized space (normalized effect size (NES)), where magnitude has no direct
473 biological interpretation. Here, NES reflects the effects of our GWAS A1 alleles (that
474 are not necessarily the alternative alleles relative to the reference alleles, as
475 reported in the GTEx database). **Supplementary Table 13** lists transcripts-SNPs
476 associations with significant eQTL effects.

477 With the UKBEC dataset that comprises 134 brains (<http://www.braineac.org/>), we
478 searched for *cis*-eQTLs in 10 brain regions, including the cerebellar cortex (CRBL),
479 frontal cortex (FCTX), hippocampus (HIPPO), medulla (specifically inferior olivary
480 nucleus, MEDU), occipital cortex (specifically primary visual cortex, OCTX), putamen
481 (PUTM), substantia nigra (SNIG), thalamus (THAL), temporal cortex (TCTX) and
482 intralobular white matter (WHMT), as well as across all brain tissues (aveALL).
483 MatrixEQTL⁷⁷ generated *P*-values for each expression profile (either exon-level or
484 gene-level) against the respective SNP were obtained for the 10 different tissues and
485 overall (aveALL). **Supplementary Table 14** lists transcripts-SNPs associations with a
486 eQTL *P*-value < 0.0045 in at least one brain tissue. Subsequent data analysis was
487 performed in R (<http://www.R-project.org/>).

488 We carried out over-representation enrichment analysis using a list of 146 GTEx
489 eQTL genes that were derived from the single-variant analysis and a list of 160 eQTL
490 genes that were derived from both single-variant and gene-based analysis. Ingenuity
491 pathway analysis (IPA[®], QIAGEN Inc.) was performed on these lists using ontology
492 annotations from all available databases except those derived from low-confidence
493 computational predictions.

494 **Magnetic Resonance Imaging Data**

495 We used the most recent release of magnetic resonance imaging (MRI) data on
496 brain, heart and liver for UKB participants to investigate genetic associations with the
497 46 novel SNPs for alcohol consumption.

498

499 **Brain imaging**

500

501 *Brain MRI acquisition and pre-processing*

502 We used the T1 data from UKB to elucidate volumetric brain structures, including the
503 cortical and the sub-cortical areas. The T1 data were acquired and pre-processed
504 centrally by UKB. The brain regions were defined by combining the Harvard-Oxford
505 cortical and subcortical atlases⁷⁸ (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases>) and
506 the Diedrichsen cerebellar atlas⁷⁹
507 (<http://www.diedrichsenlab.org/imaging/propatlas.htm>). FAST (FMRIB's Automated
508 Segmentation Tool)⁸⁰ was then used to estimate the grey matter partial volume
509 within each brain region. Subcortical region volumes were also modelled by using
510 FIRST (FMRIB's Integrated Registration and Segmentation Tool). More details about
511 the MRI scanning protocol and pre-processing has been provided in UKB
512 documentation (https://biobank.ctsu.ox.ac.uk/crystal/docs/brain_mri.pdf).

513 *Association Analyses*

514 We performed association analyses on N = 9,702 individuals between all novel SNPs
515 and the grey matter volume of brain regions using Pearson correlation, adjusting for
516 age, age², sex, age × sex, age² × sex, and head size. All, brain volume features, log
517 transformed alcohol intake data (g/d), and the confounders were firstly transformed
518 by using a rank-based inverse Gaussian transformation. Significance levels were set
519 at $P < 0.05$ adjusted using the false-discovery rate method for multiple comparisons.

520

521 *Mediation analysis*

522 To assess if the effect of a SNP on alcohol consumption is mediated through a brain
523 region, we performed a single-level mediation analysis based on a standard three-
524 variable path model (SNP-brain region-alcohol consumption) with corrected and
525 accelerated percentile bootstrapping 10,000 times to calculate the significance of
526 the mediation effect. We considered as mediator variable the grey matter volume of
527 brain regions that had a significant association on alcohol consumption. We
528 calculated the significance of path a, path b and a*b mediation (SNP-brain region-
529 alcohol consumption) using a multilevel mediation and moderation (M3) toolbox^{81,82}.
530 To exclude the possibility of an inverse causal pathway we performed additional
531 analyses in UKB non-drinkers (N =589). performing 10,000 random permutations,
532 associations of rs13107325 with both left and right putamen.

533

534 **Cardiac Imaging**

535

536 *Cardiac MRI acquisition and pre-processing*

537 Details of the cardiac image acquisition in UKB are reported previously⁸³. Cardiac
538 MRI was acquired using a clinical wide bore 1.5T scanner (MAGNETOM Aera, Syngo
539 Platform VD13A, Siemens Healthcare, Erlangen, Germany) with 48 receiver channels,
540 a 45 mT/m and 200 T/m/s gradient system, an 18-channel anterior body surface coil
541 used in combination with 12 elements of an integrated 32 element spine coil and
542 electrocardiogram gating for cardiac synchronization. A two-dimensional short-axis
543 cardiac MRI was obtained using a balanced steady state free precession to cover the
544 entire left and right ventricle (echo time, 1.10msec; repetition time, 2.6msec; flip
545 angle, 80°; slice thickness, 8mm with 2mm gap; typical field of view, 380×252mm;
546 matrix size, 208×187, acquisition of 1 slice per breath-hold).

547 The cardiac images were segmented to provide left ventricular mass (LVM), left end-
548 diastolic (LVEDV), left end-systolic volume (LVESV), and right end-diastolic (RVEDV)
549 and right end-systolic volume (RVESV) using a fully convolutional network as
550 described previously⁸⁴. Left (LVEF) and right ventricular ejection fraction (RVEF) were
551 derived from $(LVEDV-LVESV)/LVEDV \times 100$ and $(RVEDV-RVESV)/RVEDV \times 100$,
552 respectively.

553 *Association Analyses*

554 To test associations between cardiac MRI measures and alcohol consumption-
555 related SNPs, we carried out a regression of LVM, LVEDV, LVEF, RVEDV, and RVEF
556 onto each of the 46 SNPs adjusting for age, sex, height, weight, hypertension
557 (defined as systolic blood pressure >140mmHg and or diastolic blood pressure
558 >90mmHg or under antihypertensive treatment), diabetes, and smoking history on
559 N=10,706 participants. Significance levels were set at $P < 0.05$ adjusted using the
560 false-discovery rate method for multiple comparisons.

