# Genome-wide study of risk tolerance and risky behaviors reveals shared genetic influences

# Title: Genome-wide study of risk tolerance and risky behaviors reveals shared genetic influences [90 characters]

All authors are listed at the end of the manuscript.

**Abstract**:

Risk tolerance is an important variable in the behavioral and social sciences and one of the most studied phenotypes in social science genetics, but few genetic variants have so far been found to robustly associate with it or with risky behaviors. We conducted genome-wide association studies (GWAS) of general risk tolerance ($n$ = 939,908) and of six related phenotypes: adventurousness, and measures of risky behaviors in the driving, drinking, smoking, and sexual domains. We identified ~600 independent loci associated with the phenotypes, including 124 with general tolerance. We report evidence of substantial pleiotropy and estimate large genetic correlations—which exceed the corresponding measurement error-adjusted phenotypic correlations—between general risk tolerance and a range of risky behaviors (e.g., $\hat{r}_g = 0.50$ with a first principal component of measures of risky behaviors). Bioinformatic analyses imply that genes near general risk tolerance-associated SNPs are highly expressed in brain tissues and point to a role for glutamate and GABA neurotransmitters; we find no evidence of enrichment for genes that had previously been hypothesized to relate to risk tolerance. [173 words]

**One Sentence Summary:**

We identify ~600 independent loci associated with risk tolerance and risky behaviors and find evidence of substantial pleiotropy between these phenotypes [141 characters]

**Main Text:**

Risk pervades almost every aspect of human life, since many decisions involve tradeoffs between the amount of uncertainty and the magnitude of the expected payoff they entail. For instance, how car drivers deal with traffic risks, how individuals trade off the pleasure they derive from smoking and drinking alcohol with the associated health risks, and how entrepreneurs meet the challenges of business risks are decisive determinants of human well-being and of a society's innovation capacity. Risk tolerance—defined as the willingness to take risks to obtain greater rewards—is also an important variable in decision theory, finance, and economics.

Risk tolerance has been one of the most intensively studied phenotypes in social science genetics. Twin studies have established that risk tolerance (both self-reported and experimentally elicited) is moderately heritable, with estimates of its heritability ranging from 20% to 60%(*1*–*3*). To date, however, nearly all published studies attempting to discover the genetic variants associated with risk tolerance have been conducted in relatively small samples, ranging from a few hundred to a few thousand individuals (**Table S11.1)**, and still little is known about the genetic underpinnings of human risk taking behavior.

Here, we report results of what is by far the largest genetic study of risk tolerance and risky behaviors to date, in samples totaling over one million individuals. We report results of a genome-wide association study (GWAS) of self-reported 'general risk tolerance': the self-reported tendency or willingness to take risks in general. Self-reported general risk tolerance has been shown to be a good all-around predictor of risky behavior—such as portfolio allocation, occupational choice, smoking, drinking alcohol, and starting one's own business(*1*, *4*, *5*)—and to be highly correlated with a general factor of risk preference(*1*, *6*).

We also report results of GWAS of six supplementary phenotypes related to risk tolerance: 'adventurousness' (defined as the self-reported tendency to be adventurous vs. cautious); four risky behaviors—'automobile speeding propensity' (the tendency to drive faster than the speed limit), 'drinks per week' (the average number of alcoholic drinks consumed per week), 'ever smoker' (whether one has ever been a smoker), and 'number of sexual partners' (the lifetime number of sexual partners); and the first principal component (PC) of these four risky behaviors,

which we interpret as capturing the general tendency to take risks across domains. **Table 1** lists, for each GWAS, the datasets we analyzed and the GWAS sample size.

GWAS analyses

We followed a pre-specified analysis plan (available at https://osf.io/cjx9m/) for all seven GWAS. All analyses were performed for the autosomal SNPs and were restricted to individuals of European descent.

For self-reported general risk tolerance, we conducted a discovery GWAS ($n$ = 939,908) based on a sample-size-weighted meta-analysis of results from the UK Biobank (UKB, $n$ = 431,126) and from a sample of research participants from 23andMe ($n$ = 508,782). We also conducted a replication GWAS based on a meta-analysis of 10 other cohorts (total $n$ = 35,445). In the UKB cohort, we analyzed responses to the question: "Would you describe yourself as someone who takes risks? Yes / No." The exact phenotype measures vary across the 23andMe and replication cohorts in wording and number of response categories, but all measures are self-reported and broadly similar to the one analyzed in the UKB cohort (**Table S1.2)**.

Our discovery GWAS of general risk tolerance identified 124 approximately independent (pairwise $r^2 < 0.1$) genome-wide significant SNPs ("lead SNPs") (**Table S3.1**). **Fig. 1A** shows a Manhattan plot. The Q-Q plot of the discovery GWAS (before adjustment of the standard errors) (**Fig. S3.1A**) exhibits some inflation ($\lambda_{GC}$ = 1.41), as expected under polygenicity(*7*). Additional analyses suggest that population stratification is not a significant source of confounding bias. The lead SNPs' $R^2$'s are all less than 0.02%, and an additional risk-tolerance-increasing allele at the SNP with the largest estimated effect size increases general risk tolerance by ~0.026 standard deviations. The estimated effect sizes are almost all smaller than those of the lead SNPs of previous GWAS of selected phenotypes, including educational attainment(*8*) (**Fig. S3.2**).

To assess the credibility of these results, we compared them to our estimates from the replication GWAS. 123 of the 124 lead SNPs were directly available (or in high LD with a SNP) in the summary statistics of the replication GWAS. Of these 123 SNPs, 94 have a concordant sign ($P = 1.7\times10^{-9}$) and 23 of these 94 SNPs are significant at the 10% level ($P = 4.5\times10^{-8}$) (**Fig. S5.1**). To benchmark these results, we conducted a Bayesian analysis to estimate the posterior distribution

of the 123 SNPs' true effect sizes as well as their expected replication record(*9*). Our actual replication record very closely matches the estimated expected replication record.

Across our six supplementary GWAS, we identified a total of 741 lead SNPs distributed across ~550 approximately independent loci. **Table 1** reports statistics summarizing the results of the supplementary GWAS, **Fig. 1** (**B** to **G**) shows Manhattan plots, and **Fig. S3.1** (**C** to **H**) shows Q-Q plots, which also exhibit some inflation.

Genetic overlap

There is substantial overlap across the results of our seven GWAS. We identified a total of 831 different lead SNPs across the seven GWAS, but these are located across only ~611 approximately independent loci, and 34 of these lead SNPs were identified in two separate GWAS. In particular, 72 of the 124 general risk tolerance lead SNPs are in loci that also contain lead SNPs for at least one of the other GWAS, including 45 for adventurousness and 49 for at least one of the four risky behaviors or their first PC.

