ARTICLES
https://doi.org/10.1038/s41562-019-0653-z
nature human behaviour

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

Evangelos Evangelou[1,2,94], He Gao[1,3,94], Congying Chu[4,94], Georgios Ntritsos[2,94], Paul Blakeley[1,5], Andrew R. Butts[6], Raha Pazoki[1], Hideaki Suzuki[7,8], Fotios Koskeridis[2], Andrianos M. Yiorkas[9,10], Ibrahim Karaman[1,11], Joshua Elliott[1], Qiang Luo[12,13], Stefanie Aeschbacher[14], Traci M. Bartz[15,16], Sebastian E. Baumeister[17,18], Peter S. Braund[19,20], Michael R. Brown[21], Jennifer A. Brody[15], Toni-Kim Clarke[22], Niki Dimou[2], Jessica D. Faul[23], Georg Homuth[24], Anne U. Jackson[25], Katherine A. Kentistou[26,27], Peter K. Joshi[26], Rozenn N. Lemaitre[15], Penelope A. Lind[28], Leo-Pekka Lyytikäinen[29,30,31], Massimo Mangino[32,33], Yuri Milaneschi[34], Christopher P. Nelson[19,20], Ilja M. Nolte[35], Mia-Maria Perälä[36,37], Ozren Polasek[38], David Porteous[39,40], Scott M. Ratliff[41], Jennifer A. Smith[23,41], Alena Stančáková[42], Alexander Teumer[17,43], Samuli Tuominen[44], Sébastien Thériault[45,46], Jagadish Vangipurapu[42], John B. Whitfield[47], Alexis Wood[48], Jie Yao[49], Bing Yu[21], Wei Zhao[41], Dan E. Arking[50], Juha Auvinen[51,52], Chunyu Liu[53], Minna Männikkö[54], Lorenz Risch[55,56,57], Jerome I. Rotter[58], Harold Snieder[35], Juha Veijola[59,60,61], Alexandra I. Blakemore[9,10], Michael Boehnke[25], Harry Campbell[26], David Conen[45], Johan G. Eriksson[62,63,64], Hans J. Grabe[65,66], Xiuqing Guo[49], Pim van der Harst[67,68,69], Catharina A. Hartman[70], Caroline Hayward[71], Andrew C. Heath[72], Marjo-Riitta Jarvelin[73,74,75,76,77], Mika Kähönen[78,79], Sharon L. R. Kardia[41], Michael Kühne[14], Johanna Kuusisto[80], Markku Laakso[80], Jari Lahti[44], Terho Lehtimäki[29,30], Andrew M. McIntosh[22,40], Karen L. Mohlke[81], Alanna C. Morrison[21], Nicholas G. Martin[47], Albertine J. Oldehinkel[70], Brenda W. J. H. Penninx[34], Bruce M. Psaty[82,83], Olli T. Raitakari[84,85,86], Igor Rudan[26], Nilesh J. Samani[19,20], Laura J. Scott[25], Tim D. Spector[32], Niek Verweij[67], David R. Weir[23], James F. Wilson[26,71], Daniel Levy[87,88], Ioanna Tzoulaki[1,3], Jimmy D. Bell[89,94], Paul M. Matthews[7,11,94], Adrian Rothenfluh[6,90,94], Sylvane Desrivières[4,94], Gunter Schumann[4,91,94*] and Paul Elliott[1,3,11,92,93,94*]

Excessive alcohol consumption is one of the main causes of death and disability worldwide. Alcohol consumption is a heritable complex trait. Here we conducted a meta-analysis of genome-wide association studies of alcohol consumption (g d^{-1}) from the UK Biobank, the Alcohol Genome-Wide Consortium and the Cohorts for Heart and Aging Research in Genomic Epidemiology Plus consortia, collecting data from 480,842 people of European descent to decipher the genetic architecture of alcohol intake. We identified 46 new common loci and investigated their potential functional importance using magnetic resonance imaging data and gene expression studies. We identify genetic pathways associated with alcohol consumption and suggest genetic mechanisms that are shared with neuropsychiatric disorders such as schizophrenia.

Excessive alcohol consumption is a major public health problem that is responsible for 2.2% and 6.8% of age-standardized deaths of women and men, respectively[1]. Most genetic studies of alcohol consumption focus on alcohol dependency, although the population burden of alcohol-related disease mainly reflects a broader range of behaviours associated with alcohol consumption[2]. Small reductions in alcohol consumption could have major public health benefits; even moderate decreases in alcohol consumption per day may have a substantial effect on rates of mortality[3].

Alcohol consumption is a heritable complex trait[4], but genetic studies to date have robustly identified only a small number of associated genetic variants[5–8]. These include variants in the aldehyde dehydrogenase (ADH) gene family—a group of enzymes that

A full list of affiliations appears at the end of the paper.

catalyse the oxidation of aldehydes[9]—including a cluster of genes on chromosome 4q23 (*ADH1B*, *ADH1C*, *ADH5*, *ADH6* and *ADH7*)[6].

Here we report a genome-wide association studies (GWAS) meta-analysis of alcohol intake (log-transformed alcohol intake ($g\,d^{-1}$)) among individuals of European ancestry drawn from the UK Biobank (UKB)[10], the Alcohol Genome-Wide Consortium (AlcGen) and the Cohorts for Heart and Aging Research in Genomic Epidemiology Plus (CHARGE+) consortia. In brief, UKB is a prospective cohort study comprising approximately 500,000 individuals recruited between the ages of 40 yr and 69 yr. Participants were asked to report their average weekly and monthly alcohol consumption through a self-completed touchscreen questionnaire[10]. On the basis of these reports, we calculated the alcohol intake ($g\,d^{-1}$; see Methods). Participants were genotyped using a customized array with imputation from the Haplotype Reference Consortium (HRC) panel[11], yielding approximately 7 million common single nucleotide polymorphisms (SNPs) with a minor allele frequency (MAF) of ≥1% and imputation quality INFO score of ≥0.1. After quality control (QC) and exclusions (see Methods), we performed a GWAS of alcohol consumption using data from 404,731 participants of European descent from the UKB under an additive genetic model (see Methods; Supplementary Table 1). We found that genomic inflation in the UKB analysis was $\lambda_{GC}=1.45$, but we did not adjust for inflation because the linkage disequilibrium score regression (LDSR) intercept was 1.05, indicating that the genomic inflation was due to polygenicity rather than to population stratification[12]. The estimated SNP-wide heritability of alcohol consumption in the data from the UKB was 0.09.

We also performed a GWAS on the basis of 25 independent studies from the AlcGen and CHARGE+ consortia, including 76,111 participants of European descent for whom alcohol consumption ($g\,d^{-1}$) could be calculated (Supplementary Table 2). Various arrays were used for genotyping, and imputation was performed using either the 1000 Genomes reference panel or the HRC platforms (Supplementary Table 3). After QC, we applied genomic control at the individual-study level and obtained summary results for approximately 7 million SNPs with an imputation quality score of ≥0.3 (see Methods).

We combined the UKB, AlcGen and CHARGE+ results using a fixed-effects inverse-variance-weighted approach for a total of 480,842 individuals[13]. To maximize power, we performed a single-stage analysis to test common SNPs with a MAF of ≥1%. We set a stringent *P*-value threshold of $P<5\times10^{-9}$ to denote significance in the combined meta-analysis[14] and required signals to be at $P<5\times10^{-7}$ in UKB—with the same direction of effect in UKB and AlcGen plus CHARGE+—to minimize false-positive findings. We excluded SNPs within 500 kb of variants reported as genome-wide significant in previous GWAS of alcohol consumption[5,6], identified new loci by requiring SNPs to be independent of each other (linkage disequilibrium (LD; $r^2<0.1$) and selected the sentinel SNP within each locus according to lowest *P* value (see Methods).

We then tested for correlations of alcohol-associated SNPs with imaging phenotypes of brain, heart and liver (as measured by MRI), and gene expression. We tested the sentinel SNPs for association with other traits and diseases and used *Drosophila* mutant models to investigate functional effects of mutations on behaviours induced by ethanol.

Results

Our meta-analysis identified 46 new loci associated with alcohol consumption (log-transformed alcohol intake ($g\,d^{-1}$); Fig. 1, Table 1). All inferential statistics for the new loci are reported in Table 1, whereas heterogeneity metrics are presented in Supplementary Table 4. We also discovered a further eight variants in the combined analysis at nominal genome-wide significance ($P<1\times10^{-8}$) that may be associated with alcohol intake (Supplementary Table 5). The most

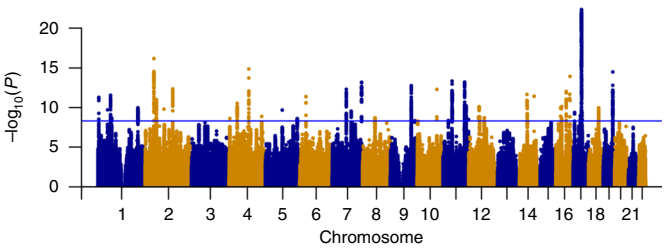

Fig. 1 | Results of the discovery genome-wide association meta-analysis with alcohol consumption. Manhattan plot showing *P* values from discovery genome-wide association meta-analysis with alcohol consumption (log-transformed alcohol intake, in g d^{-1}) among 480,842 individuals across UKB, AlcGen and CHARGE+, excluding known variants. The *P* value 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. The horizontal blue line represents the threshold of $P=5\times10^{-9}$.

significantly associated variant—rs1991556 ($P=4.5\times10^{-23}$)—is an intronic variant in *MAPT*, which encodes the microtubule-associated protein tau, that was found through PhenoScanner to be associated with not only dementia[15] and Parkinson's disease[16,17], but also with neuroticism, schizophrenia[18] and other traits[19–21] (see Methods; Fig. 2, Supplementary Table 6). The second most significantly associated variant, rs1004787 ($P=6.7\times10^{-17}$), is located near *SIX3*, which encodes a member of the sine oculis homeobox transcription factor family involved in the development of the eyes[22]. The third most significantly associated SNP, rs13107325 ($P=1.3\times10^{-15}$), is a missense SNP in *SLC39A8* (NCBI Gene 64116)—which encodes a member of the SLC39 family of metal-ion transporters—that has been associated with schizophrenia[23] as well as inflammatory bowel disease, cardiovascular and metabolic phenotypes[24–27] in previous GWAS (Fig. 2, Supplementary Table 6).

Another of our most-significant variants—an intronic SNP rs7121986 ($P=6.2\times10^{-14}$) in *DRD2* (NCBI Gene 1813)—is located in the gene that encodes the dopamine receptor D2 that has been associated with cocaine addiction, neuroticism and schizophrenia[18]. We also found significant associations with SNP rs988748 ($P=4.4\times10^{-9}$) in the *BDNF* gene (NCBI Gene 627), which encodes a member of the nerve growth factor family of proteins, and rs7517344, which is located near *ELAVL4* (NCBI Gene 1996; $P=2.0\times10^{-10}$), the product of which is involved in the regulation of *BDNF*[28]. Previous studies have suggested that a variant of *BDNF* is associated with alcohol consumption and that alcohol consumption modulates the expression of *BDNF*[29].

We also found that alcohol consumption is associated with SNP rs838145 ($P=3.2\times10^{-15}$), which has been associated with macronutrient intake in a previous GWAS[30]. This variant is nearest to *IZUMO* (NCBI Gene 284359) in a locus of around 50 kb that spans a number of genes including *FGF21* (NCBI Gene 26291), the gene product of which, FGF21, is a liver hormone that is involved in the regulation of alcohol preference, glucose and lipid metabolism[31]. We previously reported a significant association of alcohol intake with SNP rs11940694 in *KLB* (NCBI Gene 152831)—an obligate receptor of FGF21 in the brain[5]—and we strongly replicated that finding here ($P=3.3\times10^{-68}$).

As well as variants in *KLB* and in the alcohol dehydrogenase locus (smallest $P=1.2\times10^{-125}$), we found support ($P=1\times10^{-5}$) for the association of common variants in three other loci related to alcohol intake that were previously reported in GWAS (Supplementary Table 7), including SNP rs6943555 in *AUTS2* (NCBI Gene 26053; $P=2.9\times10^{-6}$). We also found that a new SNP related to alcohol intake, rs1421085 in *FTO* (NCBI Gene 79068), is in high LD ($r^2=0.92$) with a variant that has been previously reported to have a genome-wide significant association with alcohol dependence[32].

