Genome-wide analysis of risk-taking behaviour and cross-disorder genetic correlations in 116 255 individuals from the UK Biobank cohort.

Rona J. Strawbridge PhD[1,2*+], Joey Ward MSc[1*] Breda Cullen DClinPsy [1], Elizabeth M. Tunbridge DPhil[3,4], Sarah Hartz MD[5], Laura Bierut MD[5], Amy Horton PhD[5,6], Mark E. S. Bailey PhD[7], Nicholas Graham MBChB[1], Amy Ferguson BSc[1], Donald M. Lyall PhD[1], Daniel Mackay PhD[1], Laura M. Pidgeon PhD[1], Jonathan Cavanagh MD[1], Jill P. Pell MD[1], Michael O'Donovan PhD[8], Valentina Escott-Price PhD[8], Paul J. Harrison FRCPsych[3,4], Daniel J. Smith MD[1].

[1]Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK. [2]Department of Medicine Solna, Karolinska Institute, Stockholm, Sweden. [3]Department of Psychiatry, University of Oxford, Oxford, UK. [4]Oxford Health NHS Foundation Trust, Oxford, UK. [5]Department of Psychiatry, Washington University School of Medicine in St. Louis, St. Louis, Missouri, USA. [6]Transmontane Analytics, Tuscon, Arizona, USA. [7]School of Life Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK. [8]MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, UK.

*Authors contributed equally

[+]**Corresponding author:** Dr Rona Strawbridge, Institute of Health and Wellbeing, University of Glasgow, Room 111, Public Health, 1 Lilybank Gardens, Glasgow, G12 8RZ, UK. Tel: +44 141 211 3935. E-mail: rona.strawbridge@glasgow.ac.uk

Running title: Genetics of risk-taking behaviour

Abstract

Risk-taking behaviour is a key component of several psychiatric disorders and could influence lifestyle choices such as smoking, alcohol use and diet. Risk-taking behaviour therefore fits within a Research Domain Criteria (RDoC) approach, whereby elucidation of the genetic determinants of this trait has the potential to improve our understanding across different psychiatric disorders. Here we report a genome wide association study in 116 255 UK Biobank participants who responded yes/no to the question "would you consider yourself a risk-taker?" Risk-takers (compared to controls) were more likely to be men, smokers and have a history of mental illness. Genetic loci associated with risk-taking behaviour were identified on chromosomes 3 (rs13084531) and 6 (rs9379971). The effects of both lead SNPs were comparable between men and women. The chromosome 3 locus highlights *CADM2*, previously implicated in cognitive and executive functions, but the chromosome 6 locus is challenging to interpret due to the complexity of the HLA region. Risk-taking behaviour shared significant genetic risk with schizophrenia, bipolar disorder, attention deficit hyperactivity disorder and post-traumatic stress disorder, as well as with smoking and total obesity. Despite being based on only a single question, this study furthers our understanding of the biology of risk-taking behaviour, a trait which has a major impact on a range of common physical and mental health disorders.

Introduction

Risk-taking behaviour is an important aspect of several psychiatric disorders, including attention deficit hyperactivity disorder (ADHD), bipolar disorder (BD) and schizophrenia, as well as problem behaviours such as smoking and drug and alcohol misuse [1-3]. Physical health problems such as obesity might also be considered to be related to increased propensity towards risk-taking: obesity includes aspects of reward processing, response inhibition and decision-making[4]. The Research Domain Criteria (RDoC) approach suggests that studying dimensional psychopathological traits (rather than discrete diagnostic categories) may be a more useful strategy for identifying biology which cuts across psychiatric diagnoses [5]. In this respect, risk-taking behaviour is an important phenotype for investigation. It may also be useful for investigating the overlap between psychiatric disorders and conditions such as obesity, which include components of poor impulse control and abnormal reward processing.

To date, there has been no focused genetic study of risk-taking behaviour. GWAS of related phenotypes, such as impulsivity and behavioural disinhibition, have so far been underpowered for detecting associations at a genome-wide level. Here we conduct a GWAS of self-reported risk-taking behaviour in 116 255 participants from the UK Biobank cohort. We use expression quantitative trait loci analysis to highlight plausible candidate genes and we assess the extent to which there is a genetic correlation between risk-taking and several mental and physical health disorders, including ADHD, BD, major depressive disorder (MDD), post-traumatic stress disorder (PTSD), anxiety, alcohol use disorder, smoking and obesity.

Materials and methods

Sample

UK Biobank is a large population cohort which aims to investigate a diverse range of factors influencing risk of diseases which are common in middle and older age. Between 2006 and 2010, more than 502 000 participants (age range from 40 and 69 years) were recruited from 22 centres across the UK [6]. Comprehensive baseline assessments included social circumstances, cognitive abilities, lifestyle and measures of physical health status. The present study used the first release of genetic data on approximately one third of the UK Biobank cohort. In order to maximise homogeneity, we included only participants of (self-reported) white United Kingdom (UK) ancestry.

Informed consent was obtained by UK Biobank from all participants. This study was carried out under the generic approval from the NHS National Research Ethics Service (approval letter dated 13 May 2016, Ref 16/NW/0274) and under UK Biobank approval for application #6553 "Genome-wide association studies of mental health" (PI Daniel Smith).

Risk-taking phenotype

The baseline assessment of UK Biobank participants included the question "Would you describe yourself as someone who takes risks?" (data field #2040), to which participants replied yes or no. Of the sample with available genetic data, a total of 29 703 individuals responded 'yes' to this question (here referred to as 'risk-takers') and 86 552 responded 'no' (here referred to as controls). N=3 769 individuals preferred not to say or were missing this data and were excluded from analysis. Characteristics of the individuals included in this analysis are described in Table 1.