561

562 **Liver Imaging**

563 *Liver MRI acquisition and pre-processing*

564 Details of the liver image acquisition protocol have been reported previously⁸⁵.
565 Briefly, all participants were scanned in a Siemens MAGNETOM Aera 1.5-T MRI
566 scanner (Siemens Healthineers, Erlangen, Germany) using a 6-minute dual-echo
567 Dixon Vibe protocol, providing a water and fat separated volumetric data set for fat
568 and muscle covering neck to knees. For liver proton density fat fraction (PDFF)
569 quantification, an additional single multi-echo gradient slice was acquired over the
570 liver. Liver images were analysed by computing specific ROI for water, fat and T2* by

571 magnitude-based chemical shift technique with a 6-peak lipid model, correcting for
572 T1 and T2*.

573

574 *Association Analyses*

575

576 We performed association analyses between 46 alcohol consumption-related SNPs
577 and liver PDDF (%), from 8,479 samples, using a linear regression model adjusting for
578 age, age², sex, T2D, BMI, genotyping chip and first three PCs. Liver PDDF was firstly
579 transformed by using a rank-based inverse transformation. Significance levels were
580 set at $P < 0.05$ adjusted using the false-discovery rate method for multiple
581 comparisons.

582

583 ***Drosophila* experiments**

584 Flies were kept on standard cornmeal/molasses fly food in a 12:12hr light:dark cycle
585 at 25°C. Transgenic flies were obtained from the Bloomington *Drosophila* Stock
586 Center: *UAS-hZip8* BL#66125, *UAS-dZIP71B-TRiP-RNAi^{HMC04064}* BL#55376,
587 *dZip71B^{MI13940}* BL#59234, and *dZip71B^{MB11703}* BL#29928. For behavioral experiments,
588 crosses were set up such that experimental and control flies were sibling progeny
589 from a cross, and both were therefore in the same hybrid genetic background (*w*
590 *Berlin / unknown*). Flies aged 1-5 days of adult age were collected, exposed to
591 100/50 (flowrates) ethanol/air vapor in the Booze-o-Mat 2 days later, and their loss
592 of righting determined by slight tapping, as described⁸⁶. For tolerance, flies were put
593 back onto regular food after a 30-min initial exposure and were then re-exposed to
594 the same vapor 4 hours later. Note that tolerance is not connected to initial
595 sensitivity, and flies naively sensitive to ethanol-induced sedation can have no, or a
596 reduced tolerance phenotype. Flies overexpressing *hZip8* (and their sibling controls)
597 were placed at 28°C for two days to increase the expression levels of the transgene,
598 as we did not detect a phenotype when they were kept at 25°C (data not shown).
599 Data from experimental and control flies were compared by two-sided Student's t-
600 tests. Data were normally distributed according to Shapiro-Wilk testing with
601 Bonferroni adjustment for each of the three experiments.

602

603 **Effects on other traits and diseases**

604 We queried SNPs against GWAS results included in PhenoScanner
605 (<http://www.phenoscaner.medschl.cam.ac.uk>), to investigate cross-trait effects,
606 extracting all association results with genome-wide significance at $P < 5 \times 10^{-8}$ for all
607 SNPs in high LD ($r^2 \geq 0.8$) with the 46 sentinel novel SNPs, to highlight the loci with
608 strongest evidence of association with other traits. At the gene level,

609 overrepresentation enrichment analysis (ORA) with WebGestalt⁴¹ on the nearest
610 genes to all alcohol consumption loci was carried out.

611 The genetic correlations between alcohol consumption and 235 other traits and
612 diseases were obtained in the online software LD Hub. LD hub is a centralized
613 database of summary-level GWAS results and a web interface for LD score regression
614 analysis

615 To estimate the potential causal effect of alcohol consumption-related variants on
616 schizophrenia, we performed a Mendelian randomization analysis utilizing publicly
617 available GWAS data on schizophrenia and the Mendelian randomization package in
618 R. The effect was estimated using the inverse-variance weighted (IVM) method.
619 Pleiotropy was tested by applying the MR-Egger regression method and
620 heterogeneity statistics were obtained. In presence of heterogeneity the random
621 effects inverse-variance method was applied⁸⁷.

622 **Genetic risk scores and percentage of variance explained**

623 We calculated an unbiased weighted GRS in 14,004 unrelated participants in
624 Airwave, an independent cohort with high quality HRC imputed genetic data³³. All
625 previously reported and novel variants were used for the construction of the GRS.
626 We weighted the alcohol-increasing alleles by the beta coefficients of the meta-
627 analysis. We assessed the association of the GRS with alcohol intake and calculated
628 the alcohol consumption levels for individuals in the top vs the bottom quintiles of
629 the distribution. To calculate the percent of variance of alcohol consumption
630 explained by genetic variants, we generated the residuals from a regression of
631 alcohol consumption in Airwave. We then fit a second linear model for the trait
632 residuals with all novel and known variants plus the top 10 principal components and
633 estimated the percentage variance of the dependent variable explained by the
634 variants.

635 **Statistical analysis**

636 All inferential statistics for the analyses described above are provided in the text or
637 in tables and figures. All performed tests were two-sided.

638 **Data availability statement**

639 The UKB GWAS data can be accessed from the UK Biobank data repository
640 (<http://biota.osc.ox.ac.uk/>). The genetic and phenotypic UKB data are available upon
641 application to the UK Biobank (<https://www.ukbiobank.ac.uk>). Summary GWAS data

642 data can be assessed by request to the corresponding authors and will be available
643 via LDHub (<http://ldsc.broadinstitute.org/ldhub/>).