To further investigate genetic overlap, we used bivariate LD score regression(*10*) to estimate genetic correlations with self-reported general risk tolerance (using the summary statistics from the meta-analysis of our discovery and replication GWAS). The estimated genetic correlations with our six supplementary phenotypes are all positive, larger than ~0.25, and highly significant ($P < 2.3\times10^{-30}$; **Fig. 2**), indicating that SNPs associated with higher risk tolerance also tend to be associated with riskier behavior. Of note, general risk tolerance is highly genetically correlated with adventurousness ($\hat{r}_g = 0.83$, $P < 1\times10^{-100}$), automobile speeding propensity (0.45, $P = 1.21\times10^{-102}$), number of sexual partners (0.52, $P = 1.6\times10^{-171}$), and the first PC of the four risky behaviors (0.50, $P = 4.9\times10^{-167}$). We also estimated genetic correlations between general risk tolerance and eight additional risky behaviors (**Fig. 2**). The estimates are significant at the 5% level and in the expected direction for five of the eight additional risky behaviors (the other three estimates are not significant). In particular, general risk tolerance is moderately or highly genetically correlated with self-employment ($\hat{r}_g = 0.67$, $P = 0.01$), age at first sexual intercourse (-0.33, $P = 1.6\times10^{-25}$), and lifetime cannabis use (0.31, $P = 3.5\times10^{-8}$).

We also estimated genetic correlations between general risk tolerance and 20 additional phenotypes, including the Big Five personality traits and some cognitive, anthropometric, neuropsychiatric, and socioeconomic phenotypes (**Fig. 2** and in **Table S7.1**). The estimated genetic correlations are especially large for the personality traits extraversion ($\hat{r}_g$ = 0.51, $P = 5.5\times10^{-79}$), neuroticism (-0.42, $P = 8.6\times10^{-29}$), and openness to experience (0.33, $P = 4.9\times10^{-61}$). After Bonferroni correction for 34 tests, we also find significant genetic correlations with some neuropsychiatric phenotypes (including ADHD and schizophrenia) and some socioeconomic phenotypes.

Interestingly, the genetic correlations between general risk tolerance and the supplementary risky behaviors are substantially higher than the corresponding phenotypic correlations, even after adjustment of the phenotypic correlations for measurement error (**Tables S1.3** and **S7.1**). The correlated effects of genetic variants are thus an important contributor to phenotypic correlations between risk tolerance and risky behaviors. Taken together, these results suggest that a common factor for risk tolerance partially accounts for cross-domain variation in risky behavior(*1*, *6*, *11*), and that this factor is genetically influenced (although our results do not rule out some domain-specificity(*12*)).

Several regions of the genome stand out for being associated both with general risk tolerance and with all or most of the supplementary phenotypes, which further points toward the existence of substantial shared genetic influences on general risk tolerance and the six supplementary phenotypes. **Fig. 1B** and **Fig. S3.3** show local Manhattan plots for some of these. Of note, a long-range LD region(*13*) on chromosome 3 (~83.4 to 86.9 Mb) contains lead SNPs from all seven GWAS as well as the most significant lead SNP from the general risk tolerance GWAS, rs993137 ($P = 2.14\times10^{-40}$), which is located in the gene *CADM2*. *CADM2* has previously been found to be associated with a range of phenotypes, including multiple personality traits, age at menarche, BMI, educational attainment, and information processing speed(*14*). Further, a candidate inversion (i.e., a genomic region that is highly prone to inversion polymorphisms) on chromosome 18 (~49.1 to 55.5 Mb) contains lead SNPs from all seven GWAS and has previously been found to be associated with autism spectrum disorder, ADHD, depression, educational attainment, schizophrenia, and subcortical brain region volumes, among other phenotypes(*14*). Another long-range LD region, on chromosome 6 (~25.3 to 33.4 Mb), covers

the HLA-complex(*15*) and contains lead SNPs from all GWAS except drinks per week (for which we obtained a suggestive association ($P = 3.83\times10^{-7}$)). Finally, two other candidate inversions, on chromosomes 7 (~124.6 to 132.7 Mb) and 8 (~7.89 to 11.8 Mb), contain lead SNPs from six and five of our GWAS. The chromosome 8 region contains more than 500 genes and has previously been associated with several phenotypes including neuroticism(*9*), extraversion, schizophrenia, and chronotype. Importantly, the general risk tolerance-increasing alleles of the 29 general risk tolerance lead SNPs that are located in these five genomic regions are all associated with increased risk-taking in the summary statistics of all six of our supplementary GWAS.

To leverage the high degree of genetic overlap between general risk tolerance, adventurousness, and risky behaviors, we used Multi-trait Analysis of GWAS (MTAG)(*16*) to increase the precision of our estimates of the SNPs' effects on general risk tolerance. Using as input the summary statistics of our discovery and replication GWAS of general risk tolerance, of our supplementary GWAS (except for the first PC of the risky behaviors), as well as summary statistics from a GWAS of lifetime cannabis use(*17*), MTAG increased the number of general risk tolerance lead SNPs from 124 to 312 (**Fig. S9.1** and **Table S9.1**).

Biological annotation

We conducted a number of analyses to gain insights into the biological mechanisms through which genetic variants affect variation in self-reported general risk tolerance, using the summary statistics from the meta-analysis of our discovery and replication GWAS. First, we systematically reviewed the voluminous literature that has attempted to link risk tolerance to biological pathways. Our literature review covered both candidate genes studies (which test specific genetic variants as proxies for biological pathways) and studies employing other research designs, and identified 138 articles that matched our search criteria (**Table S11.1**).

Five main biological pathways have been tested by this literature: the steroid hormone cortisol, the monoamines dopamine and serotonin, and the steroid sex hormones estrogen and testosterone. Using a MAGMA(*18*) competitive gene-set analysis, we found no evidence that SNPs within genes associated with these five pathways tend to be more significantly associated with general risk tolerance than SNPs in other genes in the genome (**Table S11.3**). About 30 of

the articles identified by our literature review tested candidate genes in humans, most of which are within the dopamine or serotonin pathways. We verified that the SNPs analyzed in our summary statistics either coincide with or are likely to tag well most of the genetic variants used by the literature to test the 15 most commonly tested autosomal genes within those two pathways. We found no evidence the SNPs within those genes are particularly strongly associated with general risk tolerance (**Fig. 1C** and Table **S11.4**).