Table 1 | Association results of 46 new alcohol variants identified through the meta-analysis of UKB and AlcGen and CHARGE+

Lead SNP						Combined			UKB			AlcGen and CHARGE+		
Nearest gene	Annotated gene	Lead SNP rsID	CP	EA	EAF	β	s.e.	P	β	s.e.	P	β	s.e.	P
MAPT	*STH*	rs1991556	17:44083402	A	0.22	−0.012	0.001	4.5×10^{-23}	−0.013	0.001	2.4×10^{-21}	−0.011	0.004	4.0×10^{-3}
RP11-89K21.1	*SIX3*	rs1004787	2:45159091	A	0.54	0.009	0.001	6.7×10^{-17}	0.009	0.001	1.1×10^{-15}	0.007	0.003	1.4×10^{-2}
SLC39A8	*SLC39A8*	rs13107325	4:103188709	T	0.07	−0.016	0.002	1.3×10^{-15}	−0.017	0.002	4.8×10^{-16}	−0.006	0.006	3.6×10^{-1}
IZUMO1, RASIP1, FUT1	*IZUMO1*	rs838145	19:49248730	A	0.55	−0.008	0.001	3.2×10^{-15}	−0.009	0.001	2.4×10^{-15}	−0.004	0.003	1.7×10^{-1}
NA	*PSMD7*	rs1104608	16:73912588	C	0.43	−0.008	0.001	1.2×10^{-14}	−0.009	0.001	4.9×10^{-15}	−0.003	0.003	2.5×10^{-1}
MYBPC3	*MYBPC3*	rs2071305	11:47370957	A	0.69	0.009	0.001	4.5×10^{-14}	0.009	0.001	3.9×10^{-13}	0.007	0.003	3.1×10^{-2}
NA	*DRD2*	rs7121986	11:113355444	T	0.37	−0.008	0.001	6.2×10^{-14}	−0.008	0.001	1.3×10^{-13}	−0.005	0.003	1.1×10^{-1}
NA	*DPP6*	rs6969458	7:153489725	A	0.47	0.008	0.001	6.4×10^{-14}	0.008	0.001	1.3×10^{-12}	0.007	0.003	1.5×10^{-2}
RP11-308N19.1	*ZNF462*	rs74424378	9:109331094	T	0.76	0.009	0.001	1.7×10^{-13}	0.009	0.001	4.5×10^{-13}	0.006	0.003	8.4×10^{-2}
ARHGAP15, AC096558.1, *RP11*-570L15.2	*ARHGAP15*	rs13024996	2:144225215	A	0.37	−0.008	0.001	4.4×10^{-13}	−0.008	0.001	6.6×10^{-13}	−0.004	0.003	1.4×10^{-1}
MLXIPL	*MLXIPL*	rs34060476	7:73037956	A	0.87	−0.011	0.002	5.0×10^{-13}	−0.012	0.002	1.4×10^{-13}	−0.004	0.004	4.1×10^{-1}
NA	*FAM178A*	rs61873510	10:102626510	T	0.33	−0.008	0.001	5.1×10^{-13}	−0.008	0.001	9.8×10^{-12}	−0.008	0.003	1.7×10^{-2}
FTO	*FTO*	rs1421085	16:53800954	T	0.60	0.008	0.001	9.2×10^{-13}	0.007	0.001	1.7×10^{-10}	0.010	0.003	9.2×10^{-4}
NA	*PMFBP1*	rs11648570	16:72356964	T	0.89	−0.012	0.002	2.1×10^{-12}	−0.011	0.002	1.5×10^{-10}	−0.013	0.005	3.4×10^{-3}
OTX2, *RP11*-1085N6.6	*OTX2*	rs2277499	14:57271127	T	0.34	−0.008	0.001	2.2×10^{-12}	−0.007	0.001	2.4×10^{-9}	−0.012	0.003	9.1×10^{-5}
PDE4B	*PDE4B*	rs2310752	1:66392405	A	0.43	−0.007	0.001	2.8×10^{-12}	−0.008	0.001	1.8×10^{-11}	−0.006	0.003	4.2×10^{-2}
SERPINA1	*SERPINA1*	rs112635299	14:94838142	T	0.02	−0.025	0.004	3.7×10^{-12}	−0.027	0.004	9.8×10^{-12}	−0.017	0.010	9.9×10^{-2}
NA	*AJAP1*	rs780569	1:4569436	A	0.71	−0.008	0.001	5.2×10^{-12}	−0.008	0.001	1.1×10^{-11}	−0.005	0.003	1.2×10^{-1}
NA	*VRK2*	rs10496076	2:57942987	T	0.37	−0.007	0.001	9.7×10^{-12}	−0.007	0.001	1.3×10^{-9}	−0.009	0.003	1.6×10^{-3}
ACTR10, C14orf37	*ACTR10*	rs71414193	14:58685301	A	0.19	−0.009	0.001	1.8×10^{-11}	−0.008	0.001	5.8×10^{-9}	−0.013	0.004	4.5×10^{-4}
BEND4	*BEND4*	rs16854020	4:42117559	A	0.13	0.010	0.002	2.9×10^{-11}	0.010	0.002	5.8×10^{-9}	0.016	0.005	6.4×10^{-4}
NA	*SORL1*	rs485425	11:121544984	C	0.45	−0.007	0.001	6.1×10^{-11}	−0.007	0.001	7.3×10^{-11}	−0.004	0.003	1.9×10^{-1}
SEZ6L2	*SEZ6L2*	rs113443718	16:29892184	A	0.31	−0.007	0.001	7.4×10^{-11}	−0.008	0.001	4.5×10^{-11}	−0.003	0.003	2.9×10^{-1}
CBX5, *RP11*-968A15.2	*CBX5*	rs57281063	12:54660427	A	0.41	0.007	0.001	7.9×10^{-11}	0.007	0.001	1.8×10^{-9}	0.007	0.003	1.2×10^{-2}
NA	*TNRC6A*	rs72768626	16:24693048	A	0.94	0.014	0.002	9.7×10^{-11}	0.015	0.002	1.7×10^{-9}	0.014	0.006	1.8×10^{-2}
SYT14	*SYT14*	rs227179	1:210216731	A	0.59	−0.007	0.001	1.1×10^{-10}	−0.007	0.001	1.4×10^{-9}	−0.006	0.003	2.8×10^{-2}
TCF4	*TCF4*	rs9320010	18:53053897	A	0.60	0.007	0.001	1.1×10^{-10}	0.007	0.001	1.6×10^{-9}	0.007	0.003	2.2×10^{-2}
SBK1	*NPIPB6*	rs2726034	16:28336882	T	0.68	0.007	0.001	1.4×10^{-10}	0.007	0.001	1.1×10^{-9}	0.006	0.003	4.7×10^{-2}
ANKRD36	*ANKRD36*	rs13390019	2:97797680	T	0.87	0.010	0.002	1.6×10^{-10}	0.011	0.002	7.0×10^{-11}	0.004	0.005	4.5×10^{-1}
NA	*ELAVL4*	rs7517344	1:50711961	A	0.17	0.009	0.001	1.9×10^{-10}	0.008	0.001	2.5×10^{-7}	0.016	0.004	2.1×10^{-5}
LINC00461	*MEF2C*	rs4916723	5:87854395	A	0.58	0.007	0.001	2.1×10^{-10}	0.007	0.001	5.1×10^{-10}	0.005	0.003	1.1×10^{-1}
ARPC1B, ARPC1A	*ARPC1B*	rs10249167	7:98980879	A	0.87	0.010	0.002	2.9×10^{-10}	0.009	0.002	8.1×10^{-8}	0.015	0.004	3.8×10^{-4}
EFNB3, WRAP53	*EFNB3*	rs7640	17:7606722	C	0.80	0.008	0.001	4.3×10^{-10}	0.009	0.001	1.3×10^{-9}	0.006	0.004	9.9×10^{-2}

Continued

Table 1 | Association results of 46 new alcohol variants identified through the meta-analysis of UKB and AlcGen and CHARGE+ (continued)

Lead SNP						Combined			UKB			AlcGen and CHARGE+		
Nearest gene	Annotated gene	Lead SNP rsID	CP	EA	EAF	β	s.e.	P	β	s.e.	P	β	s.e.	P
RP11-501C14.5	*IGF2BP1*	rs4794015	17:47067826	A	0.41	0.007	0.001	4.3×10^{-10}	0.006	0.001	5.4×10^{-8}	0.009	0.003	1.2×10^{-3}
TCAP, PNMT, STARD3	*TCAP*	rs1053651	17:37822311	A	0.27	−0.007	0.001	1.1×10^{-9}	−0.008	0.001	8.4×10^{-10}	−0.003	0.003	2.8×10^{-1}
NA	*AADAT*	rs7698119	4:171070910	A	0.49	−0.006	0.001	1.3×10^{-9}	−0.006	0.001	1.6×10^{-7}	−0.009	0.003	1.6×10^{-3}
STAT6, AC023237.1	*STAT6*	rs12312693	12:57511734	T	0.55	−0.006	0.001	1.5×10^{-9}	−0.006	0.001	9.5×10^{-9}	−0.005	0.003	5.6×10^{-2}
SCN8A	*SCN8A*	rs7958704	12:51984349	T	0.41	−0.006	0.001	1.6×10^{-9}	−0.006	0.001	1.7×10^{-8}	−0.006	0.003	3.5×10^{-2}
ACSS3	*ACSS3*	rs11114787	12:81595700	T	0.27	0.007	0.001	2.0×10^{-9}	0.007	0.001	2.7×10^{-8}	0.007	0.003	2.4×10^{-2}
RP11-32K4.1	*BHLHE22*	rs2356369	8:64956882	T	0.52	−0.006	0.001	2.0×10^{-9}	−0.006	0.001	4.1×10^{-8}	−0.007	0.003	1.6×10^{-2}
ZRANB2-AS2	*ZRANB2*	rs12031875	1:71585097	A	0.82	−0.008	0.001	2.2×10^{-9}	−0.008	0.001	7.6×10^{-8}	−0.010	0.004	8.7×10^{-3}
MSANTD1, HTT	*MSANTD1*	rs12646808	4:3249828	T	0.66	0.007	0.001	2.4×10^{-9}	0.007	0.001	1.1×10^{-9}	0.002	0.003	4.7×10^{-1}
TENM2	*TENM2*	rs10078588	5:166816176	A	0.52	0.006	0.001	2.5×10^{-9}	0.006	0.001	4.3×10^{-8}	0.007	0.003	1.9×10^{-2}
IGSF9B	*IGSF9B*	rs748919	11:133783232	T	0.79	0.008	0.001	3.3×10^{-9}	0.008	0.001	1.0×10^{-8}	0.005	0.003	1.1×10^{-1}
AC010967.2	*GPR75-ASB3*	rs785293	2:53023304	A	0.57	−0.006	0.001	3.3×10^{-9}	−0.006	0.001	3.2×10^{-8}	−0.006	0.003	3.8×10^{-2}
BDNF, RP11-587D21.4	*BDNF*	rs988748	11:27724745	C	0.21	−0.008	0.001	4.4×10^{-9}	−0.007	0.001	1.2×10^{-7}	−0.010	0.004	8.3×10^{-3}

Results are ordered by *P* values of the combined analysis. CP, chromosome and position (build hg19/37); EA, effect allele of the discovered SNP; EAF, frequency of the effect allele.

Conditional analysis using Genome-wide Complex Trait Analysis (GCTA) software did not reveal any independent secondary signals related to alcohol consumption. Among approximately 14,000 individuals in the independent Airwave cohort[33] (see Methods), 7% of the variance in alcohol consumption was explained by the new and known common variants. Using weights from our analysis, we constructed an unbiased weighted genetic risk score (GRS) in the Airwave cohort (see Methods) and found a strong association of the new and known variants with levels of alcohol consumption ($P=2.75\times10^{-14}$), with a mean difference in sex-adjusted alcohol intake of 2.6 g d^{-1} when comparing the top and the bottom quintile of the GRS (Supplementary Table 8).