644

645 **References**

646

- 647 1. GBD 2016 Alcohol Collaborators. Alcohol use and burden for 195 countries
648 and territories, 1990-2016: a systematic analysis for the Global Burden of
649 Disease Study 2016. *Lancet* **5**, 987-1012 (2018).
- 650 2. World Health Organization. Global status report on alcohol and health 2018.
651 Eds: Poznyak V and Rekve D,
652 https://www.who.int/substance_abuse/publications/global_alcohol_report/gsr_2018/en/ (2018).
- 653 3. Wood, A.M. *et al.* Risk thresholds for alcohol consumption: combined analysis
654 of individual-participant data for 599 912 current drinkers in 83 prospective
655 studies. *Lancet* **391**, 1513-1523 (2018).
- 656 4. Verhulst, B., Neale, M.C. & Kendler, K.S. The heritability of alcohol use
657 disorders: a meta-analysis of twin and adoption studies. *Psychol Med* **45**,
658 1061-72 (2015).
- 659 5. Schumann, G. *et al.* KLB is associated with alcohol drinking, and its gene
660 product beta-Klotho is necessary for FGF21 regulation of alcohol preference.
661 *Proc Natl Acad Sci U S A* **113**, 14372-14377 (2016).
- 662 6. Clarke, T.K. *et al.* Genome-wide association study of alcohol consumption and
663 genetic overlap with other health-related traits in UK Biobank (N=112 117).
664 *Mol Psychiatry* **22**, 1376-1384 (2017).
- 665 7. Jorgenson, E. *et al.* Genetic contributors to variation in alcohol consumption
666 vary by race/ethnicity in a large multi-ethnic genome-wide association study.
667 *Mol Psychiatry* **22**, 1359-1367 (2017).
- 668 8. Baik, I., Cho, N.H., Kim, S.H., Han, B.G. & Shin, C. Genome-wide association
669 studies identify genetic loci related to alcohol consumption in Korean men.
670 *Am J Clin Nutr* **93**, 809-16 (2011).
- 671 9. Jackson, B. *et al.* Update on the aldehyde dehydrogenase gene (ALDH)
672 superfamily. *Hum Genomics* **5**, 283-303 (2011).
- 673 10. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the
674 causes of a wide range of complex diseases of middle and old age. *PLoS Med*
675 **12**, e1001779 (2015).
- 676 11. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype
677 imputation. *Nat Genet* **48**, 1279-83 (2016).
- 678 12. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from
679 polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).
- 680 13. Evangelou, E. & Ioannidis, J.P. Meta-analysis methods for genome-wide
681 association studies and beyond. *Nat Rev Genet* **14**, 379-89 (2013).
- 682

- 683 14. Evangelou, E. *et al.* Genetic analysis of over 1 million people identifies 535
684 new loci associated with blood pressure traits. *Nat Genet* **50**, 1412-1425
685 (2018).
- 686 15. Desikan, R.S. *et al.* Genetic overlap between Alzheimer's disease and
687 Parkinson's disease at the MAPT locus. *Mol Psychiatry* **20**, 1588-95 (2015).
- 688 16. Do, C.B. *et al.* Web-based genome-wide association study identifies two novel
689 loci and a substantial genetic component for Parkinson's disease. *PLoS Genet*
690 **7**, e1002141 (2011).
- 691 17. Pankratz, N. *et al.* Meta-analysis of Parkinson's disease: identification of a
692 novel locus, RIT2. *Ann Neurol* **71**, 370-84 (2012).
- 693 18. Okbay, A. *et al.* Genetic variants associated with subjective well-being,
694 depressive symptoms, and neuroticism identified through genome-wide
695 analyses. *Nat Genet* **48**, 624-33 (2016).
- 696 19. Couch, F.J. *et al.* Genome-wide association study in BRCA1 mutation carriers
697 identifies novel loci associated with breast and ovarian cancer risk. *PLoS*
698 *Genet* **9**, e1003212 (2013).
- 699 20. Ikram, M.A. *et al.* Common variants at 6q22 and 17q21 are associated with
700 intracranial volume. *Nat Genet* **44**, 539-44 (2012).
- 701 21. van der Harst, P. *et al.* Seventy-five genetic loci influencing the human red
702 blood cell. *Nature* **492**, 369-75 (2012).
- 703 22. Samuel, A. *et al.* Six3 regulates optic nerve development via multiple
704 mechanisms. *Sci Rep* **6**, 20267 (2016).
- 705 23. Schizophrenia Working Group of the Psychiatric Genomics, C. Biological
706 insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-7
707 (2014).
- 708 24. Liu, J.Z. *et al.* Association analyses identify 38 susceptibility loci for
709 inflammatory bowel disease and highlight shared genetic risk across
710 populations. *Nat Genet* **47**, 979-986 (2015).
- 711 25. International Consortium for Blood Pressure Genome-Wide Association
712 Studies *et al.* Genetic variants in novel pathways influence blood pressure
713 and cardiovascular disease risk. *Nature* **478**, 103-9 (2011).
- 714 26. Speliotes, E.K. *et al.* Association analyses of 249,796 individuals reveal 18 new
715 loci associated with body mass index. *Nat Genet* **42**, 937-48 (2010).
- 716 27. Teslovich, T.M. *et al.* Biological, clinical and population relevance of 95 loci
717 for blood lipids. *Nature* **466**, 707-13 (2010).
- 718 28. Lim, C.S. & Alkon, D.L. Protein kinase C stimulates HuD-mediated mRNA
719 stability and protein expression of neurotrophic factors and enhances
720 dendritic maturation of hippocampal neurons in culture. *Hippocampus* **22**,
721 2303-19 (2012).
- 722 29. Barker, J.M., Taylor, J.R., De Vries, T.J. & Peters, J. Brain-derived neurotrophic
723 factor and addiction: Pathological versus therapeutic effects on drug seeking.
724 *Brain Res* **1628**, 68-81 (2015).