Next, to search in a hypothesis-free manner for genes that are significantly associated with general risk tolerance, we again used the software tool MAGMA(*18*), this time to perform a gene analysis for each of ~18,000 protein coding genes. After Bonferroni correction, 285 genes showed a significant association with general risk tolerance (**Fig. S12.1** and **Table S12.3**). We then used the Gene Network(*19*) co-expression database to gain insight into the functions and expression patterns of the 285 MAGMA genes. We also used the software tool DEPICT(*20*) to identify tissues in which genes near general risk tolerance-associated lead SNPs are highly expressed and to identify biological pathways associated with risk tolerance.

Both the Gene Network and the DEPICT analyses separately point to a role for glutamate and GABA neurotransmitters, which are the main excitatory and inhibitory neurotransmitters in the brain, respectively(*21*). Although glutamate neurotransmission has been implicated in schizophrenia(*22*) and major depression(*23*), to our knowledge no other published GWAS of cognition, personality, or neuropsychiatric phenotypes has pointed to a clear role for both glutamate and GABA. The relative balance between excitatory and inhibitory neurotransmission may thus be a relatively strong contributor to risk tolerance.

These results stand in sharp contrast to the lack of enrichment reported above for genes associated with cortisol, dopamine, serotonin, estrogen, and testosterone (and none of our bioinformatics analyses point to these pathways either). We note, however, that some brain regions identified in our analyses (discussed next) are areas where dopamine and serotonin play important roles.

The Gene Network and the DEPICT tissue enrichment analyses also both separately point to enrichment of the prefrontal cortex and the basal ganglia (**Tables S12.4** and **S12.6-7**). The cortical and subcortical regions highlighted by DEPICT (**Fig. 5B**) are some of the major components of the cortical-basal ganglia circuit, which is known as the reward system in human

and non-human primates, and is critically involved in learning, motivation and decision-making, notably under risk and uncertainty(*24*, *25*). We caution, however, that because the gene expression data used in the DEPICT analysis does not cover all brain regions, the results are not conclusive about the relative involvement of this circuit in risk tolerance.

Lastly, we used stratified LD score regression(*26*) to test for the enrichment of SNPs associated with histone marks in 10 selected tissue or cell types. Central nervous system tissues are the most enriched, accounting for 44% of the heritability while comprising only 15% of the SNPs (**Fig. 3A** and **Table S12.2**). Interestingly given prior evidence for involvement of the immune system in several neuropsychiatric disorders (*22*, *23*), immune/hematopoietic tissues are also significantly enriched.

Polygenic prediction

We constructed polygenic scores with summary statistics from meta-analyses of our discovery and replication GWAS of general risk tolerance (the meta-analyses excluded each validation cohort). We used the Add Health, HRS, NTR, STR and Zurich cohorts as validation cohorts. Our measure of predictive power is the incremental $R^2$ (or pseudo-$R^2$) from adding the score to a regression with controls for sex, birth year, and the top ten principal components of the genetic relatedness matrix.

Our preferred score, constructed with LDpred(*27*), explains ~1.0% of the variation in general risk tolerance, and up to ~1.4% of the variation in several alternative measures of risk tolerance (**Fig. S10.1** and **Table S10.1**). The score is also predictive of several personality phenotypes, including sensation seeking, behavioral inhibition and, consistent with our estimated genetic correlations and with previously reported phenotypic correlations(*1*, *28*), the Big Five personality traits openness to experience and extraversion (**Fig. S10.1** and **Table S10.2**).

To gauge the potential for using polygenic scores of general risk tolerance in empirical research in the behavioral sciences, we estimated the predictive power of the score for 20 real-world measures of risky behaviors in the health, financial, career, and other domains. Although most incremental $R^2$ estimates are low (**Fig. S10.2** and **Table S10.3**), several results stand out: in the STR cohort, a one-standard-deviation increase in our preferred LDpred score is associated with a

1.6 and a 3.4 percentage-point increase in the probability of being an entrepreneur and of having owned a business, respectively. A third score, constructed in the Add Health cohort with the general risk tolerance summary statistics outputted by our MTAG analysis, is associated with 2.5 additional lifetime sexual partners and with a 5.2 percentage-point increase in the probability of ever having smoked cannabis.

## Discussion

Researchers have developed numerous constructs and instruments to describe and measure various dimensions of risk tolerance(*29*). These range from psychologists' measures of sensation seeking and impulse control to economists' coefficient of relative risk aversion. Our use of simple, self-reported, and commonly-used measures of general risk tolerance allowed us to assemble a much larger sample than previous studies and to conduct the largest study on the genetics of risk tolerance and risky behaviors to date.

Importantly, and consistent with the phenotypic evidence(*1*, *4*, *5*), we nonetheless found strong evidence that our general risk tolerance measure shares substantial genetic etiology with a wide range of risky behaviors. Moreover, our bioinformatic analyses point to a role for glutamate and GABA neurotransmission and to involvement of specific brain regions whose likely role in decision-making has been documented by a large body of neuroscientific studies(*24*, *25*). As available GWAS samples sizes increase, we expect further convergence of results from genetic and neuroimaging studies of risk tolerance and decision-making.

Although our focus has been on the genetics of risky tolerance and of risky behaviors, we recognize that environmental and demographic factors account for much of these phenotypes' variation. Indeed, our estimates of the SNP heritabilities of general risk tolerance and the six supplementary phenotypes range from ~6% to 17% (**Fig. S6.1**), leaving ample room for environmental influences. In fact, we observe sizeable effects of gender and age on risk tolerance in the UKB data (**Fig. S1.1**), and there is abundant evidence that life experiences affect both measured risk preferences and risky behaviors(e.g., *35*, *36*). We speculate the impact of demographic factors and life experiences on risk tolerance and risky behaviors is partly moderated by genetic factors, and anticipate future research will uncover multiple instances of such gene-environment interactions.