Associations with MRI phenotypes. We functionally characterized new variants by carrying out single-SNP analyses of the imaging phenotypes in the UKB (see Methods), focusing on brain ($n=9{,}702$), heart ($n=10{,}706$) and liver ($n=8{,}479$).

Using Bonferroni correction (corrected $P=6.6\times10^{-6}$, corresponding to 0.05/46 SNPs × 164 imaging phenotypes), we found significant positive associations between SNP rs13107325 in *SLC39A8* and the volumes of multiple brain regions; all inferential statistics for these associations are reported in Supplementary Table 9. The strongest associations were found for the putamen (left: $P=2.5\times10^{-45}$; right: $P=2.8\times10^{-47}$), the ventral striatum (left: $P=9.5\times10^{-53}$; right: $P=9.6\times10^{-51}$) and the cerebellum (strongest association for left I–IV volume, $P=1.2\times10^{-9}$; Supplementary Table 9). Similar findings were recently reported in a GWAS using brain imaging in UKB[34]. We also showed a significant association between rs1991556 and the parahippocampal gyrus ($P=1.2\times10^{-6}$).

We then tested these brain regions for association with alcohol consumption and found a significant effect for the left ($t_{8,601}=-3.7$, effect size (β) ± s.e. = −0.0019 ± 0.0005, $P=2.0\times10^{-4}$) and right ($t_{8,601}=-3.65$, $\beta\pm$s.e. = −0.0070 ± 0.0005, $P=2.6\times10^{-4}$) putamen. Finally, we used data from $n=8{,}610$ individuals and performed a mediation analysis using a standard three-variable path model, bootstrapping 10,000 times, to calculate the significance of the mediation effect of putamen volume for genetic influences on alcohol consumption (see Methods). We found evidence that the effect of SNP rs13107325 in *SLC39A8* on alcohol intake is partially mediated by its association with the volume of the left ($t_{8,601}=-3.03$, $\beta\pm$s.e. = −0.27 ± 0.09, $P=1.9\times10^{-3}$) and right ($t_{8,601}=-2.82$, $\beta\pm$s.e. = −0.27 ± 0.09, $P=1.7\times10^{-3}$) putamen (Fig. 3, Supplementary Table 10). To exclude the possibility of an inverse causal pathway, we performed additional analyses in UKB non-drinkers ($n=589$). With 10,000 random permutations, associations of rs13107325 with both left and right putamen remained significant (left putamen: $t_{541}=1.06$, $P=0.02$; right putamen: $t_{541}=0.38$, $P=0.04$), indicating that the association between rs13107325 and putamen regions is not mediated by alcohol intake.

We did not find any significant associations between the new SNPs and either cardiac (left ventricular mass or end diastolic volume or right ventricular end diastolic volume; Supplementary Table 11) or liver fat measurements from MRI analysis (Supplementary Table 12) after adjustment for multiple testing.

Effects of SNPs on gene expression. We carried out expression quantitative trait loci (eQTL) analyses using the Genotype-Tissue Expression (GTEx) and the UK Brain Expression Consortium (UKBEC) datasets. We found that 34 of the 53 new and known SNPs associated with alcohol consumption have a significant effect on gene expression in at least one tissue type, including 33 SNPs that affect gene expression in the brain (Supplementary Tables 13 and 14, Supplementary Figs. 1–3). We found that the most significant eQTLs often do not involve the nearest gene and that several of the SNPs affect expression of different genes in different tissues. For example, SNP rs1991556 in *MAPT* (NCBI Gene 4137) affects expression of 33

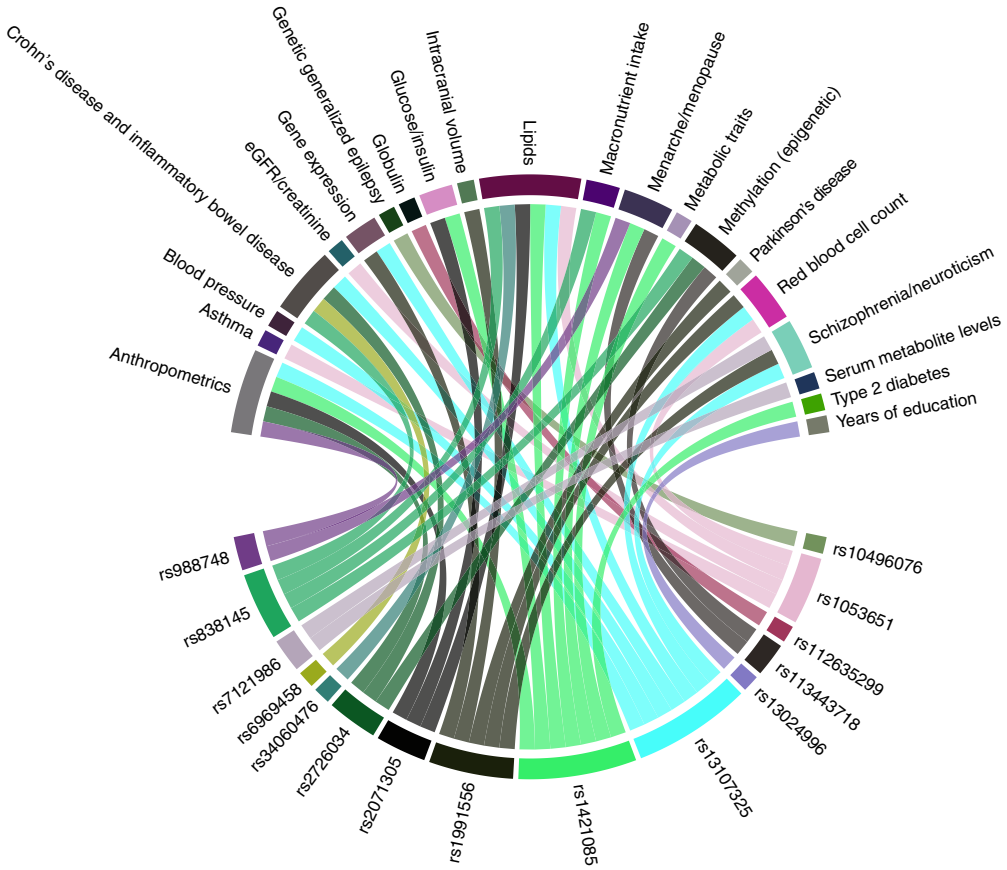

Fig. 2 | Association of alcohol intake loci with other traits. Associations with other traits were identified using the PhenoScanner database for the 46 new sentinel SNPs, including proxies in LD ($r^2 \geq 0.8$) with genome-wide significant associations. Each coloured line connects a specific variant with the associated traits and diseases. eGFR, estimated glomerural filtration rate.

genes overall—with the most significant effects on the expression of the non-protein-coding genes *CRHR1-IT1 (*also known as *C17orf69* or *LINC02210*; NCBI Gene 147081) and *LRRC37A4P* (NCBI Gene LRRC37A4P)—that are located near *MAPT* and expressed in a wide range of tissues including the brain, adipose tissue and skin ($P = 7.2 \times 10^{-126}$ to $P = 2.5 \times 10^{-6}$; Supplementary Fig. 2). Similarly, the A allele at SNP rs2071305 within *MYBPC3* (NCBI Gene 4607) affects the expression of several genes and is most significantly associated with increased expression of *C1QTNF4* (NCBI Gene 114900) across several tissues ($P = 1.9 \times 10^{-25}$ to $P = 8.4 \times 10^{-5}$).

Several of these eQTLs were found to affect expression of genes known to be involved in reward and addiction. SNP rs1053651 in the *TCAP–PNMT–STARD3* gene cluster affects expression of *PPP1R1B* (also known as *DARPP-32*; NCBI Gene 84152), which encodes a protein that mediates the effects of dopamine in the mesolimbic reward pathway[35]. Other known addiction-related genes include *ANKK1* (NCBI Gene 255239) and *DRD2* (expression affected by SNP rs7121986), genes implicated in alcohol and nicotine dependence[36,37], *CRHR1* (NCBI Gene 1394; affected by SNP rs1991556), which is involved in stress-mediated alcohol dependence[38,39], and *PPM1G* (SNP rs1260326; NCBI Gene 5496), epigenetic modification of which was previously reported to be associated with alcohol abuse[40].

Over-representation enrichment analyses on the basis of functional annotations and disease-related terms indicated that genes of which the expression is affected by the identified eQTLs are most significantly enriched for terms related to abdominal ($n = 91$) and other malignant cancers, motor function ($n = 5$) and cellular homeostasis ($n = 22$; Supplementary Fig. 4). We performed a gene-based analysis and repeated the over-representation enrichment analysis, adding the new set of identified genes (Supplementary Table 15). The results were similar, supporting an enrichment for abdominal ($n = 100$) and other cancers, as well as motor function ($n = 5$) and cellular homeostasis ($n = 24$; Supplementary Fig. 5).

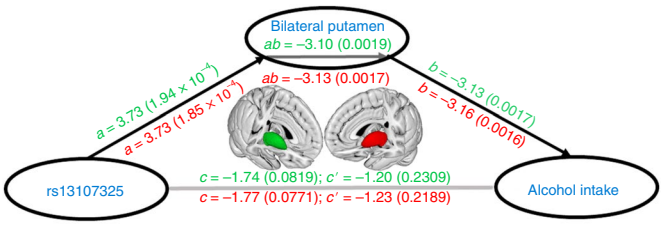

Fig. 3 | Mediation effect of the grey-matter volume of bilateral putamen on the relationship between SNP rs13107325 and alcohol intake. The left putamen is indicated in green and the right putamen is indicated in red. We use *a* for the relationship between rs13107325 and putamen, *b* for the relationship between putamen and alcohol consumption, *c* for the relationship between rs13107325 and alcohol consumption, *c′* for the relationship between rs13107325 and alcohol consumption after excluding the effect of putamen, and *ab* as the mediation effect. The significance tests are based on the bootstrapping method (10,000 times). *z*-statistics and the corresponding *P* values are provided in parentheses. The brain icon was created using Mango software v4.1 (http://ric.uthscsa.edu/mango/).

Other traits and diseases. Using LDSR[12], we assessed genetic correlations between alcohol consumption and 235 complex traits and

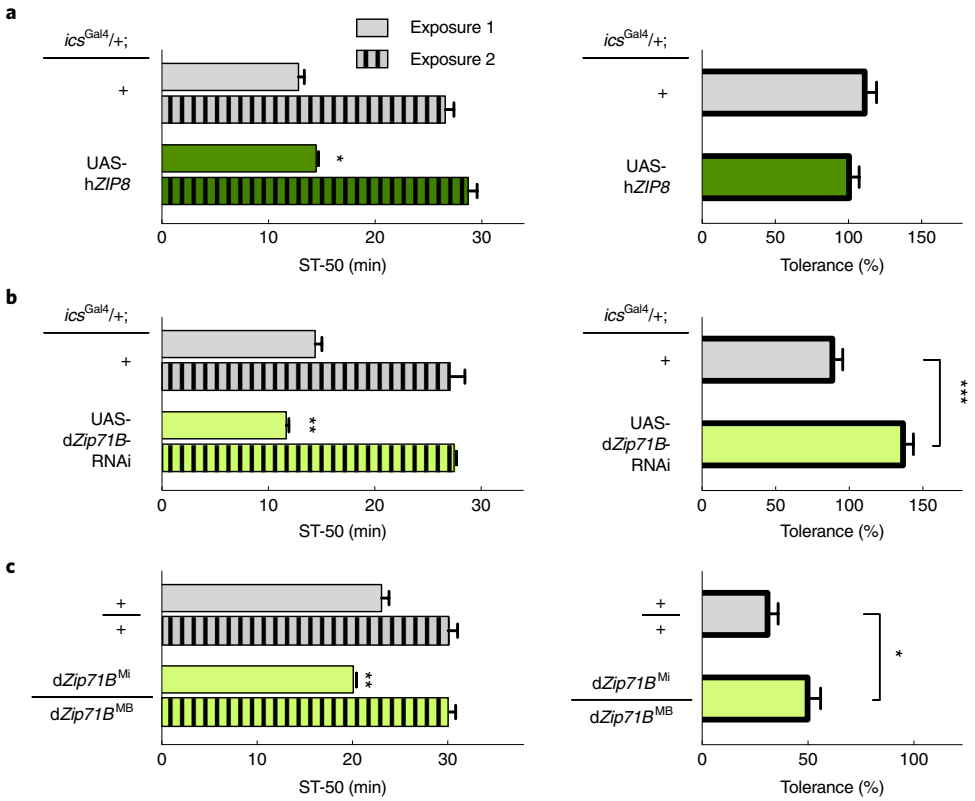

Fig. 4 | Comparison of *ZIP8* alcohol phenotypes in *Drosophila*. a–c, Flies were exposed to 100/50 (flowrates) ethanol/air vapour for 30 min for exposure 1, and the time to 50% loss-of-righting was determined (ST-50, sedation time). After recovery on food for 4 h, flies were re-exposed to the same vapours, and the second ST-50 was recorded (left). The resulting increase in ST-50—that is, tolerance—is shown on the right. *Drosophila* overexpressing human h*ZIP8* in *ics*-expressing cells (**a**), *Drosophila* with knockdown of the fly orthologue d*Zip71B* (**b**) and *Drosophila* carrying two transposon insertions in the endogenous d*Zip71B* gene (**c**) were compared with control flies. ***$P<0.001$, **$P<0.01$, *$P<0.05$; exact *P* values are presented in the text.