- 725 30. Tanaka, T. *et al.* Genome-wide meta-analysis of observational studies shows
726 common genetic variants associated with macronutrient intake. *Am J Clin*
727 *Nutr* **97**, 1395-402 (2013).
- 728 31. Talukdar, S. *et al.* FGF21 Regulates Sweet and Alcohol Preference. *Cell Metab*
729 **23**, 344-9 (2016).
- 730 32. Grant, S.F. *et al.* Association analysis of the FTO gene with obesity in children
731 of Caucasian and African ancestry reveals a common tagging SNP. *PLoS One* **3**,
732 e1746 (2008).
- 733 33. Elliott, P. *et al.* The Airwave Health Monitoring Study of police officers and
734 staff in Great Britain: rationale, design and methods. *Environ Res* **134**, 280-5
735 (2014).
- 736 34. Elliott, L.T. *et al.* Genome-wide association studies of brain imaging
737 phenotypes in UK Biobank. *Nature* **562**, 210-216 (2018).
- 738 35. Stipanovich, A. *et al.* A phosphatase cascade by which rewarding stimuli
739 control nucleosomal response. *Nature* **453**, 879-84 (2008).
- 740 36. Yang, B.Z. *et al.* Association of haplotypic variants in DRD2, ANKK1, TTC12 and
741 NCAM1 to alcohol dependence in independent case control and family
742 samples. *Hum Mol Genet* **16**, 2844-53 (2007).
- 743 37. Gelernter, J. *et al.* Haplotype spanning TTC12 and ANKK1, flanked by the
744 DRD2 and NCAM1 loci, is strongly associated to nicotine dependence in two
745 distinct American populations. *Hum Mol Genet* **15**, 3498-507 (2006).
- 746 38. Treutlein, J. *et al.* Genetic association of the human corticotropin releasing
747 hormone receptor 1 (CRHR1) with binge drinking and alcohol intake patterns
748 in two independent samples. *Mol Psychiatry* **11**, 594-602 (2006).
- 749 39. Timpl, P. *et al.* Impaired stress response and reduced anxiety in mice lacking a
750 functional corticotropin-releasing hormone receptor 1. *Nat Genet* **19**, 162-6
751 (1998).
- 752 40. Ruggieri, B. *et al.* Association of Protein Phosphatase PPM1G With Alcohol
753 Use Disorder and Brain Activity During Behavioral Control in a Genome-Wide
754 Methylation Analysis. *Am J Psychiatry* **172**, 543-52 (2015).
- 755 41. Wang, J., Vasaiakar, S., Shi, Z., Greer, M. & Zhang, B. WebGestalt 2017: a more
756 comprehensive, powerful, flexible and interactive gene set enrichment
757 analysis toolkit. *Nucleic Acids Res* **45**, W130-W137 (2017).
- 758 42. Gonzalez, D.A. *et al.* The Arf6 activator Efa6/PSD3 confers regional specificity
759 and modulates ethanol consumption in *Drosophila* and humans. *Mol*
760 *Psychiatry* **23**, 621-628 (2018).
- 761 43. Ojelade, S.A. *et al.* Rsu1 regulates ethanol consumption in *Drosophila* and
762 humans. *Proc Natl Acad Sci U S A* **112**, E4085-93 (2015).
- 763 44. Rademakers, R., Cruts, M. & van Broeckhoven, C. The role of tau (MAPT) in
764 frontotemporal dementia and related tauopathies. *Hum Mutat* **24**, 277-95
765 (2004).
- 766 45. Higashi, Y. *et al.* Influence of extracellular zinc on M1 microglial activation. *Sci*
767 *Rep* **7**, 43778 (2017).

- 768 46. Chen, G. *et al.* Striatal involvement in human alcoholism and alcohol
769 consumption, and withdrawal in animal models. *Alcohol Clin Exp Res* **35**,
770 1739-48 (2011).
- 771 47. Okada, N. *et al.* Abnormal asymmetries in subcortical brain volume in
772 schizophrenia. *Mol Psychiatry* **21**, 1460-6 (2016).
- 773 48. van Erp, T.G. *et al.* Subcortical brain volume abnormalities in 2028 individuals
774 with schizophrenia and 2540 healthy controls via the ENIGMA consortium.
775 *Mol Psychiatry* **21**, 547-53 (2016).
- 776 49. Meyers, J.L. *et al.* The association between DRD2/ANKK1 and genetically
777 informed measures of alcohol use and problems. *Addict Biol* **18**, 523-36
778 (2013).
- 779 50. Logrip, M.L., Barak, S., Warnault, V. & Ron, D. Corticostriatal BDNF and
780 alcohol addiction. *Brain Res* **1628**, 60-7 (2015).
- 781 51. Boschen, K.E., Criss, K.J., Palamarchouk, V., Roth, T.L. & Klintsova, A.Y. Effects
782 of developmental alcohol exposure vs. intubation stress on BDNF and TrkB
783 expression in the hippocampus and frontal cortex of neonatal rats. *Int J Dev*
784 *Neurosci* **43**, 16-24 (2015).
- 785 52. Monaco, A. *et al.* A complex network approach reveals a pivotal substructure
786 of genes linked to schizophrenia. *PLoS One* **13**, e0190110 (2018).
- 787 53. Nielsen, S.M., Toftdahl, N.G., Nordentoft, M. & Hjorthoj, C. Association
788 between alcohol, cannabis, and other illicit substance abuse and risk of
789 developing schizophrenia: a nationwide population based register study.
790 *Psychol Med* **47**, 1668-1677 (2017).
- 791 54. Nivard, M.G. *et al.* Connecting the dots, genome-wide association studies in
792 substance use. *Mol Psychiatry* **21**, 733-5 (2016).
- 793 55. Gaziano, J.M. *et al.* Moderate alcohol intake, increased levels of high-density
794 lipoprotein and its subfractions, and decreased risk of myocardial infarction.
795 *N Engl J Med* **329**, 1829-34 (1993).
- 796 56. Linn, S. *et al.* High-density lipoprotein cholesterol and alcohol consumption in
797 US white and black adults: data from NHANES II. *Am J Public Health* **83**, 811-6
798 (1993).
- 799 57. Vu, K.N. *et al.* Causal Role of Alcohol Consumption in an Improved Lipid
800 Profile: The Atherosclerosis Risk in Communities (ARIC) Study. *PLoS One* **11**,
801 e0148765 (2016).
- 802 58. Chaput, J.P., McNeil, J., Despres, J.P., Bouchard, C. & Tremblay, A. Short sleep
803 duration is associated with greater alcohol consumption in adults. *Appetite*
804 **59**, 650-5 (2012).
- 805 59. Walters, R.K. *et al.* Transancestral GWAS of alcohol dependence reveals
806 common genetic underpinnings with psychiatric disorders. *Nat Neurosci* **21**,
807 1656-1669 (2018).
- 808 60. Bagnardi, V. *et al.* Alcohol consumption and site-specific cancer risk: a
809 comprehensive dose-response meta-analysis. *Br J Cancer* **112**, 580-93 (2015).