**References and Notes:**


1. J. Beauchamp, D. Cesarini, M. Johannesson, The psychometric and empirical properties of measures of risk preferences. *J. Risk Uncertain.* (2017) (available at http://scholar.harvard.edu/jonathanpbeauchamp/publications/).
2. D. Cesarini, C. T. Dawes, M. Johannesson, P. Lichtenstein, B. Wallace, Genetic variation in preferences for giving and risk taking. *Q. J. Econ.* **124**, 809–842 (2009).
3. M. J. Zyphur, J. Narayanan, R. D. Arvey, G. J. Alexander, The genetics of economic risk preferences. *J. Behav. Decis. Mak.* **22**, 367–377 (2009).
4. T. Dohmen *et al.*, Individual risk attitudes: Measurement, determinants, and behavioral consequences. *J. Eur. Econ. Assoc.* **9**, 522–550 (2011).
5. A. Falk, T. Dohmen, A. Falk, D. Huffman, "The Nature and Predictive Power of Preferences : Global Evidence," *IZA Discussion Paper* (2015).
6. R. Frey, A. Pedroni, R. Mata, J. Rieskamp, R. Hertwig, Risk preference shares the psychometric structure of major psychological traits. *Sci. Adv.* **3**, e1701381 (2017).
7. J. Yang *et al.*, Genomic inflation factors under polygenic inheritance. *Eur. J. Hum. Genet.* **19**, 807–812 (2011).
8. A. Okbay *et al.*, Genome-wide association study identifies 74 loci associated with educational attainment. *Nature*. **533**, 539–542 (2016).
9. A. Okbay *et al.*, Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat. Genet.* **48**, 624–633 (2016).
10. B. K. Bulik-Sullivan *et al.*, An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
11. B. L. Einav, A. Finkelstein, I. Pascu, M. R. Cullen, How General Are Risk Preferences? Choices under Uncertainty in Different Domains. *Am. Econ. Rev.* **102**, 2606–2638 (2016).
12. E. U. Weber, A. E. Blais, N. E. Betz, A Domain-specific Risk-attitude Scale: Measuring Risk Perceptions and Risk Behaviors. *J. Behav. Decis. Mak. J. Behav. Dec. Mak.* **15**, 263–290 (2002).
13. A. L. Price *et al.*, Long-range LD can confound genome scans in admixed populations. *Am. J. Hum. Genet.* **83**, 132-135-139 (2008).
14. D. Welter *et al.*, The NHGRI GWAS Catalog, a curated resource of SNP-trait

associations. *Nucleic Acids Res.* **42**, D1001-1006 (2014).

15. S. T. Sherry *et al.*, dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–11 (2001).
16. P. Turley *et al.*, "MTAG: Multi-Trait Analysis of GWAS" (2017), , doi:https://doi.org/10.1101/118810.
17. S. Stringer *et al.*, Genome-wide association study of lifetime cannabis use based on a large meta-analytic sample of 32 330 subjects from the International Cannabis Consortium. *Transl. Psychiatry*. **6**, e769 (2016).
18. C. A. de Leeuw, J. M. Mooij, T. Heskes, D. Posthuma, MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput. Biol.* (2015), doi:10.1371/journal.pcbi.1004219.
19. R. S. N. Fehrmann *et al.*, Gene expression analysis identifies global gene dosage sensitivity in cancer. *Nat. Genet.* **47**, 115–125 (2015).
20. T. H. Pers *et al.*, Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6**, 5890 (2015).
21. O. A. C. Petroff, GABA and Glutamate in the Human Brain. *Neurosci.* **8**, 562–573 (2002).
22. S. Ripke *et al.*, Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. **511**, 421–427 (2014).
23. N. R. Wray *et al.*, "Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression" (2017), , doi:10.1101/167577.
24. S. N. Haber, B. Knutson, The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology*. **35**, 4–26 (2010).
25. P. N. Tobler, E. U. Weber, in *Neuroeconomics* (Elsevier, 2014; http://linkinghub.elsevier.com/retrieve/pii/B9780124160088000097), pp. 149–172.
26. H. K. Finucane *et al.*, Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
27. B. J. Vilhjálmsson *et al.*, Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *Am. J. Hum. Genet.* **97**, 576–592 (2015).
28. A. Becker, T. Deckers, T. Dohmen, A. Falk, F. Kosse, The relationship between economic preferences and psychological personality measures. *Annu. Rev. Econom.* **4**, 453–478 (2012).
29. K. P. Harden *et al.*, Beyond dual systems: a genetically-informed, latent factor model of

behavioral and self-report measures related to adolescent risk-taking. *Dev. Cogn. Neurosci.*, 1–14 (2016).

30. U. Malmendier, S. Nagel, Depression babies: do macroeconomic experiences affect risk taking? *Q. J. Econ.* **126**, 373–416 (2011).
31. C. R. Sahm, How Much Does Risk Tolerance Change? *Q. J. Financ.* **2**, 1250020 (2012).
32. B. J. Frey, D. Dueck, Clustering by passing messages between data points. *Science (80-. ).* **315** (2007).
33. H. Furberg *et al.*, Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat. Genet.* **42**, 441–447 (2010).

**Acknowledgments:** This research was carried out under the auspices of the Social Science Genetic Association Consortium (SSGAC). The research has also been conducted using the UK Biobank Resource under Application Number 11425. The study was supported by funding from the Ragnar Söderberg Foundation (E9/11 and E42/15), the Swedish Research Council (421-2013-1061), The Jan Wallander and Tom Hedelius Foundation, an ERC Consolidator Grant to Philipp Koellinger (647648 EdGe), the Pershing Square Fund of the Foundations of Human Behavior, and the NIA/NIH through grants P01-AG005842, P01-AG005842-20S2, P30-AG012810, and T32-AG000186-23 to NBER and R01-AG042568-02 to the University of Southern California. We thank the International Cannabis Consortium, the Psychiatric Genomics Consortium, and the eQTLgen Consortium for sharing with us summary statistics from the GWAS of lifetime cannabis use, summary statistics from the GWAS of ADHD, and eQTL summary statistics, respectively. A full list of acknowledgments is provided in **Supplementary Section 13**. Adam Auton, Pierre Fontanillas, David A Hinds, and Aaron Kleinman are employees of 23andMe; the authors declare no other competing financial interests. Upon publication, results can be downloaded from the SSGAC website (http://thessgac.org/data).