Genotyping, imputation and quality control

The first release of genotypic data from UK Biobank, in June 2015, included 152 729 UK Biobank participants. Samples were genotyped with either the Affymetrix UK Biobank Axiom array (Santa

Clara, CA, USA; approximately 67%) or the Affymetrix UK BiLEVE Axiom array (33%), which share at least 95% of content. Autosomal data only were available.

Imputation of the data has previously been described in the UK Biobank interim release documentation [7]. In brief, SNPs were excluded prior to imputation if they were multiallelic or had minor allele frequency (MAF) <1%. A modified version of SHAPEIT2 was used for phasing and IMPUTE2 (implemented on a C++ platform) was used for the imputation [8, 9]. A merged reference panel of 87 696 888 biallelic variants on 12 570 haplotypes constituted from the 1000 Genomes Phase 3 and UK10K haplotype panels [10] was used as the basis for the imputation. Imputed variants with MAF <0.001% were filtered out of the dataset used for subsequent analysis.

The Wellcome Trust Centre for Human Genetics applied stringent quality control, as described in UK Biobank documentation [11], before release of the genotypic data set. UK Biobank genomic analysis exclusions were applied (Biobank Data Dictionary item #22010). Participants were excluded from analyses due to relatedness (#22012: genetic relatedness factor; one member of each set of individuals with KING-estimated kinship coefficient >0.0442 was removed at random), sex mismatch (reported compared to genetic) (#22001: genetic sex), non-Caucasian ancestry (#22006: ethnic grouping; self-reported and based on principal component analysis of genetic data), and quality control failure (#22050: UK BiLEVE Affymetrix quality control for samples and #22051: UK BiLEVE genotype quality control for samples). SNPs were removed due to deviation from Hardy–Weinberg equilibrium at $P<1x10^{-6}$, MAF <0.01, imputation quality score <0.4 and >10% missingness in the sample after excluding genotype calls made with <90% posterior probability.

Association analyses

A total of 116 255 individuals and 8 781 003 variants were included in the analysis. 29 703 participants were classed as risk-takers and 86 552 were controls. Association analysis was conducted in PLINK [12] using logistic regression, assuming a model of additive allelic effects and models were adjusted for sex, age, genotyping array, and the first 8 genetic PCs (Biobank Data Dictionary items #22009.01

to #22009.08) to control for hidden population stratification. The threshold for GWAS significance was set at $p<5x10^{-8}$.

Data mining

SNPs associated (at genome-wide significance) with risk-taking behaviour were further investigated for influence on nearby genes (Variant Effect Predictor, VEP [13]) and for reported associations with relevant traits (GWAS catalogue [14]). Descriptions and known or predicted functions of implicated genes were compiled (GeneCards www.genecards.org and Entrez Gene www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene) and global patterns of tissue expression were assessed (GTEx [15]). Exploratory analyses of the impact of significant loci on the expression of nearby genes were carried out using the GTEx Portal "Test your own eQTL" function [15]. In the 13 brain regions available in the GTEx dataset, we tested for associations between rs13084531 and *CADM2* expression, and between rs9379971 and the expression of *POM121L2, PRSS16*, *ZNF204P* and *VN1R10P*.

SNP heritability and genetic correlation analyses

Linkage Disequilibrium Score Regression (LDSR) [16] was applied to the GWAS summary statistics to estimate the risk-taking SNP heritability (h^2_{SNP}). LDSR was also used to assess genetic correlations between risk-taking behaviour and ADHD, schizophrenia, BD, MDD, anxiety, PTSD, smoking status (ever smoked), fluid intelligence, years of education, obesity and alcohol use disorder. Two measures of obesity were included: Body-Mass Index (BMI) as a measure of total obesity [17] and waist-to-hip ratio adjusted for BMI (WHRadjBMI), reflecting metabolically-detrimental central obesity [18]. For the ADHD, schizophrenia, BD, MDD, anxiety, PTSD, and smoking status, we used GWAS summary statistics provided by the Psychiatric Genomics Consortium (http://www.med.unc.edu/pgc/) [19-25]. For the two obesity phenotypes, GWAS summary statistics for BMI [17] and WHRadjBMI [18] were taken from the consortium for the Genetic Investigation of Anthropometric Traits (http://portals.broadinstitute.org/collaboration/giant). Summary statistics for years of education [26] and fluid intelligence [27] were downloaded as instructed in the respective publications. Alcohol use

disorder was defined using DSM-5 criteria [28]. For this phenotype, a GWAS meta-analysis on genotypes imputed to 1000 Genomes was run with five datasets: COGEND, COGEND2, COGEND-23andMe, COGA, and FSCD. In total there were N=2 983 cases with alcohol use disorder and N=1 169 controls. Descriptions of the datasets are in the Supplementary information.

Results

Demographic characteristics

Small differences were observed between controls and risk-takers with regard to age and BMI (Table 1), but striking differences were observed for sex distribution, smoking and history of mood disorders: risk-takers (compared to controls) were more often men, more likely to be current or ever-smokers and more likely to have a lifetime history of severe depression or BD. Risk-takers were also more likely to have a university/college degree.

GWAS of risk-taking behaviour

GWAS results for risk-taking are summarised in Figure 1 (Manhattan plot), Figure 1 inset (QQ plot) and Supplementary Table 1 (Genome-wide significant loci associated with risk-taking in UK Biobank: basic and conditional analyses).