diseases from publicly available GWAS summary statistics (see Methods). All results—including their statistics (that is, r_g, s.e., *z* value and *P* value)—are included in Supplementary Table 16. The strongest positive genetic correlations on the basis of false-discovery rate (FDR)-adjusted values of $P<0.02$ were found for smoking ($r_g=0.42$, $P=1.0\times10^{-23}$) and levels of high-density lipoprotein (HDL) cholesterol ($r_g=0.26$, $P=5.1\times10^{-13}$). We also found negative correlations for sleep duration ($r_g=-0.14$, $P=3.8\times10^{-7}$) and fasting insulin levels ($r_g=-0.25$, $P=4.5\times10^{-6}$). A significant genetic correlation was also found with schizophrenia ($r_g=0.07$, $P=3.9\times10^{-3}$) and bipolar disorder ($r_g=0.15$, $P=5.0\times10^{-4}$; Supplementary Table 16). Over-representation enrichment analysis using WebGestalt[41] (http://www.webgestalt.org) showed that our list of new and known variants is significantly enriched for several diseases and traits, including developmental disorder in children ($P=7.3\times10^{-5}$), epilepsy ($P=1.4\times10^{-4}$), heroin dependence ($P=5.7\times10^{-4}$) and schizophrenia ($P=8.4\times10^{-4}$; Supplementary Fig. 6). Mendelian randomization analysis (see Methods) was performed to assess a potential causal effect of alcohol on schizophrenia risk using the inverse-variance weighted approach, but this analysis did not show significant results ($P=0.089$), showing large heterogeneity of the estimates of the tested variants.

Functional studies in *Drosophila*. On the basis of our GWAS and brain imaging findings, we performed further testing on the SNP rs13107325 in *SLC39A8* (also known as *ZIP8*) in *Drosophila*, making use of conserved mechanisms to modulate behaviours that are induced by ethanol[42,43]. First, we overexpressed human *ZIP8* using a Gal4 driver that included expression in neurons involved in multiple ethanol-induced behaviours[43]. Flies carrying ics^{Gal4}/+ UAS-*hZIP8*/+ showed a slight—albeit significant—resistance to sedation induced by ethanol compared with control flies ($t_{30}=2.3$, Hedge's $g=0.80$, 95% confidence interval (CI) = 0.08–1.53, $P=0.026$, $n=16$ per genotype). Ethanol tolerance, induced by repeated exposures that were separated by recovery for 4 h, was unchanged in these flies ($t=1.0$, $P=0.33$; Fig. 4a). Next, we used the same Gal4 driver to knock down the endogenous *Drosophila* orthologue of h*ZIP8*, d*Zip71B*. This caused the flies to display naive sensitivity to sedation induced by ethanol ($t_{14}=3.98$, Hedge's $g=-1.84$, 95% CI = −0.67 to −3.01, $P=0.0014$; $n=8$ per genotype) and, furthermore, these flies developed greater tolerance to ethanol after repeated exposure ($t_{14}=4.80$; Hedge's $g=2.29$, 95% CI = 1.03–3.55, $P=0.0003$; Fig. 4b). To corroborate this phenotype, we then tested flies that were transheterozygous for two independent transposon insertions in the middle of the d*Zip71B* gene (Supplementary Fig. 7) and found that these $dZip71B^{Mi/MB}$ flies also displayed naive sensitivity to sedation induced by ethanol ($t_{14}=3.23$, Hedge's $g=-1.54$, 95% CI = −0.42 to −2.65, $P=0.006$) and increased ethanol-induced tolerance ($t_{14}=2.39$, Hedge's $g=1.13$, 95% CI = 0.07–2.18, $P=0.032$) compared with the controls ($n=8$ each; Fig. 4c).

Discussion

Our discovery utilizing data on common variants from more than 480,000 people of European descent extends our knowledge of the genetic architecture of alcohol intake, increasing the number of identified loci to 46. We identified loci involved in neuropsychiatric

conditions such as schizophrenia, Parkinson's disease and dementia, as well as *BDNF*, gene expression of which is affected by alcohol abuse. Our findings illustrate that large-scale studies of genetic associations with alcohol intake in the general population—rather than studies on alcohol dependency alone—can provide further insights into the genetic mechanisms that regulate the consumption of alcohol.

We highlight the role of the highly pleiotropic *MAPT* and *SLC39A8* genes in the genetics of alcohol consumption. *MAPT* plays an important role in tau-associated dementia[44] and both *MAPT* and *SLC39A8* are implicated in other neuropsychiatric conditions, including neuroticism, schizophrenia and Parkinson's disease[16–18]. *SLC39A8* encodes a member of the SLC39 family of metal-ion transporters. SLC39A8 is glycosylated and found in the plasma membrane and mitochondria, and is involved in the cellular transport of zinc, the modulation of which could affect microglial inflammatory responses[45]. Our gain-of-function and loss-of-function studies in *Drosophila* indicate a potential causal role of *SLC39A8* in alcohol-consumption behaviour, although these results should be interpreted with caution owing to the small sample size in our experiment. Using MRI brain imaging, we demonstrate a significant association between SNP rs13107325 in *SLC39A8* and differences in the volume of the putamen, and that these structural differences seem to partially mediate associations between rs13107325 and alcohol consumption. The putamen has been associated with alcohol consumption and the withdrawal syndrome after chronic administration to rodents and non-human primates[46]. On the basis of the mediation analysis, we suggest a plausible causal pathway that links rs13107325 in *SLC39A8* with alcohol intake through an effect on putamen volume, but further studies are needed to conclusively demonstrate causal links. Differences in the volume of the putamen have also been associated with both schizophrenia and psychosis[47,48] and a robust association between SNP rs13107325 in *SLC39A8* and schizophrenia was reported in a previous GWAS[23].

We also report SNP rs7121986—which is located near *DRD2*—as a new variant associated with alcohol intake in GWAS. The gene product of *DRD2*, D2 dopamine receptor, is a G-protein-coupled receptor expressed on post-synaptic dopaminergic neurons that has long been implicated in alcoholism[49]. We also identify SNP rs988748 in *BDNF* as a new alcohol intake variant; BDNF expression is differentially affected by alcohol exposure in animal models[50,51]. *DRD2* and *BDNF* (along with *PPP1R1P*) are centrally involved in reward-mediating mesocorticolimbic pathways and are both implicated in the development of schizophrenia. For example, there is a robust GWAS association between schizophrenia and SNP rs4938021 in *DRD2* (in perfect LD with our new alcohol intake-related variant rs7121986) and *DRD2* seems to be pivotal in network analyses of genes involved in schizophrenia[52]. Taken together, our results suggest that there are shared genetic mechanisms between the regulation of alcohol intake and susceptibility to schizophrenia, as well as other neuropsychiatric disorders. Related to this, large prospective epidemiological studies report a threefold risk of schizophrenia in relation to alcohol abuse[53].

We previously reported genome-wide significant associations between alcohol intake and *KLB*, and identified a liver–brain axis that linked the liver hormone FGF21 with central regulation of alcohol intake, which involved the β-klotho receptor (the gene product of *KLB*) in the brain[5]. Here we identified a significant variant near *FGF21* and strongly replicate the previously reported *KLB* gene variant, strengthening the genetic evidence for the importance of this pathway in the regulation of alcohol consumption.

The LDSR analysis showed a positive genetic correlation between alcohol consumption, smoking and levels of HDL cholesterol. This confirms previous studies that reported an almost-identical genetic correlation between alcohol consumption and the number of cigarettes smoked per day[54]. Furthermore, the observed genetic correlation with HDL levels is consistent with previous observations of an association between alcohol consumption and HDL[55,56], including results of a Mendelian randomization study that suggested a possible causal role linking alcohol intake with increased HDL levels[57]. Moreover, we found an inverse genetic correlation between the duration of sleep and consumption of alcohol—an association that has been previously reported only in a few small epidemiological studies[58]. We also found a significant genetic correlation between schizophrenia and bipolar disorder, a result that is supported by a recently published trans-ethnic meta-analysis of case–control studies on alcohol dependence[59]. We could not test for a genetic association between alcohol and risk of alcohol-related cancers[60] owing to the limited availability of summary data. However, our gene-set enrichment analysis showed a significant enrichment for genes related to abdominal as well as other cancers.

The strengths of our study include its size, detailed attention to the alcohol phenotype, dense coverage of the genome through imputation, and incorporation of brain and other imaging data to explore potential mechanisms. More than 80% of the data came from UKB, which combines high-quality phenotypic data and imputed genome-wide genetic data with strict attention to QC[61]. We adopted a stringent approach to claim new variants, including a conservative *P*-value threshold, internal replication in UKB and consistent direction of effect with the other studies to minimize the reporting of false-positive signals.

However, because alcohol intake is socioculturally and genetically determined, it is influenced by other lifestyle and environmental factors that may modify or dilute the genetic signal. A key limitation is that assessment of alcohol intake relies on self-reporting, which is prone to errors and biases, including recall bias and systematic under-reporting by heavy drinkers[62,63]. Furthermore, questionnaires on alcohol intake covered a short duration (for example, a day or week) for a single period, which may not be representative of broader drinking patterns of cohort participants. We harmonized data across cohorts by converting alcohol intake into a common metric of $g\,d^{-1}$, with imputation where necessary in UKB for participants that reported consumption of small amounts of alcohol. Taking this approach, we were able to detect strong genetic associations with alcohol intake that explained 7% of the variance in alcohol consumption in an independent cohort, whereas our GRS analysis indicates that individuals in the lower fifth of the GRS distribution were consuming approximately one third of a standard drink ($2.6\,g\,d^{-1}$ alcohol) less each day compared with those in the upper fifth.

We should also point out that our eQTL analyses are a first step in the identification of causal genes. Yet, as the most significant eQTLs affect expression of many genes—which are not necessarily the nearest genes—there is a need to further prioritize potential causal genes. Unbiased strategies that leverage information from multiple datasets, including extensive genomic annotations and high-throughput functional screening in a broad range of tissues, will be essential to effectively prioritize genes and uncover underlying causal mechanisms[64]. Establishing confidence in the prioritized genes in such a way is a prerequisite for performing functional follow-up studies in appropriate model systems, as demonstrated by the identification of the causal genes and potential disease mechanisms at the obesity-associated *FTO* locus[65].

In summary, in this large study of genetic associations with alcohol consumption, we identified common variants in 46 new loci. Several of these variants are located in genes that are expressed in the brain as well as other tissues. Our findings suggest that there may be shared genetic mechanisms that underpin the regulation of alcohol intake and development of neuropsychiatric disorders, including schizophrenia. This may form the basis for a greater understanding of observed associations between alcohol consumption, schizophrenia[66] and other disorders.