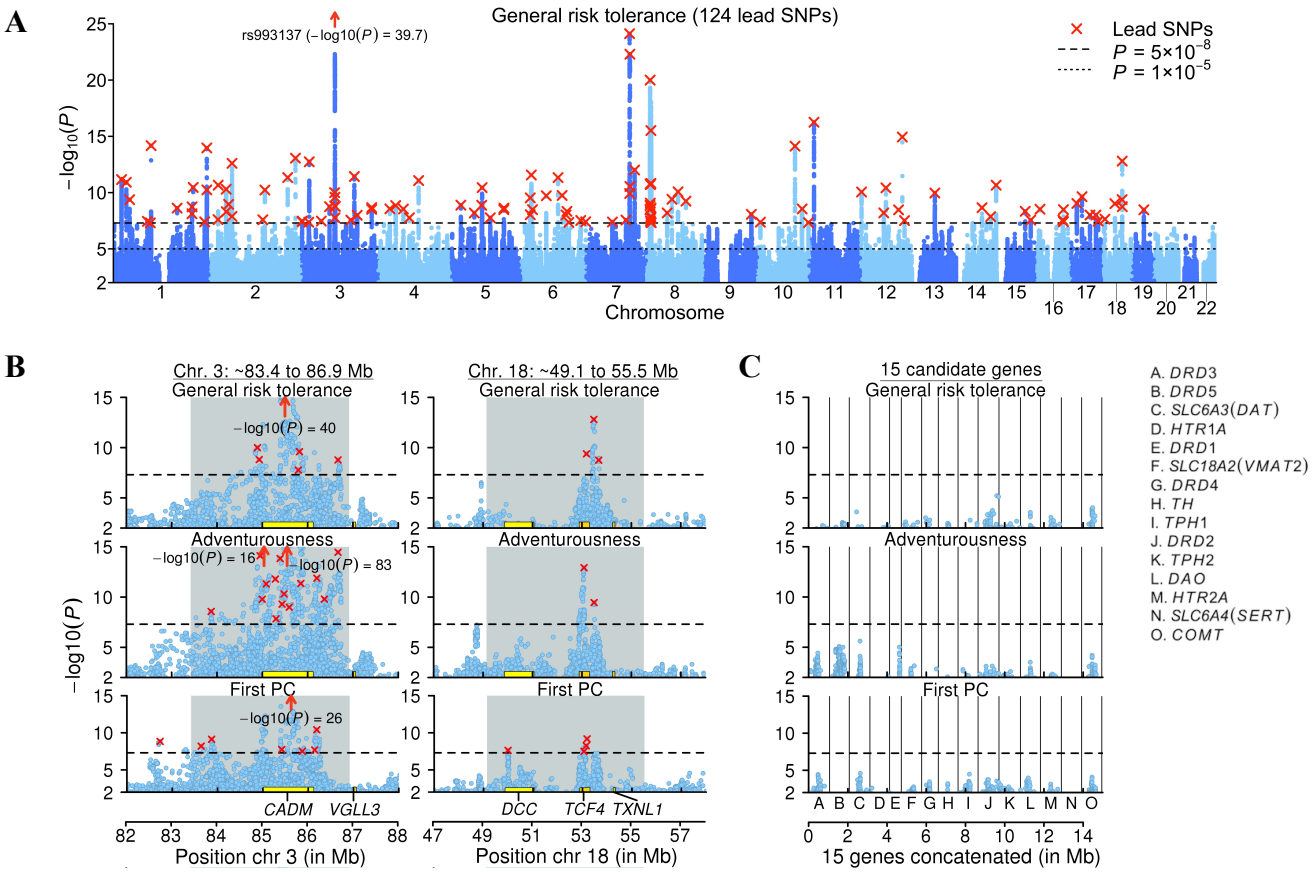


**Fig. 1. Manhattan plots.** In all panels, the *x*-axis is chromosomal position; the *y*-axis is the significance on a $-\log_{10}$ scale; the horizontal dashed line marks the threshold for genome-wide significance ($P = 5\times10^{-8}$); and each approximately independent genome-wide significant association ("lead SNP") is marked by a red ×. (**A**) Manhattan plots for the discovery GWAS of general risk tolerance. (**B**) Local Manhattan plots of two genomic regions that contain lead SNPs for all seven of our GWAS. The gray background marks the locations of candidate inversions or long-range LD regions. (**C**) Local Manhattan plots of the loci around the 15 most commonly tested candidate genes in the prior literature on the genetics of risk tolerance. Each locus comprises all SNPs within 500 kb of the gene's borders that are in LD ($r^2 > 0.1$) with a SNP in the gene. The 15 plots are concatenated and shown together in the panel, divided by the black vertical lines. The 15 genes are not particularly strongly associated with risk tolerance or the risky behaviors, as can be seen by comparing the results within each row across panels (**B**) and (**C**) (the three rows correspond to the GWAS of general risk tolerance, adventurousness, and the first PC of the risky behaviors).