The GWAS data test statistics showed modest deviation from the null (λ_{GC} =1.13). Considering the sample size, the deviation was negligible (λ_{GC} 1000=1.002). LDSR suggested that deviation from the null was due to a polygenic architecture in which h^2_{SNP} accounted for approximately 4% of the population variance in risk-taking behaviour (observed scale h^2_{SNP}=0.058 (SE 0.006)), rather than inflation due to unconstrained population structure (LD regression intercept=1.003 (SE 0.008)).

Two loci were associated with risk-taking behaviour at genome-wide significance, on chromosome 3 and chromosome 6 (Figure 1 and Supplementary Table 1). The index SNP on chromosome (chr) 3, rs13084531, lies within the *CADM2* gene, however linkage disequilibrium (LD) suggests that the signal also encompasses *miR5688*, and borders a *CADM2* anti-sense transcript (*CADM2-AS2*, Figure 2a). The

minor allele of rs13084531 was associated with increased risk-taking (G allele, Odds Ratio (OR) 1.07, Confidence interval (CI) 1.04-1.09, P 8.75×10^{-9}). Conditional analysis of the chr3 locus (including rs13084531 as a covariate) is suggestive of a second signal (index SNP rs62250716, OR 0.96, CI 0.94-0.98, P 8.53×10^{-5}, LD r^2=0.16 with rs13084531, Figure 2b and Supplementary Table 1). The LD structure across the chr3 locus supports the possibility of two distinct signals (Supplementary Figure 1).

The chr6 locus lies within the gene-rich HLA region (Figure 2c), where index SNP rs9379971 demonstrated an association between the minor allele and decreased risk-taking (A allele, OR 0.95, CI 0.93-0.97, P 2.31×10^{-9}). Conditional analysis (including rs9379971 as a covariate) and assessment of the LD structure across this locus indicated that the associated region probably includes only one signal (Figure 2d, Supplementary Table 1 and Supplementary Figure 2).

As there was a predominance of men in the risk-taking group, within a secondary analysis, the lead SNP from each locus was assessed in men and women separately. The effect of rs13084531 (chr3) was comparable between men (N=51 662, Beta 0.061, se 0.016, p=0.0001), women (N=48 812, Beta 0.075, se 0.018, p=4.37×10^{-5}) and the combined dataset (Beta 0.067, se 0.012 p=8.75×10^{-9}). For rs9379971 (chr6) the signal (Beta -0.063, se 0.011, p=2.31×10^{-9}) appeared stronger in women (Beta -0.074, se 0.017, p=1.05×10^{-5}) compared to men (Beta -0.044, se 0.014 p=0.0022), but sex-interaction analysis did not support a significant difference (p=0.1764).

Data mining

As with the majority of SNPs identified by GWAS, the genome-wide significant SNPs in both loci are non-coding. Current prediction models ascribe only non-coding modifier functions to the 81 genome-wide significant SNPs (VEP [13], Supplementary Table 2). Expression quantitative trait analysis directly tests association of the index SNPs with expression of nearby transcripts. The chr3 index SNP (rs13084531) lies within the *CADM2* gene and adjacent to miR5688 and *CADM2-AS2* (Figure 2 and Supplementary Table 3). Currently most miRs are predicted (but not reliably proven) to influence

transcription of hundreds or thousands of genes. Furthermore, analysing transcription levels of miRs is challenging. Similarly, the importance of antisense transcripts such as *CADM2-AS2* is unclear and difficult to assess. *CADM2,* which encodes cell adhesion molecule 2 (also known as synaptic cell adhesion molecule, SynCAM2*),* is a plausible target gene as it is predominantly expressed in the brain (Supplementary Figure 3a). The risk allele at rs13084531 was associated with increased *CADM2* mRNA levels in several regions of the brain (including the caudate basal ganglia and putamen basal ganglia, hippocampus and hypothalamus, Supplemental Figure 4). *CADM1*, a related cell adhesion molecule, demonstrates overlapping and co-regulated (albeit inversely) expression patterns [29]. It is worth noting that *CADM1* shows a similar, albeit less brain-specific, expression pattern (Supplementary Figure 3b) and that genetic deletion of *Cadm1* in mice results in behavioural abnormalities, including anxiety [30].

A recent GWAS of executive functioning and information processing speed in non-demented older adults from the CHARGE (Cohorts for Heart and Aging Research in Genomic Epidemiology) consortium found that genetic variation in the *CADM2* gene was associated with individual differences in information processing speed [31]. In the same study, pathway analysis implicated *CADM2* in glutamate signalling and gamma-aminobutyric acid (GABA) transport [31]. The allele of rs17518584 (LD r^2=0.45 with rs13084531, LD r^2=0.34 with rs62250716) associated with increased processing speed was associated with reduced (self-reported) risk-taking in the current study (Supplementary Table 4, $p=1.17x10^{-7}$). Furthermore, a GWAS of educational attainment in the UK Biobank cohort demonstrated a significant signal in *CADM2* [32]. The effect allele of rs56262138 (LD r^2=0.00 with rs13084531, LD r^2=0.00 with rs62250716) for increased educational attainment showed a negative effect on risk-taking behaviour (Supplementary Table 4, p=0.0210).

Day *et al* reported an association between the *CADM2* locus and age of reproductive onset in UK Biobank [33]. In a secondary analysis, they also report an association between the same locus, *CADM2*, and risk-taking behaviour. However, we have applied more stringent quality control procedures (exclusion of 1 of each related pair of individuals, exclusion of UK Biobank-defined non-Caucasians and

exclusion of SNPs deviating from Hardy-Weinberg equilibrium) and standard analysis methods (logistic regression using PLINK and using eight principlal components to adjust for possible population structure). The lead SNP reported by Day *et al* failed quality control in our analysis, making direct comparison between that study and the current difficult.