- 810 61. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and
811 genomic data. *Nature* **562**, 203-209 (2018).
- 812 62. Boniface, S., Kneale, J. & Shelton, N. Drinking pattern is more strongly
813 associated with under-reporting of alcohol consumption than socio-
814 demographic factors: evidence from a mixed-methods study. *BMC Public*
815 *Health* **14**, 1297 (2014).
- 816 63. Greenfield, T.K. & Kerr, W.C. Alcohol measurement methodology in
817 epidemiology: recent advances and opportunities. *Addiction* **103**, 1082-99
818 (2008).
- 819 64. Grotz, A.K., Gloyn, A.L. & Thomsen, S.K. Prioritising Causal Genes at Type 2
820 Diabetes Risk Loci. *Curr Diab Rep* **17**, 76 (2017).
- 821 65. Claussnitzer, M. *et al.* FTO Obesity Variant Circuitry and Adipocyte Browning
822 in Humans. *N Engl J Med* **373**, 895-907 (2015)
- 823 66. Hambrecht, M. & Hafner, H. Substance abuse and the onset of schizophrenia.
824 *Biol Psychiatry* **40**, 1155-63 (1996).
- 825 67. Loh, P.R. *et al.* Efficient Bayesian mixed-model analysis increases association
826 power in large cohorts. *Nat Genet* **47**, 284-90 (2015).
- 827 68. Georgiopoulos, G. & Evangelou, E. Power considerations for lambda inflation
828 factor in meta-analyses of genome-wide association studies. *Genet Res*
829 *(Camb)* **98**, e9 (2016).
- 830 69. Winkler, T.W. *et al.* Quality control and conduct of genome-wide association
831 meta-analyses. *Nat Protoc* **9**, 1192-212 (2014).
- 832 70. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of
833 genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
- 834 71. Stacey, D. *et al.* RASGRF2 regulates alcohol-induced reinforcement by
835 influencing mesolimbic dopamine neuron activity and dopamine release. *Proc*
836 *Natl Acad Sci U S A* **109**, 21128-33 (2012).
- 837 72. Bakshi, A. *et al.* Fast set-based association analysis using summary data from
838 GWAS identifies novel gene loci for human complex traits. *Sci Rep* **6**, 32894
839 (2016).
- 840 73. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet*
841 **45**, 580-5 (2013).
- 842 74. Ramasamy, A. *et al.* Genetic variability in the regulation of gene expression in
843 ten regions of the human brain. *Nat Neurosci* **17**, 1418-1428 (2014).
- 844 75. Ongen, H., Buil, A., Brown, A.A., Dermitzakis, E.T. & Delaneau, O. Fast and
845 efficient QTL mapper for thousands of molecular phenotypes. *Bioinformatics*
846 **32**, 1479-85 (2016).
- 847 76. Storey, J.D. & Tibshirani, R. Statistical significance for genomewide studies.
848 *Proc Natl Acad Sci U S A* **100**, 9440-5 (2003).
- 849 77. Shabalin, A.A. Matrix eQTL: ultra fast eQTL analysis via large matrix
850 operations. *Bioinformatics* **28**, 1353-8 (2012).

- 851 78. Brown, C.A. *et al.* Development, validation and application of a new fornix
852 template for studies of aging and preclinical Alzheimer's disease. *Neuroimage*
853 *Clin* **13**, 106-115 (2017).
- 854 79. Diedrichsen, J. *et al.* Imaging the deep cerebellar nuclei: a probabilistic atlas
855 and normalization procedure. *Neuroimage* **54**, 1786-94 (2011).
- 856 80. Zhang, Y., Brady, M. & Smith, S. Segmentation of brain MR images through a
857 hidden Markov random field model and the expectation-maximization
858 algorithm. *IEEE Trans Med Imaging* **20**, 45-57 (2001).
- 859 81. Wager, T.D., Davidson, M.L., Hughes, B.L., Lindquist, M.A. & Ochsner, K.N.
860 Prefrontal-subcortical pathways mediating successful emotion regulation.
861 *Neuron* **59**, 1037-50 (2008).
- 862 82. Wager, T.D. *et al.* Brain mediators of cardiovascular responses to social
863 threat: part I: Reciprocal dorsal and ventral sub-regions of the medial
864 prefrontal cortex and heart-rate reactivity. *Neuroimage* **47**, 821-35 (2009).
- 865 83. Petersen, S.E. *et al.* UK Biobank's cardiovascular magnetic resonance
866 protocol. *J Cardiovasc Magn Reson* **18**, 8 (2016).
- 867 84. Bai, W. *et al.* Automated cardiovascular magnetic resonance image analysis
868 with fully convolutional networks. *J Cardiovasc Magn Reson* **20**, 65 (2018).
- 869 85. Linge, J. *et al.* Body Composition Profiling in the UK Biobank Imaging Study.
870 *Obesity (Silver Spring)* (2018).
- 871 86. Peru, Y.C.d.P.R.L. *et al.* Adult neuronal Arf6 controls ethanol-induced
872 behavior with Arfaptin downstream of Rac1 and RhoGAP18B. *J Neurosci* **32**,
873 17706-13 (2012).
- 874 87. Dimou, N.L. & Tsilidis, K.K. A Primer in Mendelian Randomization
875 Methodology with a Focus on Utilizing Published Summary Association Data.
876 *Methods Mol Biol* **1793**, 211-230 (2018).

877

878 **Acknowledgements**

879 H.G. was funded by the NIHR Imperial College Health Care NHS Trust and Imperial
880 College London Biomedical Research Centre. I.K. was supported by the EU
881 PhenoMeNal project (Horizon 2020, 654241) and the UK Dementia Research
882 Institute, which is supported by the MRC, the Alzheimer's Society and Alzheimer's
883 Research UK. S.Thériault was supported by the Canadian Institutes of Health
884 Research and Université Laval (Quebec City, Canada). L.R. was supported by
885 Forschungs- und Förder-Stiftung INOVA, Vaduz, Liechtenstein. D.C. holds a McMaster
886 University Department of Medicine Mid-Career Research Award. M.B. is supported
887 by NIH grant R01-DK062370. P.v.d.H. was supported by ICIN-NHI and Marie
888 Sklodowska-Curie GF (call: H2020-MSCA-IF-2014, Project ID: 661395). C.H. was
889 supported by a core MRC grant to the MRCHGU QTL in Health and Disease research
890 programme. N.V. was supported by Marie Sklodowska-Curie GF grant (661395) and
891 ICIN-NHI. P.E. acknowledges support from the NIHR Biomedical Research Centre at