Methods

UKB data. We conducted a GWAS analysis using data from 458,577 UKB participants of European descent who were identified on the basis of a combination of self-reported and genetic data. The details of the selection of participants have been described elsewhere[14]. These participants comprised 408,951 individuals from the UKB, genotyped at 825,927 variants using a custom Affymetrix UKB Axiom Array chip, and a further 49,626 individuals from the UK BiLEVE study—which is a subset of UKB—genotyped at 807,411 variants using a custom Affymetrix UK BiLEVE Axiom Array chip. For our analyses, we used SNPs imputed centrally by UKB using the HRC panel.

Alcohol intake. We calculated alcohol intake as $g\,d^{-1}$ on the basis of self-reported levels of alcohol consumption that were provided through a touch-screen questionnaire. The quantity of each type of drink (red wine, white wine, beer or cider, fortified wine or spirits) was multiplied by its standard drink size and reference alcohol content. Drink-specific intake during the reported drinking period (per week for frequent drinkers, defined as: daily or almost daily, once or twice a week, three or four times a week, or per month for occasional drinkers, defined as: one to three times a month or special occasions only) was summed and converted to $g\,d^{-1}$ alcohol intake for all participants who provided a complete response to the quantitative drinking questions. The alcohol intake for participants who submitted incomplete responses was imputed by bootstrap resampling from the complete responses, stratified by drinking frequency (occasional or frequent) and sex.

Participants were defined as life-time non-drinkers if they reported 'never' on the question on alcohol drinking frequency (UKB field 1558) and 'no' for the question on former drinker (UKB field 3731); these participants were excluded from further analysis. We considered participants with alcohol consumption of more than $500\,g\,d^{-1}$ to be outliers and they were excluded from further analysis. We also excluded participants with missing covariates, leaving data on 404,732 individuals. The $\log_{10}$-transformed alcohol and sex-specific residuals were derived from the regression of $\log_{10}$-transformed alcohol intake on age, age^2, genotyping chip and weight.

UKB genetic analysis. We performed linear mixed modelling with BOLT-LMM software[67] using an additive genetic model to determine associations between measured and imputed SNPs and alcohol consumption (sex-specific residuals of the $\log_{10}$-transformed alcohol intake variable). Model building was based on SNPs with MAF > 5%, call rate > 98.5% and Hardy–Weinberg equilibrium $P > 1\times10^{-6}$. SNPs were imputed using the HRC panel with an imputation quality INFO score of ≥0.1. We estimated the LDSR intercept to assess the degree of genomic inflation beyond polygenicity as well as the genomic inflation factor[68] λ_{GC}.

AlcGen and CHARGE+ consortia. We analysed GWAS data that were available from 25 independent studies (n = 76,111) conducted by the AlcGen and the CHARGE+ consortia. All study participants were of reported European ancestry and data were imputed to either the 1000 Genomes Project or the HRC panel. Alcohol intake in $g\,d^{-1}$ was computed and the $\log_{10}$-transformed residuals were analysed as described above. Study names, cohort information and general study methods are included in Supplementary Tables 2 and 3.

All studies were subject to central QC using easyQC[69], including filtering for MAFs. Finally, we analysed data on approximately 7.1 million SNPs with MAF > 1% and imputation quality score (Impute (INFO score) or Mach (r^2)) ≥ 0.3. Genomic control (GC) was applied at the study level. We generated the available GWAS using a fixed-effects inverse-variance weighted meta-analysis and summary estimates were derived for AlcGen and CHARGE+.

One-stage meta-analysis. We performed a one-stage meta-analysis, in which we applied a fixed-effects inverse-variance weighted meta-analysis using METAL[70] to obtain summary results from the UKB and AlcGen plus CHARGE+ GWAS for up to 480,842 participants and around 7.1 million SNPs with MAF ≥ 1% for variants present in both the UKB data and AlcGen and CHARGE+ meta-analysis. We assessed the observed heterogeneity using Cochran's Q and we quantified this using the I^2 metric. We considered a Cochran's Q $P < 1\times10^{-4}$ to be significant. The LDSR intercept (s.e.) in the discovery meta-analysis was 1.05 and no further correction was applied. Quantile–quantile plots of the combined meta-analysis summary results, UKB only, and AlcGen and CHARGE+ only are presented in Supplementary Fig. 8.

Previously reported SNPs. We searched the GWAS Catalog (https://www.ebi.ac.uk/gwas/) and identified 17 SNPs associated with alcohol consumption at genome-wide level of significance ($P < 5\times10^{-8}$). We enhanced the list with reference to a recent GWAS by Clarke et al.[6] that was not covered by the GWAS Catalog at the time of the analysis, reporting 14 additional rare and common SNPs. Together with a SNP in *RASGRF2*, which has been shown to be associated with alcohol-induced reinforcement[71], we found 31 previously reported alcohol-consumption-related SNPs.

New loci. Using the locus definition of SNPs that are within a distance of ±500 kb of each other and SNPs that were in LD ($r^2 > 0.1$) as calculated using PLINK, we augmented the list of known SNPs with all SNPs present within our data that were not contained within the previously published loci. We further excluded SNPs in the human leukocyte antigen region (chromosome 6, 25–34 Mb) owing to its complex LD structure. Using PLINK, we then performed LD clumping on 4,515 unknown SNPs with $P < 1\times10^{-8}$ using an $r^2 > 0.1$ and a distance threshold of 500 kb. We further grouped the lead SNPs within 500 kb of each other into the same loci and selected the SNP with smallest P value from the locus as sentinel SNP.

We report a SNP as a new signal of association with alcohol consumption if the following criteria are fulfilled:

1. The sentinel SNP has $P < 5\times10^{-9}$ in the one-stage meta-analysis.
2. The sentinel SNP is strongly associated ($P < 5\times10^{-7}$) in the UKB GWAS alone.
3. The sentinel SNP has a concordant direction of effect between UKB and AlcGen and CHARGE+ datasets.
4. The sentinel SNP is not located within any of the previously reported loci.

We selected criteria (1)–(3) to minimize false-positive findings, including use of a conservative one-stage P-value threshold that was an order of magnitude more stringent than a genome-wide significance P value (for example, the threshold of $P < 5\times10^{-9}$ has been proposed for whole-genome sequencing-based studies). This approach allowed us to identify 46 sentinel SNPs in total. Regional plots for all 46 sentinel SNPs are presented in Supplementary Fig. 9.

Conditional analysis. We conducted locus-specific conditional analysis using the GCTA software (https://cnsgenomics.com/software/gcta). For each of the 46 new sentinel SNPs, we obtained conditional analysis results for the SNPs with a MAF of >1% and located within 500 kb of the sentinel SNP after conditioning on the sentinel SNP. The meta-analysis results of the GWAS in the UKB, AlcGen and CHARGE+ were used as input summary statistics and the individual-level genetic data from the UKB were used as the reference sample. Results for a SNP were considered to be conditionally significant if the difference between the conditional P value and the original P value was greater than 1.5-fold ($-\log_{10}[P]/-\log_{10}[P_{conditional}] > 1.5$) and the conditional P value was smaller than 5×10^{-8}.

Gene-based analysis. We performed a gene-based analysis using fastBAT, a method that performs a set-based association analysis using summary-level data from GWAS. We used the UKB dataset as a reference set for the LD calculations[72]. Gene-based associations with $P < 5\times10^{-9}$ were considered to be significant.

Gene expression analyses. To analyse the impact of genetic variants on the expression of neighbouring genes and identify eQTLs (*cis*-eQTLs; that is, SNPs associated with differences in local gene expression), we used two publicly available databases—the GTEx database[73] (https://www.gtexportal.org) and the UKBEC dataset[74] (http://www.braineac.org). We searched these databases for significant variant–transcript pairs for genes within 1 Mb of each input SNP.

Using the GTEx database, we tested for *cis*-eQTL effects in 48 tissues from 620 donors. The data described herein were obtained from the GTEx Portal v.7, and we used FastQTL[75] to map SNPs to gene-level expression data and calculate q values on the basis of β-distribution-adjusted empirical P values[76]. A FDR threshold of $P \leq 0.05$ was applied to identify genes with a significant eQTL. The effect size—defined as the slope of the linear regression—was computed in a normalized space (normalized effect size (NES)) in which magnitude has no direct biological interpretation. Here, NES reflects the effects of our GWAS effect alleles (that are not necessarily the alternative alleles relative to the reference alleles, as reported in the GTEx database). Supplementary Table 13 lists transcript–SNP associations with significant eQTL effects.

Using the UKBEC dataset that comprises 134 brains (http://www.braineac.org/), we searched for *cis*-eQTLs in 10 brain regions, including the cerebellar cortex (CRBL), frontal cortex (FCTX), hippocampus (HIPP), medulla (specifically inferior olivary nucleus; MEDU), occipital cortex (specifically primary visual cortex; OCTX), putamen (PUTM), substantia nigra (SNIG), thalamus (THAL), temporal cortex (TCTX) and intralobular white matter (WHMT), as well as across all brain tissues (aveALL). We used MatrixEQTL[77] to generate P values for each expression profile (at either the exon level or gene level) against the respective SNP for the 10 different brain regions and all brain tissues. Supplementary Table 14 lists transcript–SNP associations with a eQTL $P < 0.0045$ in at least one brain tissue. Subsequent data analysis was performed in R v3.5 (https://www.R-project.org/).

We performed over-representation enrichment analysis using a list of 146 GTEx eQTL genes that were derived from the single-variant analysis and a list of 160 eQTL genes that were derived from both single-variant and gene-based analysis. Ingenuity pathway analysis (QIAGEN) was performed on these lists using ontology annotations from all available databases, except those derived from low-confidence computational predictions.

MRI data. We used the most recent release of MRI data on brain, heart and liver from participants of the UKB to investigate genetic associations between the 46 new SNPs and alcohol consumption.

Brain imaging. *Brain MRI acquisition and preprocessing.* We used the T1 data from the UKB to elucidate volumetric brain structures, including the cortical

and the sub-cortical areas. The T1 data were acquired and preprocessed centrally by the UKB. The brain regions were defined by combining the Harvard–Oxford cortical and subcortical atlases[78] (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases) and the Diedrichsen cerebellar atlas[79] (http://www.diedrichsenlab.org/imaging/propatlas.htm). FAST (FMRIB's Automated Segmentation Tool)[80] was then used to estimate the grey-matter partial volume within each brain region. Subcortical region volumes were also modelled using FIRST (FMRIB's Integrated Registration and Segmentation Tool). More details about the MRI scanning protocol and preprocessing are provided in the UKB documentation (https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf).

Association analyses. We performed association analyses on $n = 9{,}702$ individuals between all new SNPs and the grey-matter volume of brain regions using Pearson correlation, adjusting for age, age^2, sex, age × sex, age^2 × sex and head size. Brain-volume features, log-transformed alcohol intake data (g d^{-1}) and the confounding factors were initially transformed using a rank-based inverse Gaussian transformation. Significance levels were set at $P < 0.05$ and adjusted using the FDR method for multiple comparisons.

Mediation analysis. To assess whether the effect of a SNP on alcohol consumption is mediated through a brain region, we performed a single-level mediation analysis on the basis of a standard three-variable path model (SNP–brain region–alcohol consumption) with corrected and accelerated percentile bootstrapping 10,000 times to calculate the significance of the mediation effect. We considered the grey-matter volume of brain regions that had a significant association on alcohol consumption to be a mediator variable. We calculated the significance of path *a*, path *b* and *ab* mediation (SNP–brain region–alcohol consumption) using a multilevel mediation and moderation (M3) toolbox[81,82]. To exclude the possibility of an inverse causal pathway, we performed further analyses in non-drinkers from the UKB ($n = 589$)—performing 10,000 random permutations for associations between rs13107325 and both left and right putamen.