The *CADM2* locus has also been tentatively associated with longevity [34] (rs9841144, LD r^2=0.99 with rs13084531, LD r^2=0.16 with rs62250716), but associations between *CADM2* SNPs and longevity, survival and attaining 100 years of age in that study were inconsistent, limiting the interpretation of these signals in the context of risk-taking behaviour.

The chr6 locus is a gene-dense region (Supplementary Table 3) with a highly complex LD structure. Global gene expression analysis (Supplementary Figure 5 A-C) highlighted *POM121L2* (testis-specific expression), *VN1R10P* (predominantly expressed by testis) and *ZNF192P2* (predominantly expressed by testis and brain) as potentially interesting genes. Although interpretation of the biological basis for association signals in this region is fraught with difficulty, the identification of associated eQTLs for genes expressed solely or predominantly in the testis is intriguing, particularly given the higher frequency of men in the group who consider themselves risk-takers.

No genotype-specific expression patterns were detected in the brain for *ZNF192P2* or (unsurprisingly) *POM121L2* and *VN1R10P*. Despite relatively low expression in the brain (Supplementary Figure 5 D), the minor allele of rs9379971 (associated with reduced risk-taking behaviour) was associated with increased levels of *PRSS16* mRNA (Supplementary Figure 6 A and B). Expression of *ZNF204P* levels was prominent in the brain (Supplementary Figure 5 E) and reduced levels of *ZNF204P* mRNA were associated with the minor (reduced risk-taking) allele of rs9379971 (Supplementary Figure 6 C).

Genetic correlations

We found significant positive genetic correlations between the risk-taking phenotype and ADHD (r_g=0.31, SE=0.13, p=0.01), schizophrenia (r_g=0.27, SE=0.04, p=4.54×10^{-11}), BD (r_g=0.26,

SE=0.07, p=1.73x10^{-4}), PTSD (r_g=0.51, SE=0.17, p=0.0018) and smoking (r_g=0.17, SE=0.07, p=0.01) and a negative genetic correlation with fluid intelligence (r_g=-0.15, SE=0.05, p=0.0013, Table 2). We found no significant genetic correlation between risk-taking and MDD, anxiety or years of education (Table 2). There was also a significant genetic correlation between risk-taking and BMI (r_g=0.10, SE=0.03, p=0.003), but a similar correlation was not found for WHRadjBMI. A genetic correlation between alcohol use disorder and the risk-taking phenotype was apparent (r_g=0.22, SE0.31, p=0.47), although this analysis was likely underpowered due to the modest size of the alcohol use disorder GWAS (n=4 171) and we draw no conclusions about this correlation.

Discussion

There is a growing emphasis on the importance of using phenotypic traits which cut across traditional diagnostic groups to investigate the biological basis of psychiatric disorders. Risk-taking behaviour is one such trans-nosological characteristic, recognised clinically as a feature of several disorders, including ADHD, schizophrenia and BD. In this study we identified two loci, on 3p12.1 and 6p22.1, that were associated with self-reported risk-taking behaviour.

The chr6 locus falls within the HLA region which encodes a large number of genes and is extremely complicated genetically. Further work is needed to be able to identify the genes responsible for the association with risk-taking behaviour identified here and to dissect the underlying behavioural mechanisms. The (non-significant) sex effect (that is, a stronger association in women) of the chr6 locus is intriguing, especially when considering that there are a number of testis-specific genes expressed in this locus. Further examination of this aspect of risk-taking behaviour would be of value when a larger dataset is available, thus offsetting the loss of power resulting from stratifying the analysis according to sex.

A key finding of our study was the positive association between chr3 SNP, rs13084531, and risk-taking behaviour as well as *CADM2* expression levels. Here, the allele associated with increased self-reported

risk-taking behaviour was also associated with increased *CADM2* expression. It is of interest that lack of *Cadm1* in mice was associated with anxiety-related behaviour [30] and that both *CADM1* and *CADM2* were identified as BMI-associated loci [17] suggesting that *CADM2* and related family members may be involved in balancing appetitive and avoidant behaviours.

Day and colleagues recently identified 38 genome-wide significant loci for age at first sexual intercourse within the UK Biobank cohort [33] and two of these loci were within the 3p12.1 region, close to *CADM2* (rs12714592 and rs57401290). The association between rs57401290 (and SNPs in LD) and age at first sexual intercourse was also observed for a number of behavioural traits, including number of sexual partners, number of children and risk-taking propensity. In addition, *CADM2* also showed association with information processing speed [31] and educational attainment [32], highlighting the complexity of relationships between cognitive performance and risk-taking. Taken together, this evidence suggests that *CADM2* plays a fundamental role in risk-taking behaviours, and may be a gene involved in the nexus of cognitive and reward-related processes that underlie them.

A perhaps surprising observation was the increased frequency of having a university degree in self-reported risk-takers, compared to controls, despite the negative (albeit non-significant) association between years of education and risk-taking behaviour. These observations underscore the complexity between risk-taking and educational attainment, and highlight differences between genetic and phenotypic relationships. They may also be indicative of selection bias within the UK Biobank cohort towards more highly educated individuals.