892 Imperial College Healthcare NHS Trust and Imperial College London, the NIHR Health
893 Protection Research Unit in Health Impact of Environmental Hazards (HPRU-2012-
894 10141), the Medical Research Council (MRC) and Public Health England (PHE) Centre
895 for Environment and Health (MR/L01341X/1) and Health Data Research (HDR) UK.
896 P.E. is a UK Dementia Research Institute (DRI) professor, UK DRI at Imperial College
897 London, funded by the MRC, Alzheimer's Society and Alzheimer's Research UK. This
898 work received support from the following sources: the European Union-funded FP6
899 Integrated Project IMAGEN (Reinforcement-related behaviour in normal brain
900 function and psychopathology) (LSHM-CT- 2007-037286), the Horizon 2020 funded
901 ERC Advanced Grant 'STRATIFY' (Brain network based stratification of reinforcement-
902 related disorders) (695313), ERANID (Understanding the Interplay between Cultural,
903 Biological and Subjective Factors in Drug Use Pathways) (PR-ST-0416-10004),
904 BRIDGET (JPND: BRain Imaging, cognition Dementia and next generation GENomics)
905 (MR/N027558/1), the FP7 projects IMAGEMEND(602450; IMAGING GENetics for
906 MENTAL Disorders) and MATRICS (603016), the Innovative Medicine Initiative Project
907 EU-AIMS (115300-2), the Medical Research Council Grant 'c-VEDA' (Consortium on
908 Vulnerability to Externalizing Disorders and Addictions) (MR/N000390/1), the
909 Swedish Research Council FORMAS, the Medical Research Council, the National
910 Institute for Health Research (NIHR) Biomedical Research Centre at South London
911 and Maudsley NHS Foundation Trust and King's College London, the
912 Bundesministerium für Bildung und Forschung (BMBF grants 01GS08152; 01EV0711;
913 eMED SysAlc01ZX1311A; Forschungsnetz AERIAL 01EE1406A, 01EE1406B), the
914 Deutsche Forschungsgemeinschaft (DFG grants SM 80/7-2, SFB 940/2), the Medical
915 Research Foundation and Medical research council (grant MR/R00465X/1), the
916 Human Brain Project (HBP SGA 2). Further support was provided by grants from: ANR
917 (project AF12-NEUR0008-01 - WM2NA, and ANR-12-SAMA-0004), the Fondation de
918 France, the Fondation pour la Recherche Médicale, the Mission Interministérielle de
919 Lutte-contre-les-Drogues-et-les-Conduites-Addictives (MILDECA), the Assistance-
920 Publique-Hôpitaux-de-Paris and INSERM (interface grant), Paris Sud University IDEX
921 2012; the National Institutes of Health, Science Foundation Ireland (16/ERCD/3797),
922 U.S.A. (Axon, Testosterone and Mental Health during Adolescence; RO1 MH085772-
923 01A1), and by NIH Consortium grant U54 EB020403, supported by a cross-NIH
924 alliance that funds Big Data to Knowledge Centres of Excellence.

925 The funders had no role in study design, data collection and analysis, decision to
926 publish or preparation of the manuscript.

927

928 **Competing Interests**

929 B.M.P. serves on the DSMB of a clinical trial funded by the manufacturer (Zoll LifeCor)
930 and on the Steering Committee of the Yale Open Data Access Project funded by
931 Johnson & Johnson.

932 B.W.J.H.P. has received research funding (non-related to the work reported here)
933 from Jansen Research and Boehringer Ingelheim.

934 The other authors declare no competing interests.

935

936 **Author contributions**

937

938 **Central analysis:** E.E., H.G., C.C., G.N., P.B., A.R.B., R.P., H.Suzuki, F.K., A.M.Y., I.K.,
939 J.E., Q.L., N.D., D.L., I.T., J.D.B., P.M.M., A.R., S.D., G.S., P.E.

940

941 **Writing of the manuscript:** E.E., H.G., C.C., G.N., P.B., A.R.B., R.P., H.Suzuki, F.K.,
942 A.M.Y., I.K., D.L., I.T., J.D.B., P.M.M., A.R., S.D., G.S., P.E.

943

944 **Association of MRI analysis:** C.C., H.Suzuki, A.M.Y., A.I.B., J.D.B., P.M.M., G.S.

945

946 **Alcgen and Charge+ contributor:** (ARIC): A.C.M., M.R.B., B.Y., D.E.A., (CHS):
947 B.M.P., R.N.L., T.M.B., J.A.B., (FHS): D.L., C.L., (GAPP/Swiss-AF/Beat-AF):
948 S.Thériault, S.A., D.C., L.R., M.Kühne, (GENOA): S.L.R.K., J.A.S., W.A., S.M.R.,
949 (GRAPHIC): N.J.S., C.P.N., P.S.B., (GS): A.M.M., T-K.C., C.H., D.P., (HBCS): L.J.,
950 S.Tuominen, M.M.P., J.G.E., (HRS): D.R.W., S.L.R.K., J.D.F., W.Z., J.A.S.,
951 (MESA): X.G., J.Y., A.W., J.I.R., (METSIM): M.L., A.S., J.Vangipurapu, J.K.,
952 (FUSION): M.B., K.L.M., L.J.S., A.U.J., (NESDA): B.W.J.H.P., Y.M., (NFBC): M-
953 R.J., J.Veijola, M.Männikkö, J.A., (ORCADES): H.C., P.K.J, (VIKING): J.F.W.,
954 K.A.K., (Croatia-VIS): I.R., O.P., (Croatia-KORCULA): C.H., (PREVEND): N.V.,
955 P.v.d.H., (OZALC): N.G.M., J.B.W., P.A.L., A.C.H., (SHIP): A.T., H.J.G., S.E.B.,
956 G.H., (TRAILS-pop): A.J.O, I.M.N., (TRAILS-CC): C.A.H., H.Snieder , (TwinsUK):
957 T.D.S, M.Mangino, (YFS): L-P.L., M.Kähönen, O.T.R., T.L.