Cardiac imaging. *Cardiac MRI acquisition and preprocessing.* Details of the cardiac image acquisition in the UKB were reported previously[83]. Cardiac MRI was acquired using a clinical wide bore 1.5 T scanner (MAGNETOM Aera, Syngo Platform VD13A, Siemens Healthcare) with 48 receiver channels, a 45 mT m^{-1} and 200 T m^{-1} s^{-1} gradient system, an 18-channel anterior body surface coil used in combination with 12 elements of an integrated 32 element spine coil and electrocardiogram gating for cardiac synchronization. A two-dimensional short-axis cardiac MRI was obtained using a balanced steady-state free precession to cover the entire left and right ventricle (echo time, 1.10 ms; repetition time, 2.6 ms; flip angle, 80°; slice thickness, 8 mm with 2 mm gap; typical field of view, 380 × 252 mm^2; matrix size, 208 × 187, acquisition of 1 slice per breath-hold).

The cardiac images were segmented to provide left ventricular mass (LVM), left end-diastolic volume (LVEDV), left end-systolic volume (LVESV), right end-diastolic volume (RVEDV) and right end-systolic volume (RVESV) using a fully convolutional network as described previously[84]. Left ventricular ejection fraction (LVEF) and right ventricular ejection fraction (RVEF) were derived from (LVEDV–LVESV)/LVEDV × 100 and (RVEDV–RVESV)/RVEDV × 100, respectively.

Association analyses. To test associations between cardiac MRI parameters and alcohol-consumption-related SNPs, we carried out a regression of LVM, LVEDV, LVEF, RVEDV and RVEF onto each of the 46 SNPs adjusting for age, sex, height, weight, hypertension (defined as systolic blood pressure > 140 mmHg and/or diastolic blood pressure > 90 mmHg or under antihypertensive treatment), diabetes and smoking history on $n = 10{,}706$ participants. Significance levels were set at $P < 0.05$ and adjusted using the FDR method for multiple comparisons.

Liver imaging. *Liver MRI acquisition and preprocessing.* Details of the liver image acquisition protocol have been reported previously[85]. In brief, all participants were scanned in a Siemens MAGNETOM Aera 1.5 T MRI scanner (Siemens Healthineers) using a 6-min dual-echo Dixon Vibe protocol, providing a water- and-fat-separated volumetric dataset for fat and muscle covering neck to knees. For the quantification of the liver proton density fat fraction (PDFF), an additional single multi-echo gradient slice was acquired from the liver. Liver images were analysed by computing specific regions of interest for water, fat and T2* using a magnitude-based chemical shift technique with a six-peak lipid model, correcting for T1 and T2*.

Association analyses. We performed association analyses between 46 SNPs related to alcohol consumption and liver PDFF (%), from 8,479 samples, using a linear regression model adjusted for age, age^2, sex, type 2 diabetes, body mass index, genotyping chip and the first three principal components. Liver PDDF was initially transformed using a rank-based inverse transformation. Significance levels were set at $P < 0.05$, adjusted using the FDR method for multiple comparisons.

***Drosophila* experiments.** Flies were kept on standard cornmeal/molasses fly food in a 12 h:12 h light:dark cycle at 25 °C. The following transgenic flies were obtained from the Bloomington *Drosophila* Stock Center: UAS-h*ZIP8*, BL 66125; UAS-d*ZIP71B*-TRiP-RNAi^{HMC04064}, BL 55376; d*Zip71B*^{MI13940}, BL 59234; and d*Zip71B*^{MB11703}, BL 29928. For behavioural experiments, crosses were set up such that experimental and control flies were sibling progeny from a cross, and both were therefore in the same hybrid genetic background (*w* Berlin/unknown). Flies aged 1–5 d of adult age were collected and after 2 d were exposed to 100/50 (flowrates) ethanol/air vapour in the booz-o-mat, and their loss of righting was determined by slight tapping as described previously[86]. For tolerance, flies were put back onto regular food after an initial 30 min exposure and were then re-exposed to the same vapour 4 h later. Note that tolerance is not connected to initial sensitivity, and flies naively sensitive to ethanol-induced sedation can have a no-tolerance phenotype or a reduced-tolerance phenotype. Flies overexpressing h*ZIP8* (and their sibling controls) were placed at 28 °C for 2 d to increase the expression levels of the transgene, as we did not detect a phenotype when they were kept at 25 °C (data not shown). Data from experimental and control flies were compared using two-sided Student's *t*-tests. Data were normally distributed according to Shapiro–Wilk testing with Bonferroni adjustment for each of the three experiments.

Effects on other traits and diseases. We queried SNPs against GWAS results included in PhenoScanner (http://www.phenoscanner.medschl.cam.ac.uk) to investigate cross-trait effects, and we extracted all association results with genome-wide significance at $P < 5 \times 10^{-8}$ for all of the SNPs in high LD ($r^2 \geq 0.8$) with the 46 sentinel new SNPs to highlight the loci with strongest evidence of association with other traits. At the gene level, we performed over-representation enrichment analysis using WebGestalt[41] on the nearest genes to all alcohol-consumption-associated loci.

The genetic correlations between alcohol consumption and 235 other traits and diseases were obtained using the online software LD Hub. LD Hub is a centralized database of summary-level GWAS results and a web interface for LDSR analysis.

To estimate the potential causal effect of variants related to alcohol consumption on schizophrenia, we performed a Mendelian randomization analysis using publicly available GWAS data on schizophrenia and the Mendelian randomization package in R. The effect was estimated using the inverse-variance weighted method. Pleiotropy was tested by applying the MR-Egger regression method and heterogeneity statistics were obtained. The random-effects inverse-variance method was applied in the presence of heterogeneity[87].

GRSs and percentage of variance explained. We calculated an unbiased weighted GRS in 14,004 unrelated participants from the Airwave cohort—an independent cohort with high-quality HRC-imputed genetic data[33]. All previously reported and new variants were used for the construction of the GRS. We weighted the alcohol-consumption-increasing alleles by the β coefficients of the meta-analysis. We assessed the association of the GRS with alcohol intake and calculated the alcohol consumption levels for individuals in the top versus the bottom quintiles of the distribution. To calculate the percentage variance of alcohol consumption explained by genetic variants, we generated the residuals from a regression of alcohol consumption in the Airwave cohort. We then fit a second linear model for the trait residuals with all new and known variants plus the top ten principal components and estimated the percentage variance of the dependent variable explained by the variants.

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

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The UKB GWAS data can be assessed from the UKB data repository (https://biota.osc.ox.ac.uk/). The genetic and phenotypic UKB data are available through application to the UKB (https://www.ukbiobank.ac.uk). Summary GWAS data can be assessed by request to the corresponding authors and are available at LDHub (http://ldsc.broadinstitute.org/ldhub/).

Received: 1 November 2018; Accepted: 11 June 2019;
Published online: 29 July 2019

References

1. GBD 2016 Alcohol Collaborators. Alcohol use and burden for 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* **392**, 1015–1035 (2018).
2. Poznyak, V. & Rekve, D. (eds) *Global Status Report on Alcohol and Health 2018*. https://www.who.int/substance_abuse/publications/global_alcohol_report/gsr_2018/en/(WHO, 2018).
3. Wood, A. M. et al. Risk thresholds for alcohol consumption: combined analysis of individual-participant data for 599 912 current drinkers in 83 prospective studies. *Lancet* **391**, 1513–1523 (2018).
4. Verhulst, B., Neale, M. C. & Kendler, K. S. The heritability of alcohol use disorders: a meta-analysis of twin and adoption studies. *Psychol. Med.* **45**, 1061–1072 (2015).

5. Schumann, G. et al. *KLB* is associated with alcohol drinking, and its gene product β-klotho is necessary for FGF21 regulation of alcohol preference. *Proc. Natl Acad. Sci. USA* **113**, 14372–14377 (2016).
6. Clarke, T. K. et al. Genome-wide association study of alcohol consumption and genetic overlap with other health-related traits in UK Biobank (N=112 117). *Mol. Psychiatr.* **22**, 1376–1384 (2017).
7. Jorgenson, E. et al. Genetic contributors to variation in alcohol consumption vary by race/ethnicity in a large multi-ethnic genome-wide association study. *Mol. Psychiatr.* **22**, 1359–1367 (2017).
8. Baik, I., Cho, N. H., Kim, S. H., Han, B. G. & Shin, C. Genome-wide association studies identify genetic loci related to alcohol consumption in Korean men. *Am. J. Clin. Nutr.* **93**, 809–816 (2011).
9. Jackson, B. et al. Update on the aldehyde dehydrogenase gene (*ALDH*) superfamily. *Hum. Genomics* **5**, 283–303 (2011).
10. Sudlow, C. et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
11. McCarthy, S. et al. Haplotype Reference Consortium. A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
12. Bulik-Sullivan, B. K. et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
13. Evangelou, E. & Ioannidis, J. P. Meta-analysis methods for genome-wide association studies and beyond. *Nat. Rev. Genet.* **14**, 379–389 (2013).
14. Evangelou, E. et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat. Genet.* **50**, 1412–1425 (2018).
15. Desikan, R. S. et al. Genetic overlap between Alzheimer's disease and Parkinson's disease at the *MAPT* locus. *Mol. Psychiatr.* **20**, 1588–1595 (2015).
16. Do, C. B. et al. Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease. *PLoS Genet.* **7**, e1002141 (2011).
17. Pankratz, N. et al. Meta-analysis of Parkinson's disease: identification of a novel locus, *RIT2*. *Ann. Neurol.* **71**, 370–384 (2012).
18. Okbay, A. et al. Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat. Genet.* **48**, 624–633 (2016).
19. Couch, F. J. et al. Genome-wide association study in *BRCA1* mutation carriers identifies novel loci associated with breast and ovarian cancer risk. *PLoS Genet.* **9**, e1003212 (2013).
20. Ikram, M. A. et al. Common variants at 6q22 and 17q21 are associated with intracranial volume. *Nat. Genet.* **44**, 539–544 (2012).
21. van der Harst, P. et al. Seventy-five genetic loci influencing the human red blood cell. *Nature* **492**, 369–375 (2012).
22. Samuel, A. et al. Six3 regulates optic nerve development via multiple mechanisms. *Sci. Rep.* **6**, 20267 (2016).
23. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
24. Liu, J. Z. et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat. Genet.* **47**, 979–986 (2015).
25. International Consortium for Blood Pressure Genome-Wide Association Studies. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* **478**, 103–109 (2011).
26. Speliotes, E. K. et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **42**, 937–948 (2010).
27. Teslovich, T. M. et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707–713 (2010).
28. Lim, C. S. & Alkon, D. L. Protein kinase C stimulates HuD-mediated mRNA stability and protein expression of neurotrophic factors and enhances dendritic maturation of hippocampal neurons in culture. *Hippocampus* **22**, 2303–2319 (2012).
29. Barker, J. M., Taylor, J. R., De Vries, T. J. & Peters, J. Brain-derived neurotrophic factor and addiction: pathological versus therapeutic effects on drug seeking. *Brain Res.* **1628**, 68–81 (2015).
30. Tanaka, T. et al. Genome-wide meta-analysis of observational studies shows common genetic variants associated with macronutrient intake. *Am. J. Clin. Nutr.* **97**, 1395–1402 (2013).
31. Talukdar, S. et al. FGF21 regulates sweet and alcohol preference. *Cell Metab.* **23**, 344–349 (2016).
32. Grant, S. F. et al. Association analysis of the *FTO* gene with obesity in children of Caucasian and African ancestry reveals a common tagging SNP. *PLoS One* **3**, e1746 (2008).
33. Elliott, P. et al. The airwave health monitoring study of police officers and staff in Great Britain: rationale, design and methods. *Environ. Res.* **134**, 280–285 (2014).
34. Elliott, L. T. et al. Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature* **562**, 210–216 (2018).
35. Stipanovich, A. et al. A phosphatase cascade by which rewarding stimuli control nucleosomal response. *Nature* **453**, 879–884 (2008).
36. Yang, B. Z. et al. Association of haplotypic variants in *DRD2*, *ANKK1*, *TTC12* and *NCAM1* to alcohol dependence in independent case control and family samples. *Hum. Mol. Genet.* **16**, 2844–2853 (2007).
37. Gelernter, J. et al. Haplotype spanning *TTC12* and *ANKK1*, flanked by the *DRD2* and *NCAM1* loci, is strongly associated to nicotine dependence in two distinct American populations. *Hum. Mol. Genet.* **15**, 3498–3507 (2006).
38. Treutlein, J. et al. Genetic association of the human corticotropin releasing hormone receptor 1 (*CRHR1*) with binge drinking and alcohol intake patterns in two independent samples. *Mol. Psychiatr.* **11**, 594–602 (2006).
39. Timpl, P. et al. Impaired stress response and reduced anxiety in mice lacking a functional corticotropin-releasing hormone receptor 1. *Nat. Genet.* **19**, 162–166 (1998).
40. Ruggeri, B. et al. Association of protein phosphatase *PPM1G* with alcohol use disorder and brain activity during behavioral control in a genome-wide methylation analysis. *Am. J. Psychiatr.* **172**, 543–552 (2015).
41. Wang, J., Vasaikar, S., Shi, Z., Greer, M. & Zhang, B. WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res.* **45**, W130–W137 (2017).
42. Gonzalez, D. A. et al. The Arf6 activator Efa6/PSD3 confers regional specificity and modulates ethanol consumption in *Drosophila* and humans. *Mol. Psychiatr.* **23**, 621–628 (2018).
43. Ojelade, S. A. et al. Rsu1 regulates ethanol consumption in *Drosophila* and humans. *Proc. Natl Acad. Sci. USA* **112**, E4085–E4093 (2015).
44. Rademakers, R., Cruts, M. & van Broeckhoven, C. The role of tau (*MAPT*) in frontotemporal dementia and related tauopathies. *Hum. Mutat.* **24**, 277–295 (2004).
45. Higashi, Y. et al. Influence of extracellular zinc on M1 microglial activation. *Sci. Rep.* **7**, 43778 (2017).
46. Chen, G. et al. Striatal involvement in human alcoholism and alcohol consumption, and withdrawal in animal models. *Alcohol. Clin. Exp. Res.* **35**, 1739–1748 (2011).
47. Okada, N. et al. Abnormal asymmetries in subcortical brain volume in schizophrenia. *Mol. Psychiatr.* **21**, 1460–1466 (2016).
48. van Erp, T. G. et al. Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. *Mol. Psychiatr.* **21**, 547–553 (2016).
49. Meyers, J. L. et al. The association between *DRD2*/*ANKK1* and genetically informed measures of alcohol use and problems. *Addict. Biol.* **18**, 523–536 (2013).
50. Logrip, M. L., Barak, S., Warnault, V. & Ron, D. Corticostriatal BDNF and alcohol addiction. *Brain Res.* **1628**, 60–67 (2015).
51. Boschen, K. E., Criss, K. J., Palamarchouk, V., Roth, T. L. & Klintsova, A. Y. Effects of developmental alcohol exposure vs. intubation stress on BDNF and TrkB expression in the hippocampus and frontal cortex of neonatal rats. *Int. J. Dev. Neurosci.* **43**, 16–24 (2015).
52. Monaco, A. et al. A complex network approach reveals a pivotal substructure of genes linked to schizophrenia. *PLoS One* **13**, e0190110 (2018).
53. Nielsen, S. M., Toftdahl, N. G., Nordentoft, M. & Hjorthoj, C. Association between alcohol, cannabis, and other illicit substance abuse and risk of developing schizophrenia: a nationwide population based register study. *Psychol. Med.* **47**, 1668–1677 (2017).
54. Nivard, M. G. et al. Connecting the dots, genome-wide association studies in substance use. *Mol. Psychiatr.* **21**, 733–735 (2016).
55. Gaziano, J. M. et al. Moderate alcohol intake, increased levels of high-density lipoprotein and its subfractions, and decreased risk of myocardial infarction. *N. Engl. J. Med.* **329**, 1829–1834 (1993).
56. Linn, S. et al. High-density lipoprotein cholesterol and alcohol consumption in US white and black adults: data from NHANES II. *Am. J. Publ. Health* **83**, 811–816 (1993).
57. Vu, K. N. et al. Causal role of alcohol consumption in an improved lipid profile: the atherosclerosis risk in communities (ARIC) study. *PLoS One* **11**, e0148765 (2016).
58. Chaput, J. P., McNeil, J., Despres, J. P., Bouchard, C. & Tremblay, A. Short sleep duration is associated with greater alcohol consumption in adults. *Appetite* **59**, 650–655 (2012).
59. Walters, R. K. et al. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat. Neurosci.* **21**, 1656–1669 (2018).
60. Bagnardi, V. et al. Alcohol consumption and site-specific cancer risk: a comprehensive dose-response meta-analysis. *Br. J. Cancer* **112**, 580–593 (2015).
61. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
62. Boniface, S., Kneale, J. & Shelton, N. Drinking pattern is more strongly associated with under-reporting of alcohol consumption than socio-demographic factors: evidence from a mixed-methods study. *BMC Publ. Health* **14**, 1297 (2014).