Another key finding was genetic correlation between self-reported risk-taking and obesity. Although there are likely to be a range of potential mechanisms linking risk-taking behaviour with obesity, evidence of a shared genetic component is in keeping with work that has highlighted the importance of the central nervous system in the regulation of obesity (BMI), particularly brain regions involved in cognition, learning and reward [17]. In contrast, central fat accumulation (WHRadjBMI) is primarily regulated by adipose tissue [18] which fits with the lower, non-significant genetic correlation between

risk-taking behaviour and this measure. Two SNPs (rs13078807 and rs13078960) in the *CADM2* locus have previously been associated with BMI [17, 35, 36], but whilst these SNPs tag each other (LD r^2=0.99), the LD between the risk-taking index SNP or possible secondary signal is low (LD r^2=0.31 and 0.01 for rs13084531 and rs62250716 respectively), suggesting that these are distinct signals.

It is perhaps unsurprising that we identified genetic correlations between risk-taking and smoking. Similarly, risk-taking and impulsive behaviour is a core feature of ADHD and BD, suggesting substantial genetic overlap between variants predisposing to risk-taking behaviour and these disorders. The genetic correlation between risk-taking and schizophrenia is of interest because schizophrenia is commonly comorbid with substance abuse disorders [37]. The correlation between risk-taking and PTSD is perhaps plausible if we accept that risk-takers may be more likely to find themselves in high-risk situations with the potential to cause psychological trauma. Overall, these correlations suggest that studying dimensional traits such as risk-taking has the potential to inform not only the biology of complex psychiatric disorders but might also represent a way to sub-classify or stratify individuals in terms of diagnosis and treatment.

Strengths and limitations

We acknowledge that Day *et al* have previously reported an association for risk-taking within the *CADM2* locus. Strengths of our study include the use of a more conservative and standardised methodology and reporting of results across the entire genome. A risk-taking locus was identified in the *HLA* region as well as the *CADM2* locus and we have shown that CADM2 may contain a second signal. Furthermore, we have investigated the possibility of a sex-specific effect of these loci, provided evidence highlighting possible candidate genes at both loci and confirmed the importance of this phenotype in relation to psychiatric illness. In short, our report provides a fuller understanding of the genetic basis of risk-taking behaviour. Despite this, we highlight some limitations. The risk-taking phenotype used was a self-reported measure, based on response to a single question, and is therefore open to responder bias. It is also plausible that there are distinct

subtypes of risk-taking behaviour (for example disinhibition, sensation-seeking and calculated risks). Whether the single question used in our analyses captures all, or only some, of these is not clear. It would be of interest to investigate whether the loci identified here are also associated with more quantitative and objective measures of risk taking; however, such measures were not available in the UK Biobank dataset.

Conclusion

In summary, we have identified a polygenic basis for self-reported risk-taking behaviour and two associated loci containing variants likely to play a role in predisposition to this complex but important phenotype. The identification of significant genetic correlations between risk-taking and several psychiatric disorders, as well as with smoking and obesity, suggest that future work on this trait may clarify mechanisms underlying several common psychopathological and physical health conditions, which are important for public health and wellbeing.

Acknowledgements

This research was conducted using the UK Biobank resource. UK Biobank was established by the Wellcome Trust, Medical Research Council, Department of Health, Scottish Government and Northwest Regional Development Agency. UK Biobank has also had funding from the Welsh Assembly Government and the British Heart Foundation. Data collection was funded by UK Biobank. DJS is acknowledges the support of the Brain and Behaviour Research Foundation (Independent Investigator Award 1930) and a Lister Prize Fellowship (173096). EMT is supported by a University Research Fellowship (UF140705) from the Royal Society. JC acknowledges the support of The Sackler Trust and is part of the Wellcome Trust funded Neuroimmunology of Mood and Alzheimer's consortium that includes collaboration with GSK, Lundbeck, Pfizer and Janssen & Janssen. The funders had no role in the design or analysis of this study, decision to publish, or preparation of the manuscript.

Conflict of interest

JPP is a member of UK Biobank advisory committee; this had no bearing on the study. No other conflicts of interest.

Supplementary information is available at Molecular Psychiatry's website.

References

1. de Haan L, Egberts AC, Heerdink ER. The relation between risk-taking behavior and alcohol use in young adults is different for men and women. *Drug Alcohol Depend* 2015; **155:** 222-227.

2. Kreek MJ, Nielsen DA, Butelman ER, LaForge KS. Genetic influences on impulsivity, risk taking, stress responsivity and vulnerability to drug abuse and addiction. *Nat Neurosci* 2005; **8**(11)**:** 1450-1457.

3. Reddy LF, Lee J, Davis MC, Altshuler L, Glahn DC, Miklowitz DJ *et al.* Impulsivity and risk taking in bipolar disorder and schizophrenia. *Neuropsychopharmacology* 2014; **39**(2)**:** 456-463.

4. McClelland J, Dalton B, Kekic M, Bartholdy S, Campbell IC, Schmidt U. A systematic review of temporal discounting in eating disorders and obesity: Behavioural and neuroimaging findings. *Neurosci Biobehav Rev* 2016; **71:** 506-528.

5. Cuthbert BN, Insel TR. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Med* 2013; **11:** 126.

6. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J *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* 2015; **12**(3)**:** e1001779.

7. UK Biobank. Genotype imputation and genetic association studies of UK Biobank, Interim Data Release. 2015; **11 September 2015**.

8. Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods* 2013; **10**(1)**:** 5-6.

9. Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. *G3 (Bethesda)* 2011; **1**(6)**:** 457-470.

10. Huang J, Howie B, McCarthy S, Memari Y, Walter K, Min JL *et al.* Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. *Nat Commun* 2015; **6:** 8111.

11. UK Biobank. Genotyping of 500,000 UK Biobank participants. Description of sample processing workflow and preparation of DNA for genotyping. 2015; **11 September 2015**.

12. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**(3)**:** 559-575.

13. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* 2016; **17**(1)**:** 122.

14. MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res* 2017; **45**(D1)**:** D896-D901.

15. Gene-Tissue Expression Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013; **45**(6)**:** 580-585.

16. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* 2015; **47**(11)**:** 1236-1241.

17. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015; **518**(7538)**:** 197-206.

18. Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Magi R *et al.* New genetic loci link adipose and insulin biology to body fat distribution. *Nature* 2015; **518**(7538)**:** 187-196.

19. Schizophrenia Psychiatric Genome-Wide Association Study Consortium. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* 2011; **43**(10)**:** 969-976.

20. Ruderfer DM, Fanous AH, Ripke S, McQuillin A, Amdur RL, Schizophrenia Working Group of Psychiatric Genomics C *et al.* Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. *Mol Psychiatry* 2014; **19**(9)**:** 1017-1024.

21. Major Depressive Disorder Working Group of the Psychiatric Genetics Consortium, Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM *et al.* A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry* 2013; **18**(4)**:** 497-511.

22. Neale BM, Medland SE, Ripke S, Asherson P, Franke B, Lesch KP *et al.* Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* 2010; **49**(9)**:** 884-897.

23. Tobacco Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 2010; **42**(5)**:** 441-447.

24. Duncan LE, Ratanatharathorn A, Aiello AE, Almli LM, Amstadter AB, Ashley-Koch AE *et al.* Largest GWAS of PTSD (N=20 070) yields genetic overlap with schizophrenia and sex differences in heritability. *Mol Psychiatry* 2017.

25. Otowa T, Hek K, Lee M, Byrne EM, Mirza SS, Nivard MG *et al.* Meta-analysis of genome-wide association studies of anxiety disorders. *Mol Psychiatry* 2016; **21**(10)**:** 1485.

26. Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* 2016; **533**(7604)**:** 539-542.

27. Sniekers S, Stringer S, Watanabe K, Jansen PR, Coleman JRI, Krapohl E *et al.* Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nat Genet* 2017.

28. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, DSM-5. American Psychiatric Association: Arlington, VA, 2013.

29. Pietri T, Easley-Neal C, Wilson C, Washbourne P. Six cadm/SynCAM genes are expressed in the nervous system of developing zebrafish. *Dev Dyn* 2008; **237**(1)**:** 233-246.

30. Tanabe Y, Fujita E, Hayashi YK, Zhu X, Lubbert H, Mezaki Y *et al.* Synaptic adhesion molecules in Cadm family at the neuromuscular junction. *Cell Biol Int* 2013; **37**(7)**:** 731-736.

31. Ibrahim-Verbaas CA, Bressler J, Debette S, Schuur M, Smith AV, Bis JC *et al.* GWAS for executive function and processing speed suggests involvement of the CADM2 gene. *Mol Psychiatry* 2016; **21**(2)**:** 189-197.

32. Davies G, Marioni RE, Liewald DC, Hill WD, Hagenaars SP, Harris SE *et al.* Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). *Mol Psychiatry* 2016; **21**(6)**:** 758-767.

33. Day FR, Helgason H, Chasman DI, Rose LM, Loh PR, Scott RA *et al.* Physical and neurobehavioral determinants of reproductive onset and success. *Nat Genet* 2016; **48**(6)**:** 617-623.

34. Broer L, Buchman AS, Deelen J, Evans DS, Faul JD, Lunetta KL *et al.* GWAS of longevity in CHARGE consortium confirms APOE and FOXO3 candidacy. *J Gerontol A Biol Sci Med Sci* 2015; **70**(1)**:** 110-118.

35. Berndt SI, Gustafsson S, Magi R, Ganna A, Wheeler E, Feitosa MF *et al.* Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat Genet* 2013; **45**(5)**:** 501-512.

36. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 2010; **42**(11)**:** 937-948.

37. Thoma P, Daum I. Comorbid substance use disorder in schizophrenia: a selective overview of neurobiological and cognitive underpinnings. *Psychiatry Clin Neurosci* 2013; **67**(6)**:** 367-383.

Figures and Legends

Figure 1: Results of a genome-wide association study of self-reported risk-taking behaviour. SNPs are plotted along the X axis by chromosome and position, with strength of association with self-reported risk-taking behaviour plotted on the Y axis. The red line indicates the threshold for GWAS significance ($P \leq 5e^{-8}$). Inset: QQ plot demonstrates deviation from null expectation (solid red line) of the GWAS results (black data points).

Figure 2: Regional plots for risk-taking-associated loci. A) Chr 3 main analysis results; B) results of analysis conditioned on chr 3 rs13084531; C) chr 6 main analysis results; D) results of analysis conditioned on chr 6 rs9379971. The index SNP is shown as a purple diamond.