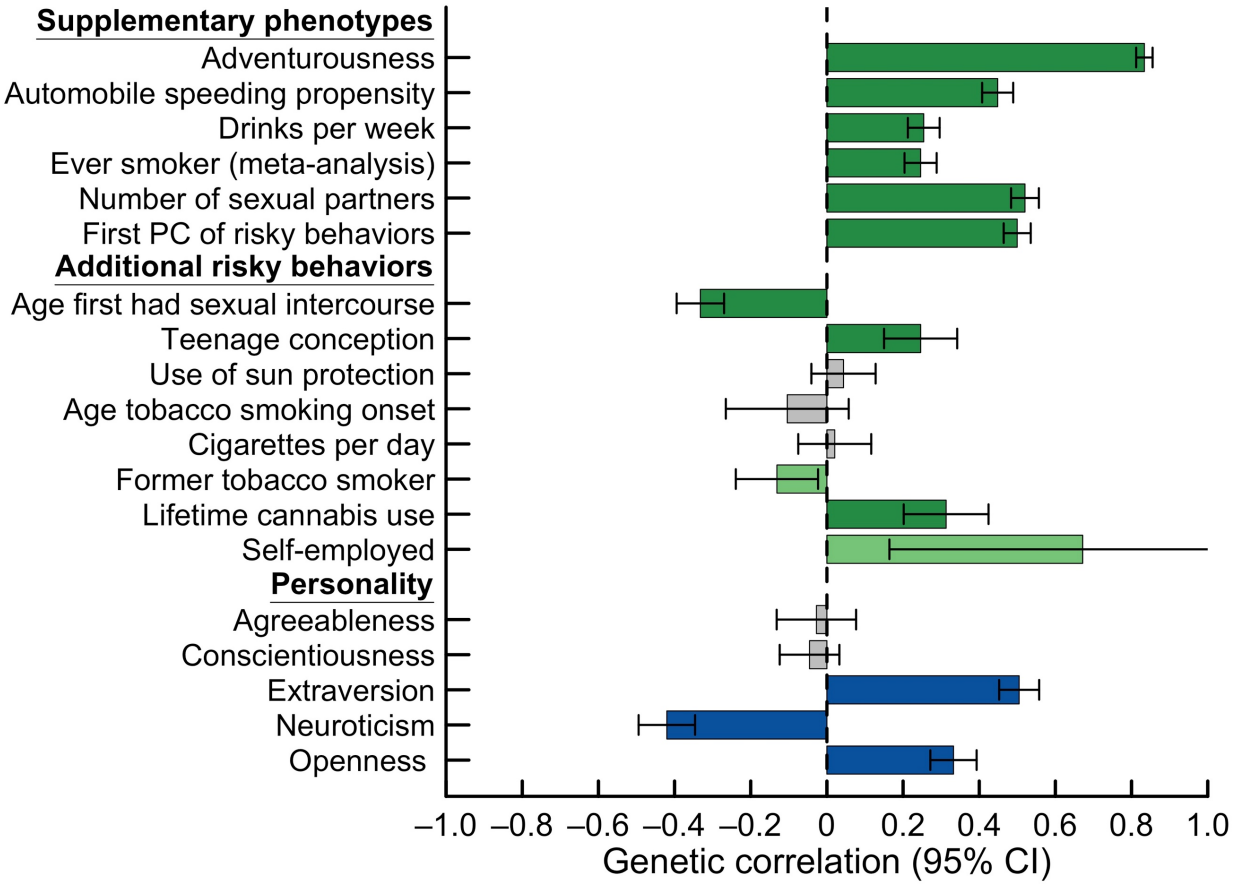


**Fig. 2. Genetic correlations with general risk tolerance.** The genetic correlations were estimated using bivariate LD score (LDSC) regression(*10*). Error bars show 95% confidence intervals. For the supplementary phenotypes and the additional risky behaviors, green bars represent significant estimates with the expected signs, where higher risk tolerance is associated with riskier behavior. For the personality phenotypes, blue bars represent significant estimates. Light green bars represent genetic correlations that are statistically significant at the 5% level, and dark green and blue bars represent correlations that are statistically significant after Bonferroni correction for 34 tests (the total number of phenotypes tested). Grey bars represent correlations that are not statistically significant at the 5% level.

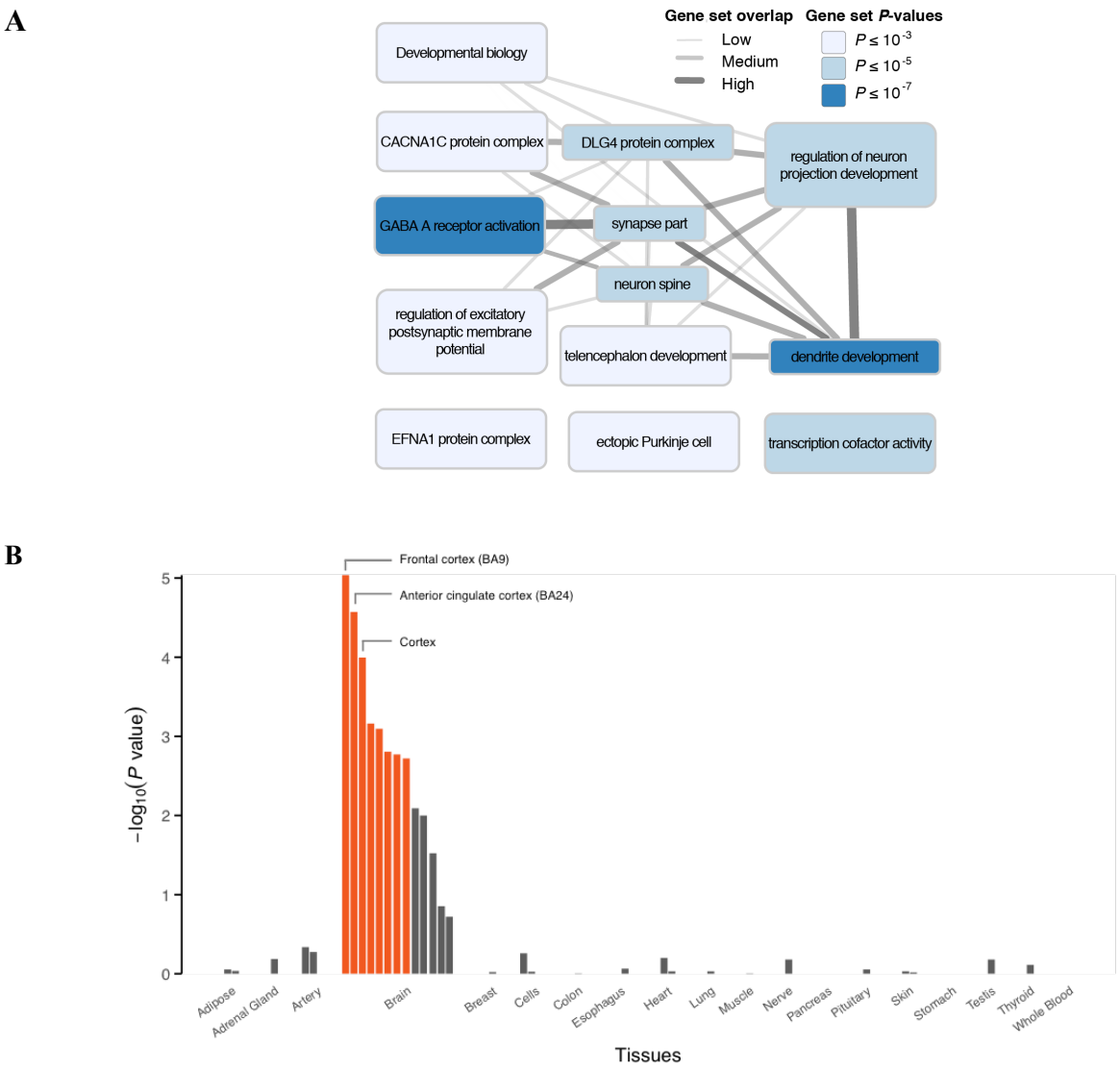


**Fig. 3. Results from selected biological analyses. (A)** DEPICT geneset enrichment diagram. We identified 93 reconstituted gene sets that are significantly enriched (FDR < 0.01) for genes overlapping loci associated with general risk tolerance; using the Affinity Propagation method(*32*), these were grouped into the 13 clusters displayed in the graph. Each cluster was named after the most significant gene set it contained, and each cluster's color represents the permutation *P* value of its most significant gene set. Overlap between the named representatives of two clusters is represented by an edge. Edge width represents the Pearson correlation $\rho$ between the two respective vectors of gene membership scores ($\rho < 0.3$, no edge; $0.3 \leq \rho < 0.5$, thin edge; $0.5 \leq \rho < 0.7$, intermediate edge; $\rho \geq 0.7$, thick edge). **(B)** Results of a DEPICT tissue enrichment analysis using GTEx RNA-sequencing gene expression data. The panel shows

whether the genes overlapping loci associated with general risk tolerance are significantly overexpressed (relative to genes in random sets of loci matched by gene density) in various tissues. Tissues grouped by organ or tissue type. The orange bars correspond to tissues with significant overexpression (FDR < 0.01). The *y*-axis is the significance on a $-\log_{10}$ scale.