63. Greenfield, T. K. & Kerr, W. C. Alcohol measurement methodology in epidemiology: recent advances and opportunities. *Addiction* **103**, 1082–1099 (2008).
64. Grotz, A. K., Gloyn, A. L. & Thomsen, S. K. Prioritising causal genes at type 2 diabetes risk loci. *Curr. Diab. Rep.* **17**, 76 (2017).
65. Claussnitzer, M. et al. FTO obesity variant circuitry and adipocyte browning in humans. *N. Engl. J. Med.* **373**, 895–907 (2015).
66. Hambrecht, M. & Hafner, H. Substance abuse and the onset of schizophrenia. *Biol. Psychiatr.* **40**, 1155–1163 (1996).
67. Loh, P. R. et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **47**, 284–290 (2015).
68. Georgiopoulos, G. & Evangelou, E. Power considerations for λ inflation factor in meta-analyses of genome-wide association studies. *Genet. Res.* **98**, e9 (2016).
69. Winkler, T. W. et al. Quality control and conduct of genome-wide association meta-analyses. *Nat. Protoc.* **9**, 1192–1212 (2014).
70. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
71. Stacey, D. et al. *RASGRF2* regulates alcohol-induced reinforcement by influencing mesolimbic dopamine neuron activity and dopamine release. *Proc. Natl Acad. Sci. USA* **109**, 21128–21133 (2012).
72. Bakshi, A. et al. Fast set-based association analysis using summary data from GWAS identifies novel gene loci for human complex traits. *Sci. Rep.* **6**, 32894 (2016).
73. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
74. Ramasamy, A. et al. Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat. Neurosci.* **17**, 1418–1428 (2014).
75. Ongen, H., Buil, A., Brown, A. A., Dermitzakis, E. T. & Delaneau, O. Fast and efficient QTL mapper for thousands of molecular phenotypes. *Bioinformatics* **32**, 1479–1485 (2016).
76. Storey, J. D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc. Natl Acad. Sci. USA* **100**, 9440–9445 (2003).
77. Shabalin, A. A. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* **28**, 1353–1358 (2012).
78. Brown, C. A. et al. Development, validation and application of a new fornix template for studies of aging and preclinical Alzheimer's disease. *Neuroimage Clin.* **13**, 106–115 (2017).
79. Diedrichsen, J. et al. Imaging the deep cerebellar nuclei: a probabilistic atlas and normalization procedure. *Neuroimage* **54**, 1786–1794 (2011).
80. Zhang, Y., Brady, M. & Smith, S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imaging* **20**, 45–57 (2001).
81. Wager, T. D., Davidson, M. L., Hughes, B. L., Lindquist, M. A. & Ochsner, K. N. Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron* **59**, 1037–1050 (2008).
82. Wager, T. D. et al. Brain mediators of cardiovascular responses to social threat: part I: reciprocal dorsal and ventral sub-regions of the medial prefrontal cortex and heart-rate reactivity. *Neuroimage* **47**, 821–835 (2009).
83. Petersen, S. E. et al. UK Biobank's cardiovascular magnetic resonance protocol. *J. Cardiovasc. Magn. Reson.* **18**, 8 (2016).
84. Bai, W. et al. Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. *J. Cardiovasc. Magn. Reson.* **20**, 65 (2018).
85. Linge, J. et al. Body composition profiling in the UK Biobank imaging study. *Obesity* (2018).
86. Peru y Colón de Portugal, R. L. et al. Adult neuronal Arf6 controls ethanol-induced behavior with Arfaptin downstream of Rac1 and RhoGAP18B. *J. Neurosci.* **32**, 17706–17713 (2012).
87. Dimou, N. L. & Tsilidis, K. K. A primer in Mendelian randomization methodology with a focus on utilizing published summary association data. *Methods Mol. Biol.* **1793**, 211–230 (2018).

Acknowledgements

H.G. was funded by the NIHR Imperial College Health Care NHS Trust and Imperial College London Biomedical Research Centre. I.K. was supported by the EU PhenoMeNal project (Horizon 2020, grant no. 654241) and the UK Dementia Research Institute, which is supported by the MRC, the Alzheimer's Society and Alzheimer's Research UK. S. Thériault was supported by the Canadian Institutes of Health Research and Université Laval (Quebec City, Canada). L.R. was supported by Forschungs- und Förder-Stiftung INOVA, Vaduz, Liechtenstein. D.C. holds a McMaster University Department of Medicine Mid-Career Research Award. M.B. is supported by NIH grant R01-DK062370. P.v.d.H. was supported by ICIN-NHI and Marie Skłodowska-Curie GF (call: H2020-MSCA-IF-2014; Project ID: 661395). C.H. was supported by a core MRC grant to the MRCHGU QTL in Health and Disease research programme. N.V. was supported by Marie Skłodowska-Curie GF (grant no. 661395) and ICIN-NHI. Q.L. is partially supported by the National Natural Sciences Foundation of China (No. 81873909), Shanghai Municipal Science and Technology Major Project(No.2018SHZDZX01) and ZJLab. P.E. acknowledges support from the NIHR Biomedical Research Centre at Imperial College Healthcare NHS Trust and Imperial College London, the NIHR Health Protection Research Unit in Health Impact of Environmental Hazards (HPRU-2012-10141), the Medical Research Council (MRC) and Public Health England (PHE) Centre for Environment and Health (MR/L01341X/1) and Health Data Research (HDR) UK. P.E. is supported by a UK Dementia Research Institute (DRI) professorship, UK DRI at Imperial College London, funded by the MRC, Alzheimer's Society and Alzheimer's Research UK. This work received support from the following sources: the European Union-funded FP6 Integrated Project IMAGEN (Reinforcement-related behaviour in normal brain function and psychopathology; LSHM-CT- 2007-037286), the Horizon 2020 funded ERC Advanced Grant 'STRATIFY' (Brain network based stratification of reinforcement-related disorders; grant no. 695313), ERANID (Understanding the interplay between cultural, biological and subjective factors in drug use pathways; PR-ST-0416-10004), BRIDGET (JPND: BRain Imaging, cognition Dementia and next generation GEnomics; MR/N027558/1), the FP7 projects IMAGEMEND (grant no. 602450; IMAging GEnetics for MENtal Disorders) and MATRICS (grant no. 603016), the Innovative Medicine Initiative Project EU-AIMS (grant. no115300-2), the Medical Research Council Grant 'c-VEDA' (Consortium on Vulnerability to Externalizing Disorders and Addictions; MR/N000390/1), the Swedish Research Council FORMAS, the Medical Research Council, the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London, the Bundesministeriumfür Bildung und Forschung (BMBF grants 01GS08152; 01EV0711; eMED SysAlc01ZX1311A; and Forschungsnetz AERIAL 01EE1406A and 01EE1406B), the Deutsche Forschungsgemeinschaft (DFG grants SM 80/7-2 and SFB 940/2), the Medical Research Foundation and Medical Research Council (MR/R00465X/1), the Human Brain Project (HBP SGA 2). Further support was provided by grants from ANR (project AF12-NEUR0008-01-WM2NA, and ANR-12-SAMA-0004), the Fondation de France, the Fondation pour la Recherche Médicale, the Mission Interministérielle de Lutte-contre-les-Drogues-et-les-Conduites-Addictives (MILDECA), the Assistance-Publique-Hôpitaux-de-Paris and INSERM (interface grant), Paris Sud University IDEX 2012; the National Institutes of Health, Science Foundation Ireland (16/ERCD/3797), USA (Axon, Testosterone and Mental Health during Adolescence; RO1 MH085772-01A1); and by NIH Consortium grant U54 EB020403, supported by a cross-NIH alliance that funds Big Data to Knowledge Centres of Excellence. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