Figure 1

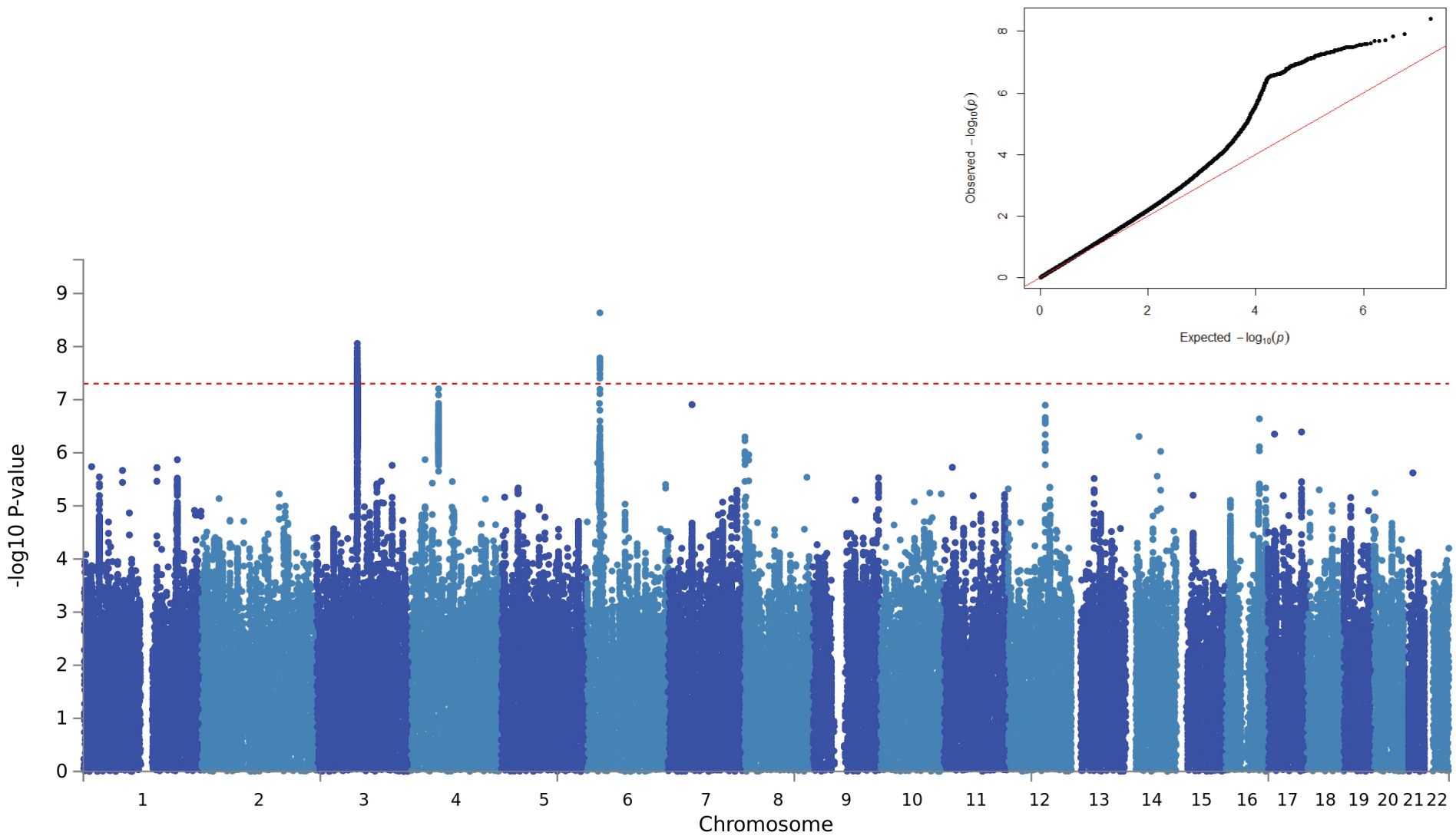
-log10 P-value
0
1
2
3
4
5
6
7
8
9
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Chromosome
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$
0
2
4
6
8

Figure 2

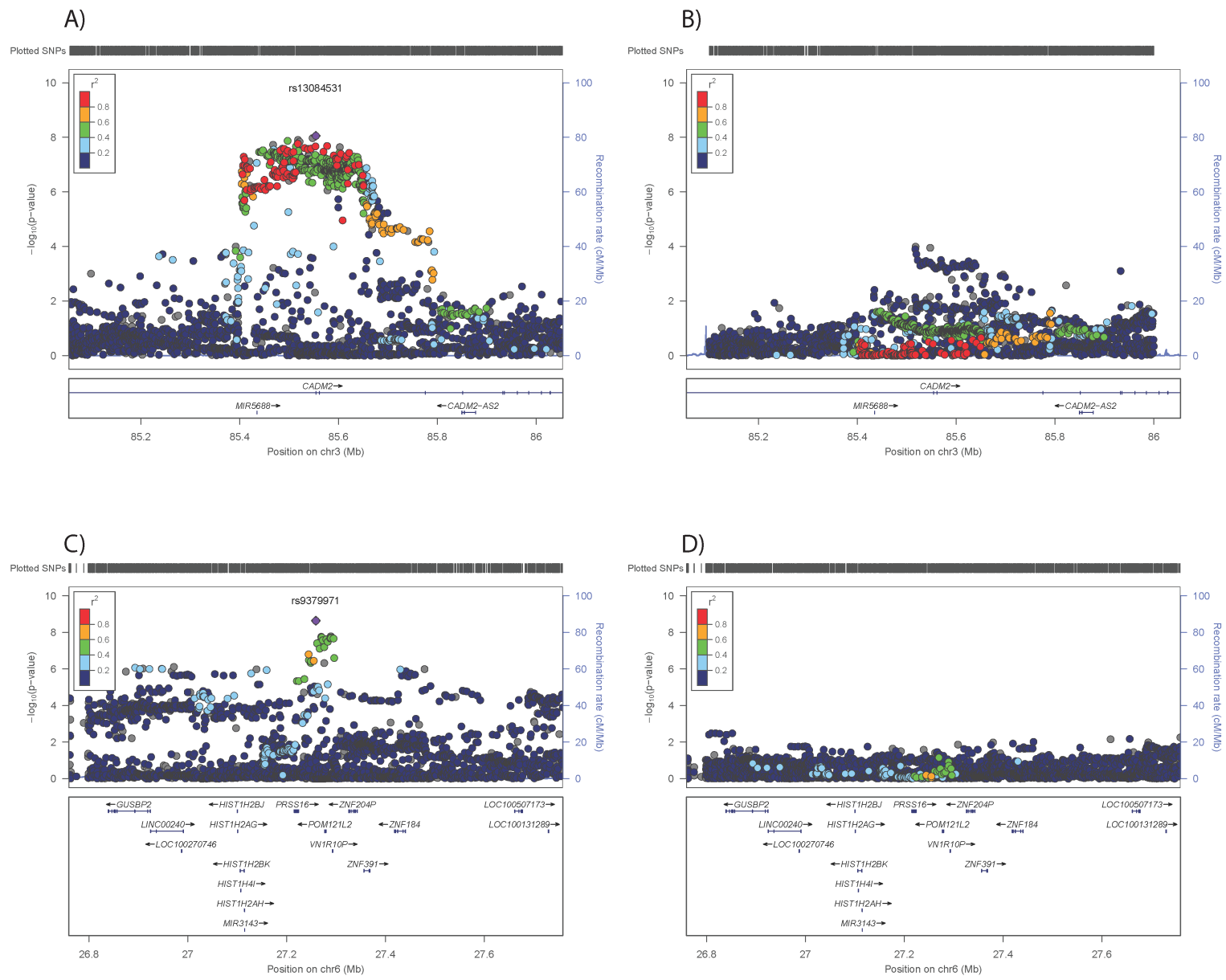
A)
B)
C)
D)
Plotted SNPs
rs13084531
rs9379971
r2
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
CADM2
MIR5688
CADM2−AS2
Position on chr3 (Mb)
GUSBP2
HIST1H2BJ
PRSS16
ZNF204P
LOC100507173
LINC00240
HIST1H2AG
POM121L2
ZNF184
LOC100131289
LOC100270746
VN1R10P
HIST1H2BK
ZNF391
HIST1H4I
HIST1H2AH
MIR3143
Position on chr6 (Mb)