**Table 1. GWAS results.**

| GWAS | Cohorts analyzed | *n* | Mean $\chi^2$ | No. lead SNPs |
|---|---|---|---|---|
| General risk tolerance (disc. GWAS) | UKB; 23andMe | 939,908 | 1.85 | 124 |
| General risk tolerance (rep. GWAS) | 10 indep. cohorts | 35,445 | 1.03 | 0 |
| General risk tolerance (disc + rep.) | UKB; 23andMe; 10 indep. cohorts | 975,353 | 1.87 | 132 |
| Adventurousness | 23andMe | 557,923 | 1.98 | 167 |
| Automobile speeding propensity | UKB | 404,291 | 1.53 | 42 |
| Drinks per week | UKB | 414,343 | 1.61 | 85 |
| Ever smoker | UKB; TAG Consortium(*33*) | 518,633 | 1.97 | 223 |
| Number of sexual partners | UKB | 370,711 | 1.77 | 118 |
| First PC of the four risky behaviors | UKB | 315,894 | 1.77 | 106 |

The table provides an overview of the GWAS of our primary and supplementary phenotypes. "disc.": discovery; "rep.": replication; "indep.": independent; "*n*": GWAS sample size; "Mean $\chi^2$": mean chi-squared statistics across HapMap3 SNPs with minor allele frequency greater than 0.01.

**Supplementary Materials:**

Materials and Methods

Supplementary Text

Figures S1.1-S12.3

Tables S1.1-S12.11

References (##-##)

**Authors:**

Richard Karlsson Linnér[1,2,3]*†, Pietro Biroli[4]†, Edward Kong[5]†, S Fleur W Meddens[1,2,3], Robbee Wedow[6,7,8]†, Mark Alan Fontana[9,10], Maël Lebreton[11,12],

Abdel Abdellaoui[13],Anke R Hammerschlag[1], Michel G Nivard[13], Aysu Okbay[3,14,15], Cornelius A Rietveld[14,2,15], Pascal N Timshel[16,17], Stephen P Tino[18], Maciej Trzaskowski[19], Ronald de Vlaming[3,14,15], Christian L Zünd[4],

Yanchun Bao[20], Laura Buzdugan[21,22], Ann H Caplin[23], Chia-Yen Chen[24,25], Peter Eibich[26,27], Pierre Fontanillas[28], Peter K Joshi[29], Ville Karhunen[30], Aaron Kleinman[28], Remy Z Levin[31], Christina M Lill[32,33], Gerardus A Meddens[34], Gerard Muntané[35], Sandra Sanchez-Roige[36], Juan Ramon Gonzalez[37,38,39], Frank J van Rooij[15], Erdogan Taskesen[1], Yang Wu[19], Futao Zhang[19],

*23andMe Research Team*, *eQTLgen Consortium*, *International Cannabis Consortium*, *Psychiatric Genomics Consortium*, *Social Science Genetic Association Consortium*,

Adam Auton[28], Jason D Boardman[8,7], David W Clark[29], Andrew Conlin[40], Conor C Dolan[13], Urs Fischbacher[41,42], Patrick JF Groenen[43,2], Kathleen Mullan Harris[44,45], Gregor Hasler[46], Albert Hofman[15,47], Mohammad A Ikram[15], Sonia Jain[48], Robert Karlsson[49], James MacKillop[50,51], Minna Männikkö[30], Carlos Morcillo-Suarez[35], Matthew B McQueen[52], Klaus M Schmidt[53], Melissa C Smart[20], Matthias Sutter[54,55], A Roy Thurik[14,2,56], Andre G Uitterlinden[57], Jon White[58], Harriet de Wit[59], Jian Yang[19,60],

Lars Bertram[32,61,62], Dorret Boomsma[13], Tõnu Esko[63], Ernst Fehr[22], David A Hinds[28], Magnus Johannesson[64], Meena Kumari[20], David Laibson[5], Patrik KE Magnusson[49], Michelle N Meyer[65], Arcadi Navarro[35,66], Abraham A Palmer[36], Tune H Pers[16,17], Danielle Posthuma[1,67], Daniel Schunk[68], Murray B Stein[36,48], Rauli Svento[40], Henning Tiemeier[15], Paul RHJ Timmers[29], Patrick Turley[69,25,70], Robert J Ursano[71], Gert G Wagner[72,27], James F Wilson[29,73],

Jacob Gratten[19], James J Lee[74],

David Cesarini[75], Daniel J Benjamin[9,70,76]*#, Philipp D Koellinger[3,77]*#, Jonathan P Beauchamp[18]*#

† Lead analysts; authors 2-5 listed alphabetically.
# These authors oversaw this work.
* Corresponding author. E-mail: jonathan.pierre.beauchamp@gmail.com (J.P.B.), r.karlssonlinner@vu.nl (R.K.L); daniel.benjamin@gmail.com (D.J.B.); p.d.koellinger@vu.nl (P.D.K.)