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., J.E., Q.L., N.D., D.L., I.T., J.D.B., P.M.M., A.R., S.D., G.S. and P.E. Writing of the manuscript: E.E., H.G., C.C., G.N., P.B., A.R.B., R.P., H. Suzuki, F.K., A.M.Y., I.K., D.L., I.T., J.D.B., P.M.M., A.R., S.D., G.S. and P.E. Association of MRI analysis: C.C., H. Suzuki, A.M.Y., A.I.B., J.D.B., P.M.M. and G.S. AlcGen and Charge+ contributors, by study: (ARIC): A.C.M., M.R.B., B.Y. and D.E.A. (CHS): B.M.P., R.N.L., T.M.B. and J.A.B. (FHS): D.L. and C.L. (GAPP/Swiss-AF/Beat-AF): S. Thériault, S.A., D.C., L.R. and M. Kühne. (GENOA): S.L.R.K., J.A.S., W.Z. and S.M.R., (GRAPHIC): N.J.S., C.P.N. and P.S.B., (GS): A.M.M., T.-K.C., C.H. and D.P., (HBCS): J.L., S. Tuominen, M.-M.P. and J.G.E. (HRS): D.R.W., S.L.R.K., J.D.F., W.Z. and J.A.S. (MESA): X.G., J.Y., A.W. and J.I.R. (METSIM): M.L., A.S., J. Vangipurapu and J.K. (FUSION): M.B., K.L.M., L.J.S. and A.U.J. (NESDA): B.W.J.H.P. and Y.M. (NFBC): M.-R.J., J. Veijola, M. Männikkö and J.A. (ORCADES): H.C. and P.K.J. (VIKING): J.F.W. and K.A.K. (Croatia-VIS): I.R. and O.P. (Croatia-KORCULA): C.H. (PREVEND): N.V. and P.v.d.H. (OZALC): N.G.M., J.B.W., P.A.L. and A.C.H. (SHIP): A.T., H.J.G., S.E.B. and G.H. (TRAILS-pop): A.J.O. and I.M.N. (TRAILS-CC): C.A.H. and H. Snieder. (TwinsUK): T.D.S. and M. Mangino. (YFS): L.-P.L., M. Kähönen, O.T.R. and T.L. All authors critically reviewed and approved the final version of the manuscript

Competing interests

B.M.P. serves on the DSMB of a clinical trial funded by the manufacturer (Zoll LifeCor) and on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson. B.W.J.H.P. has received research funding (unrelated to the work reported here) from Jansen Research and Boehringer Ingelheim. The other authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41562-019-0653-z.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to G.S. or P.E.

Peer review information: Primary Handling Editor: Stavroula Kousta.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

[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]Centre for Population Neuroscience and Precision Medicine (PONS), Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK. [5]NIHR Imperial Biomedical Research Centre, ITMAT Data Science Group, Imperial College London, London, UK. [6]Molecular Medicine, School of Medicine, University of Utah, Salt Lake City, UT, USA. [7]Centre for Restorative Neurosciences, Division of Brain Sciences, Department of Medicine, Hammersmith Campus, Imperial College London, London, UK. [8]Tohoku Medical Megabank Organization, Tohoku University, Sendai, 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, UK. [12]Institute of Science and Technology for Brain-Inspired Intelligence, MOE-Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence, Fudan University, Shanghai, China. [13]Department of Psychology and the Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge, UK. [14]Cardiology Division, University Hospital Basel, Basel, Switzerland. [15]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA. [16]Department of Biostatistics, University of Washington, Seattle, WA, USA. [17]Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany. [18]Chair of Epidemiology, Ludwig-Maximilians-Universitat Munchen, UNIKA-T Augsburg, Augsburg, Germany. [19]Department of Cardiovascular Sciences, University of Leicester, Cardiovascular Research Centre, Glenfield Hospital, Leicester, UK. [20]NIHR Leicester Biomedical Research Centre, Glenfield Hospital, Leicester, 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, USA. [22]Department of Psychiatry, University of Edinburgh, Edinburgh, UK. [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, Greifswald, Germany. [25]Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA. [26]Centre for Global Health Research, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK. [27]Centre for Cardiovascular Sciences, Queens Medical Research Institute, University of Edinburgh, Edinburgh, UK. [28]Psychiatric Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [29]Department of Clinical Chemistry, Fimlab Laboratories, Tampere, Finland. [30]Department of Clinical Chemistry, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and LHealth Technology, Tampere University, Tampere, Finland. [31]Department of Cardiology, Heart Center, Tampere University Hospital, Tampere, Finland. [32]Department of Twin Research and Genetic Epidemiology, King's College London, London, UK. [33]NIHR Biomedical Research Centre, Guy's and St Thomas Foundation Trust, London, UK. [34]Department of Psychiatry, Amsterdam Neuroscience and Amsterdam Public Health Research Institute, Amsterdam University Medical Center, Amsterdam, the Netherlands. [35]Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands. [36]Folkhälsan Research Center, Helsinki, Finland. [37]Department of Public Health Solutions, National Institute for Health and Welfare, Helsinki, Finland. [38]Faculty of Medicine, University of Split, Split, Croatia. [39]Generation Scotland, Medical Genetics Section, Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK. [40]Centre for Cognitive Ageing and Cognitive Epidemiology, Edinburgh, UK. [41]Department of Epidemiology, University of Michigan, Ann Arbor, MI, USA. [42]Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland, Kuopio, Finland. [43]German Centre for Cardiovascular Research (DZHK), partner site Greifswald, Greifswald, Germany. [44]Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland. [45]Population Health Research Institute, McMaster University, Hamilton, Ontario, Canada. [46]Department of Molecular Biology, Medical Biochemistry and Pathology, Laval University, Quebec City, Quebec, Canada. [47]Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [48]Department of Pediatrics/Nutrition, Baylor College of Medicine, Houston, TX, USA. [49]The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA, USA. [50]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA. [51]Center for Life Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Finland. [52]Oulunkaari Health Center, Ii, Finland. [53]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA. [54]Northern Finland Birth Cohorts, Faculty of Medicine, University of Oulu, Oulu, Finland. [55]Institute of Clinical Chemistry, Inselspital Bern, University Hospital, University of Bern, 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, CA, USA. [59]Department of Psychiatry, Research Unit of Clinical Neuroscience, University of Oulu, Oulu, Finland. [60]Department of Psychiatry, University Hospital of Oulu, Oulu, Finland. [61]Medical research Center Oulu, University and University Hospital of Oulu, Oulu, Finland. [62]Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland. [63]National Institute for Health and Welfare, Helsinki, Finland. [64]Unit of General Practice, Helsinki University Central Hospital, Helsinki, Finland. [65]Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Germany. [66]German Center for Neurodegenerative Diseases (DZNE), Rostock/Greifswald, Greifswald, Germany. [67]Department of Cardiology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands. [68]Department of Genetics, University Medical Center Groningen, University of Groningen, 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, Groningen, the Netherlands. [71]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh, UK. [72]Department of Psychiatry, School of Medicine, Washington University in St Louis, St Louis, MO, USA. [73]Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, UK. [74]Center for Life Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Finland. [75]Biocenter Oulu, University of Oulu, Oulu, Finland. [76]Unit of Primary Health Care, Oulu University Hospital, OYS, Oulu, Finland. [77]Department of Life Sciences, College of Health and Life Sciences, Brunel University London, London, UK. [78]Department of Clinical Physiology, Tampere University Hospital, Tampere, Finland. [79]Department of Clinical Physiology, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland. [80]Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland and Kuopio University Hospital, 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, USA. [83]Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA. [84]Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland. [85]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland. [86]Centre for Population Health Research, University of Turku and Turku University Hospital, Turku, Finland. [87]Framingham Heart Study, Framingham, MA, USA. [88]Population Sciences Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA. [89]Research Centre for Optimal Health, Department of Life Sciences, University of Westminster, London, UK. [90]Departments of Psychiatry, Neurobiology & Anatomy, Human Genetics, School of Medicine, University of Utah, Salt Lake City, UT, USA. [91]PONS Research Group, Dept of Psychiatry and Psychotherapy, Campus Charite Mitte, Humboldt University, Berlin, Germany and Institute for Science and Technology of Brain-inspired Intelligence (ISTBI), Fudan University, Shanghai, P.R. China. [92]National Institute for Health Research Imperial Biomedical Research Centre, Imperial College Healthcare NHS Trust and Imperial College London, London, UK. [93]Health Data Research UK London Substantive Site, London, UK. [94]These authors contributed equally: Evangelos Evangelou, He Gao, Congying Chu, Georgios Ntritsos, Jimmy D. Bell, Paul M. Matthews, Adrian Rothenfluh, Sylvane Desrivières, Gunter Schumann, Paul Elliott. *e-mail: gunter.schumann@kcl.ac.uk; p.elliott@imperial.ac.uk

nature research

Corresponding author(s): Gunter Schumann and Paul Elliott

Last updated by author(s): Jun 6, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Open source software was used for this project including BOLT LLM v2.3.2, R v3.5.3, PLINK v1.9, Mango v4.1, GCTA v1.92. We also used IPA for pathway analysis.
Data analysis	For the UKB GWAS analysis, BOLT-LMM software v2.3 was used for running an association analysis using linear mixed modelling; then METAL software was used for all meta-analyses with a fixed effects inverse variance weighted meta-analysis approach. We used R software for any general statistical analyses, for secondary analyses (e.g. variance explained analyses, risk score analyses) and for producing plots in the figures. We used PLINK software for LD calculations of variants. For the bioinformatics analyses, specific software was used for each different analysis. Each method and the software used is described in the Online Methods, For example, IPA was used for the over-representation enrichment analysis etc.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The UKB GWAS data can be assessed from the UK Biobank data repository (http://biota.osc.ox.ac.uk/). The genetic and phenotypic UKB data are available upon application to the UK Biobank (https://www.ukbiobank.ac.uk). Summary GWAS data data can be assessed by request to the corresponding authors and will be available via LDHub (http://ldsc.broadinstitute.org/ldhub/).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	From the ~500,000 UKB participant available, we analyzed 404,731 subjects which passed quality control (QC) of the genetic data, were of European ancestry, and met our phenotypic data QC requirements for availability. To maximize sample size in the discovery, we recruited an additional 76,111 samples of European descent from 25 cohorts participating in the AlcGen and the CHARGE+ consortia reaching a total discovery sample size of N=480,842.
Data exclusions	Within UK Biobank, we excluded samples according to both genetic data quality control (QC) and phenotypic data QC. From genetic data QC, we excluded 968 subjects listed as QC outliers for heterozygosity or missingness within the centrally provided UK Biobank sample QC files, and 378 individuals with sex discordance between the phenotypic and genetically inferred sex. We also restricted to subjects of European ancestry, according to both self-reported ethnicity status and ancestry clustering using PCA data. For phenotypic QC, we excluded any subjects defined as long-time non-drinkers. Participants with with daily alcohol consumption >500 grams were excluded as outliers. We also excluded participants with missing covariates. Similar sample QC was performed at study level within each of the AlcGen and CHARGE+ consortia.
Replication	We adopted a one-stage design with a stringent p-value threshold. Novel loci identified from our 1-stage approach met our criteria for internal replication by showing significant support within each of the UKB and AlcGen plus CHARGE+ GWAS datasets separately.
Randomization	N/A for GWAS
Blinding	N/A for GWAS (Note data collection of UK Biobank was done centrally, not performed by us)

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	Summary descriptives of UKB, AlcGen and CHARGE+ individuals are provided in Supplementary Tables 1, 2, showing descriptive summary statistics in UKB (ST1), descriptive summary statistics in AlcGen and CHARGE plus cohorts (ST2).
Recruitment	Data collection of UK Biobank was done centrally, not performed by us
Ethics oversight	Research protocol for UK Biobank data has been approved by UK Biobank (application 13375) . Summary data have been provided from the other studies. All studies had their own research protocols approved by the respective local ethics committees. Participants provided written informed consent

Note that full information on the approval of the study protocol must also be provided in the manuscript.