Table 1. Description of UK Biobank participants included in the risk-taking GWAS

	Not risk takers	Risk takers‡	Full sample	Statistic	p
N	86 552	29 703	116 255		
N men	36 679 (42.4)	18 554 (62.5)	55 233 (47.5)	3600.0	**<0.001**
Age (years)	57.2 (7.8)	56.1 (8.1)	56.9 (7.9)	22.4	**<0.001**
BMI (kg/m2)	27.4 (4.9)	27.9 (4.7)	27.5 (4.8)	-13.7	**<0.001**
Current smoker	28 575 (33.1)	11 123 (37.5)	39 698 (34.2)	192.5	**<0.001**
Ever smoker	37 782 (43.8)	16 052 (54.2)	53 834 (46.4)	959.4	**<0.001**
Age completed education#	16.6 (2.1)	16.6 (2.3)	16.6 (2.2)	-4.0	**<0.001**
Has a degree	24 442 (28.5)	10 235 (34.7)	34 677 (30.1)	406.0	**<0.001**
Townsend deprivation index	-1.6 (2.9)	-1.3 (3.1)	-1.5 (3.0)	-15.5	**<0.001**
Unstable mood¤	37 429 (44.2)	14 258 (48.9)	51 687 (45.4)	189.3	**<0.001**
Comparison group*	17 024 (73.6)	5 418 (69.2)	22 442 (72.5)	55.9	**<0.001**
single episode depression*	1 615 (7.8)	519 (6.6)	2 134 (6.9)	1.6	0.209
Moderate depression*	2 816 (12.2)	1 034 (13.2)	3 850 (12.4)	16.4	**<0.001**
Severe depression*	1 486 (6.4)	678 (8.7)	2 164 (7.0)	75.8	**<0.001**
BD*	190 (0.8)	177 (2.3)	367 (1.2)	113.2	**<0.001**

‡ participants who answered "yes" to "do you consider yourself a risk taker?"; Test statistic, Pearsons Chi2 for categorical variables or t-test for continuous variables; # based on a subset of 80 229 subjects; ¤ paticipants who answered yes to ""Does your mood often go up and down?"; * definitions as per Smith et al, Plos One, 2013, based on a subset of 29,929 subjects; BD, bipolar disorder.

Table 2. Genetic correlation between risk-taking and traits relevant to psychiatric disorders

Phenotype	rg	se	p
ADHD	0.311	0.126	**0.0132**
SCZ	0.265	0.040	**4.54×10^{-11}**
BD	0.261	0.070	**1.73×10^{-4}**
MDD	0.069	0.084	0.4120
PTSD	0.513	0.165	**0.0018**
Anxiety (case control)	-0.090	0.132	0.4963
Anxiety (quantitative)	-0.123	0.156	0.4289
Fluid intelligence	-0.151	0.047	**0.0013**
Years of education	-0.023	0.033	0.4873
BMI	0.102	0.034	**0.0028**
WHRadjBMI	0.087	0.047	0.0655
Ever smoker	0.174	0.068	**0.0102**
Alcohol use disorder	0.221	0.306	0.4700

Where: MDD, major depressive disorder; BD, bipolar disorder; SCZ, schizophrenia; ADHD, attention deficit hyperactivity disorder; BMI, body mass index; WHRadjBMI; waist:hip ratio adjusted for BMI; Alcohol dependence DSM 5 Criteria; rg, regression coefficient; se, standard error of the regression coefficient; p, pvalue for theregression analysis.