958

959 **All authors critically reviewed and approved the final version of the**
960 **manuscript**

961

Table 1: Association results of 46 novel alcohol variants identified through the meta-analysis of UK Biobank and AlcGen and CHARGE+. Results are ordered by P-value of combined analysis.

leadSNP		Combined				UKB			AlcGen and CHARGE+					
Nearest_Gene	Annotated Gene	rsID_LEAD_SNP	CP	EA	EAF	BETA	SE	P	BETA	SE	P	BETA	SE	P
MAPT	STH	rs1991556	17:44083402	A	0.22	-0.012	0.001	4.5E-23	-0.013	0.001	2.4E-21	-0.011	0.004	4.0E-03
RP11-89K21.1	SIX3	rs1004787	2:45159091	A	0.54	0.009	0.001	6.7E-17	0.009	0.001	1.1E-15	0.007	0.003	1.4E-02
SLC39A8	SLC39A8	rs13107325	4:103188709	T	0.07	-0.016	0.002	1.3E-15	-0.017	0.002	4.8E-16	-0.006	0.006	3.6E-01
IZUMO1, RASIP1, FUT1	IZUMO1	rs838145	19:49248730	A	0.55	-0.008	0.001	3.2E-15	-0.009	0.001	2.4E-15	-0.004	0.003	1.7E-01
na	PSMD7	rs1104608	16:73912588	C	0.43	-0.008	0.001	1.2E-14	-0.009	0.001	4.9E-15	-0.003	0.003	2.5E-01
MYBPC3	MYBPC3	rs2071305	11:47370957	A	0.69	0.009	0.001	4.5E-14	0.009	0.001	3.9E-13	0.007	0.003	3.1E-02
na	DRD2	rs7121986	11:113355444	T	0.37	-0.008	0.001	6.2E-14	-0.008	0.001	1.3E-13	-0.005	0.003	1.1E-01
na	DPP6	rs6969458	7:153489725	A	0.47	0.008	0.001	6.4E-14	0.008	0.001	1.3E-12	0.007	0.003	1.5E-02
RP11-308N19.1	ZNF462	rs74424378	9:109331094	T	0.76	0.009	0.001	1.7E-13	0.009	0.001	4.5E-13	0.006	0.003	8.4E-02
ARHGAP15, AC096558.1, RP11-570L15.2	ARHGAP15	rs13024996	2:144225215	A	0.37	-0.008	0.001	4.4E-13	-0.008	0.001	6.6E-13	-0.004	0.003	1.4E-01
MLXIPL	MLXIPL	rs34060476	7:73037956	A	0.87	-0.011	0.002	5.0E-13	-0.012	0.002	1.4E-13	-0.004	0.004	4.1E-01
na	FAM178A	rs61873510	10:102626510	T	0.33	-0.008	0.001	5.1E-13	-0.008	0.001	9.8E-12	-0.008	0.003	1.7E-02
FTO	FTO	rs1421085	16:53800954	T	0.60	0.008	0.001	9.2E-13	0.007	0.001	1.7E-10	0.010	0.003	9.2E-04
na	PMFBP1	rs11648570	16:72356964	T	0.89	-0.012	0.002	2.1E-12	-0.011	0.002	1.5E-10	-0.013	0.005	3.4E-03
OTX2, RP11-1085N6.6	OTX2	rs2277499	14:57271127	T	0.34	-0.008	0.001	2.2E-12	-0.007	0.001	2.4E-09	-0.012	0.003	9.1E-05
PDE4B	PDE4B	rs2310752	1:66392405	A	0.43	-0.007	0.001	2.8E-12	-0.008	0.001	1.8E-11	-0.006	0.003	4.2E-02
SERPINA1	SERPINA1	rs112635299	14:94838142	T	0.02	-0.025	0.004	3.7E-12	-0.027	0.004	9.8E-12	-0.017	0.010	9.9E-02
na	AJAP1	rs780569	1:4569436	A	0.71	-0.008	0.001	5.2E-12	-0.008	0.001	1.1E-11	-0.005	0.003	1.2E-01
na	VRK2	rs10496076	2:57942987	T	0.37	-0.007	0.001	9.7E-12	-0.007	0.001	1.3E-09	-0.009	0.003	1.6E-03
ACTR10, C14orf37	ACTR10	rs71414193	14:58685301	A	0.19	-0.009	0.001	1.8E-11	-0.008	0.001	5.8E-09	-0.013	0.004	4.5E-04
BEND4	BEND4	rs16854020	4:42117559	A	0.13	0.010	0.002	2.9E-11	0.010	0.002	5.8E-09	0.016	0.005	6.4E-04
na	SORL1	rs485425	11:121544984	C	0.45	-0.007	0.001	6.1E-11	-0.007	0.001	7.3E-11	-0.004	0.003	1.9E-01
SEZ6L2	SEZ6L2	rs113443718	16:29892184	A	0.31	-0.007	0.001	7.4E-11	-0.008	0.001	4.5E-11	-0.003	0.003	2.9E-01
CBX5, RP11-968A15.2	CBX5	rs57281063	12:54660427	A	0.41	0.007	0.001	7.9E-11	0.007	0.001	1.8E-09	0.007	0.003	1.2E-02
na	TNRC6A	rs72768626	16:24693048	A	0.94	0.014	0.002	9.7E-11	0.015	0.002	1.7E-09	0.014	0.006	1.8E-02
SYT14	SYT14	rs227179	1:210216731	A	0.59	-0.007	0.001	1.1E-10	-0.007	0.001	1.4E-09	-0.006	0.003	2.8E-02
TCF4	TCF4	rs9320010	18:53053897	A	0.60	0.007	0.001	1.1E-10	0.007	0.001	1.6E-09	0.007	0.003	2.2E-02
SBK1	NPIP6	rs2726034	16:28336882	T	0.68	0.007	0.001	1.4E-10	0.007	0.001	1.1E-09	0.006	0.003	4.7E-02
ANKRD36	ANKRD36	rs13390019	2:97797680	T	0.87	0.010	0.002	1.6E-10	0.011	0.002	7.0E-11	0.004	0.005	4.5E-01
na	ELAVL4	rs7517344	1:50711961	A	0.17	0.009	0.001	1.9E-10	0.008	0.001	2.5E-07	0.016	0.004	2.1E-05
LINC00461	MEF2C	rs4916723	5:87854395	A	0.58	0.007	0.001	2.1E-10	0.007	0.001	5.1E-10	0.005	0.003	1.1E-01
ARPC1B, ARPC1A	ARPC1B	rs10249167	7:98980879	A	0.87	0.010	0.002	2.9E-10	0.009	0.002	8.1E-08	0.015	0.004	3.8E-04
EFNB3, WRAP53	EFNB3	rs7640	17:7606722	C	0.80	0.008	0.001	4.3E-10	0.009	0.001	1.3E-09	0.006	0.004	9.9E-02
RP11-501C14.5	IGF2BP1	rs4794015	17:47067826	A	0.41	0.007	0.001	4.3E-10	0.006	0.001	5.4E-08	0.009	0.003	1.2E-03
TCAP, PNMT, STARD3	TCAP	rs1053651	17:37822311	A	0.27	-0.007	0.001	1.1E-09	-0.008	0.001	8.4E-10	-0.003	0.003	2.8E-01
na	AADAT	rs7698119	4:171070910	A	0.49	-0.006	0.001	1.3E-09	-0.006	0.001	1.6E-07	-0.009	0.003	1.6E-03
STAT6, AC023237.1	STAT6	rs12312693	12:57511734	T	0.55	-0.006	0.001	1.5E-09	-0.006	0.001	9.5E-09	-0.005	0.003	5.6E-02
SCN8A	SCN8A	rs7958704	12:51984349	T	0.41	-0.006	0.001	1.6E-09	-0.006	0.001	1.7E-08	-0.006	0.003	3.5E-02
ACSS3	ACSS3	rs11114787	12:81595700	T	0.27	0.007	0.001	2.0E-09	0.007	0.001	2.7E-08	0.007	0.003	2.4E-02
RP11-32K4.1	BHLHE22	rs2356369	8:64956882	T	0.52	-0.006	0.001	2.0E-09	-0.006	0.001	4.1E-08	-0.007	0.003	1.6E-02
ZRANB2-AS2	ZRANB2	rs12031875	1:71585097	A	0.82	-0.008	0.001	2.2E-09	-0.008	0.001	7.6E-08	-0.010	0.004	8.7E-03
MSANTD1, HTT	MSANTD1	rs12646808	4:3249828	T	0.66	0.007	0.001	2.4E-09	0.007	0.001	1.1E-09	0.002	0.003	4.7E-01
TENM2	TENM2	rs10078588	5:166816176	A	0.52	0.006	0.001	2.5E-09	0.006	0.001	4.3E-08	0.007	0.003	1.9E-02
IGSF9B	IGSF9B	rs748919	11:133783232	T	0.79	0.008	0.001	3.3E-09	0.008	0.001	1.0E-08	0.005	0.003	1.1E-01
AC010967.2	GPR75-ASB3	rs785293	2:53023304	A	0.57	-0.006	0.001	3.3E-09	-0.006	0.001	3.2E-08	-0.006	0.003	3.8E-02
BDNF, RP11-587D21.4	BDNF	rs988748	11:27724745	C	0.21	-0.008	0.001	4.4E-09	-0.007	0.001	1.2E-07	-0.010	0.004	8.3E-03