[1] Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Amsterdam Neuroscience, VU University Amsterdam, De Boelelaan 1085, Amsterdam, 1081 HV, the Netherlands
[2] Erasmus University Rotterdam Institute for Behavior and Biology, Erasmus School of Economics, Erasmus University Rotterdam, Burgemeester Oudlaan 50, Rotterdam, 3062 PA Rotterdam, the Netherlands
[3] Department of Economics, VU University Amsterdam, De Boelelaan 1105, Amsterdam, 1081 HV, the Netherlands
[4] Department of Economics, University of Zurich, Schönberggasse 1, Zurich, 8001, Switzerland
[5] Economics, Harvard University, 1805 Cambridge St, Cambridge, MA 2138, USA
[6] Institute for Behavioral Genetics, University of Colorado Boulder, 1480 30th Street, Boulder, CO 80309,USA
[7] Institute of Behavioral Science, University of Colorado Boulder, 1440 15th Street, Boulder, CO 80309, USA
[8] Department of Sociology, University of Colorado Boulder, UCB 327, Boulder, CO 80309, USA
[9] Center for the Advancement of Value in Musculoskeletal Care, Hospital for Special Surgery, 535 East 70th Street, New York, NY 10021, USA
[10] Center for Economic and Social Research, University of Southern California, 635 Downey Way, Los Angeles, CA 90089, USA
[11] Amsterdam School of Economics, University of Amsterdam, Roetersstraat 11, Amsterdam, 1018 WB
[12] Amsterdam Brain and Cognition, University of Amsterdam, Nieuwe Achtergracht 129B, Amsterdam, 1001 NK, Netherlands
[13] Biological psychology, Vrije Universiteit Amsterdam, van der Boechorstraat 1, Amsterdam, 1087 BT, Amsterdam
[14] Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, Burgemeester Oudlaan 50, Rotterdam, 3062 PA
[15] Department of Epidemiology, Erasmus Medical Center, Wytemaweg 80, Rotterdam, 3015 GE, The Netherlands
[16] Novo Nordisk Foundation Centre for Basic Metabolic Research, University of Copenhagen, Copenhagen, 2200, Denmark
[17] Department of Epidemiology Research, Statens Serum Institut, Copenhagen, 2300, Denmark
[18] Department of Economics, University of Toronto, 150 St. George St., Toronto, M5S 3G7, Ontario, Canada
[19] Institute for Molecular Bioscience, The University of Queensland, Brisbane, 4072, Australia
[20] Institute for Social & Economic Research, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK
[21] Seminar for Statistics, Department of Mathematics, ETH Zurich, Raemistrasse 101, Zurich, 8092, Switzerland
[22] Department of Economics, University of Zurich, Bluemlisalpstrasse 10, Zurich, 8006, Switzerland
[23] Stuyvesant High School, New York, 10282 NY, USA
[24] Psychiatry, Massachusetts General Hospital, Boston, MA 02114, USA
[25] Stanley Center for Psychiatric Genomics, The Broad Institute at Harvard and MIT, 75 Ames St, Cambridge, MA 02142, USA
[26] Nuffield Department of Population Health, University of Oxford, Old Road Campus, Oxford, OX3 7LF, UK
[27] Socio-Economic Panel Study, DIW Berlin, Mohrenstr. 58, Berlin, 10117, Germany
[28] Research, 23andMe, Inc., 899 W Evelyn Ave, Mountain View, CA 94041 USA
[29] Centre for Global Health Research, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Teviot Place, Edinburgh, EH8 9AG, Scotland
[30] Center for Life Course Health Research, University of Oulu, Aapistie 5, 90220 Oulu, Finland
[31] Department of Economics, University of California San Diego, La Jolla, CA 92093, USA
[32] Institute of Neurogenetics, University of Lübeck, Maria-Goeppert-Str. 1, Lübeck, 23562, Germany
[33] Institute of Human Genetics, University of Lübeck, Ratzeburger Allee 160, Lübeck, 23538, Germany
[34] Team Loyalty BV, Debbemeerstraat 25, Hoofddorp, 2131 HE, the Netherlands
[35] Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, Dr. Aiguader, 88, Barcelona, 8003, Spain
[36] Psychiatry, University of California San Diego, 9500 Gilman Drive, La Jolla, 92093, CA, USA
[37] Barcelona Institute for Global Health (ISGlobal), Barcelona, 8003, Spain
[38] Universitat Pompeu Fabra (UPF), Barcelona, 8003, Spain
[39] CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, 280229, Spain
[40] Oulu Business School, University of Oulu, Oulu, Finland
[41] Department of Economics, University of Konstanz, Universitätssstraße 10, Konstanz, 78457, Germany
[42] Thurgau Institute of Economics, Hauptstrasse 90, Kreuzlingen, 8280, Switzerland
[43] Department of Econometrics, Erasmus University Rotterdam, Rotterdam, the Netherlands

[44] Department of Sociology , University of North Carolina at Chapel Hill, 155 Hamilton Hall CB #3210, Chapel Hill, NC 27599, USA
[45] Carolina Population Center, University of North Carolina at Chapel Hill, 206 West Franklin St., Rm. 208, Chapel Hill, NC 27516, USA
[46] Division of Molecular Psychiatry, Translational Research Center, University Hospital of Psychiatry, University of Bern, Bolligenstrasse 111, Bern, 3000, Switzerland
[47] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA
[48] Family Medicine and Public Health, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA
[49] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Nobels väg 12a, Stockholm, 171 77, Sweden
[50] Psychiatry and Behavioural Neurosciences , McMaster University, 100 West 5th Street, Hamilton, L8P 3R2, Ontario, Canada
[51] Peter Boris Centre for Addictions Research, St. Joseph's Healthcare Hamilton, 100 West 5th Street, Hamilton, L8P 3R2, Ontario, Canada
[52] Department of Integrative Physiology, University of Colorado Boulder, 354 UCB, Boulder, CO 80309, USA
[53] Department of Economics, University of Munich, Ludwigstrasse 28, Munich, 80539, Germany
[54] Department of Economics, University of Cologne, Albert-Magnus-Platz, Cologne, 50923, Germany
[55] Experimental Economics Group, Max Planck Institute for Research into Collective Goods, Schumacher Strasse 10, D-51113 Bonn, Germany
[56] Montpellier Business School, Montpellier, France
[57] Internal Medicine, Erasmus Medical Center, Wytemaweg 80, Rotterdam, 3015 CN, the Netherlands
[58] UCL Genetics Institute, Department of Genetics, Evolution and Environment, University College London, Darwin Building, Gower Street, London, WC1E 6BT, UK
[59] Psychiatry and Behavioral Neuroscience, University of Chicago, 5841 S. Maryland Ave., MC 3077, Rm L485, Chicago, IL 60637, USA
[60] Queensland Brain Institute, The University of Queensland, Brisbane, 4072, Australia
[61] Institute of Integrative Genomics, University of Lübeck, Maria-Goeppert-Str. 1, Lübeck, 23562, Germany
[62] Faculty of Medicine, Imperial College London, London
[63] Estonian Genome Center, University of Tartu, Tartu, 51010, Estonia
[64] Department of Economics, Stockholm School of Economics, Box 6501, Stockholm, 113 83, Sweden
[65] Center for Translational Bioethics and Health Care Policy, Geisinger Health System, Danville, PN, USA
[66] Institució Catalana de Recerca i Estudis Avançats (ICREA), Parc de Recerca Bomèdica. Carrer Dr. Aiguader, 88, Barcelona, 8003, Spain
[67] Department of Clinical Genetics, VU Medical Centre, De Boelelaan 1085, Amsterdam, 1081 HV, The Netherlands
[68] Chair of Public Economics, Johannes Gutenberg University, Jakob-Welder-Weg 4, Mainz, 55128, Germany
[69] Analytical and Translational Genetics Unit, Massachusetts General Hospital, Richard B. Simches Research building, 185 Cambridge St, CPZN-6818, Boston, MA 02114, USA
[70] Behavioral and Health Genomics Center, University of Southern California, 635 Downey Way, Los Angeles, 90089 CA, USA
[71] Psychiatry, Uniformed Services University Health Science, Bethesda, MD, USA
[72] Max Planck Institute for Human Development, Lentzeallee 94, Berlin, 14195, Germany
[73] MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh, EH4 2XU, Scotland
[74] Department of Psychology, University of Minnesota Twin Cities, 75 East River Parkway, Minneapolis, 55455, MN, USA
[75] Department of Economics, New York University, 19 W 4th St, 6 FL, NY, 10012, USA
[76] National Bureau of Economic Research, 1050 Massachusetts Ave, Cambridge, MA 2138, USA
[77] German Institute for Economic Research, DIW Berlin, Mohrenstr. 58, Berlin, 10117, Germany