SNP: Single Nucleotide polymorphism; LocusName: Nearest Gene; rsID_LEAD_SNP: Rs ID number of the lead SNP; CP: Chromosome/Position (build hg19/37); EA: Effect allele of the discovered SNP; EAF: Frequency of the effect allele; BETA_comb: Effect size in meta-analysis; SE_comb: Standard Error of the effect in meta-analysis; P_comb: Meta-analysis P-value; BETA_UKB: Effect size in UK Biobank analysis; SE_UKB: Standard Error of the effect in the UK Biobank analysis; P_UKB: UK Biobank analysis P-value; BETA_AlcGenCHARGE+: Effect size in the AlcGen meta-analysis; SE_AlcGenCHARGE+: Standard Error of the effect in the AlcGen meta-analysis; P_AlcGenCHARGE+: AlcGen meta-analysis P-value

962 **FIGURE CAPTIONS**

963 **Figure 1. Manhattan plot showing P -values from discovery genome-wide**
964 **association meta-analysis with alcohol intake (log g/d) among 480,842 individuals**
965 **across UK Biobank, AlcGen and CHARGE+, excluding known variants.** The P -value
966 was computed using inverse variance fixed effects models. The y axis shows the $-\log_{10} P$ values and the x axis shows their chromosomal positions. Horizontal blue line
967 represents the threshold of $P = 5 \times 10^{-9}$.
968

969
970 **Figure 2. Association of alcohol intake loci with other traits.** Plot shows results from
971 associations with other traits which were extracted from the PhenoScanner database
972 for the 46 novel sentinel SNPs including proxies in Linkage Disequilibrium ($r^2 \geq 0.8$)
973 with genome-wide significant associations. Each colored line connects a specific
974 variant with the associated traits and diseases.
975

976 **Figure 3. Mediation effect of the grey matter volume of bilateral putamen on the**
977 **relationship between SNP rs13107325 and alcohol intake.** The green is for left
978 putamen, and, the red is for the right one. We use 'a' for the relationship between
979 rs13107325 and putamen, 'b' for the relationship between putamen and alcohol
980 consumption, 'c' for the relationship between rs13107325 and alcohol consumption,
981 'c'' for the relationship between rs13107325 and alcohol consumption after
982 excluding the effect of putamen, and 'ab' as the mediation effect. The significance
983 tests are based on the bootstrapping method (10,000 times). Z- statistics and the
984 corresponding P values are provided in parentheses. The brain icon was created
985 using Mango software, version 4.1 (<http://ric.uthscsa.edu/mango/>).
986

987 **Figure 4. Comparison of *Zip8* alcohol phenotypes in *Drosophila*.** Flies were exposed
988 to 100/50 Ethanol/Air vapor for 30 min for exposure 1, and the time to 50% loss of
989 righting was determined (ST-50, sedation time). After recovery on food for 4 hours,
990 flies were re-exposed to the same vapors, and the second ST-50 recorded (left side).
991 The resulting increase in ST-50, i.e. tolerance, is shown on the right. In a)
992 overexpressed human *hZIP8* in *ics*-expressing cells flies are compared against
993 controls whereas in b) knockdown of the fly ortholog *dZip71B* is compared against
994 controls. In c) flies carrying two transposon insertions in the endogenous *dZip71B*
995 gene are compared against controls. Significance levels: *** $P < 0.001$, ** $P < 0.01$, * P
996 < 0.05 . Exact P -values are presented in the text.
997