Supplement to *No support for historic candidate gene or candidate gene-by-interaction hypotheses for major depression across multiple large samples*

Contents

List of Tables

List of Figures

Part I
Supplemental methods

S1 Identification of top genes and polymorphisms

Table S1.1: Estimated lower bounds for the number of human association studies of each of the top 18 and 16 identified candidate genes and polymorphisms, respectively.

	Top candidate genes				*Top candidate polymorphisms*			
	Gene	$\widehat{N}^{\dagger}$	**95% CI**†		**Polymorphism**	**Gene**	$\widehat{N}^{\dagger}$	**95% CI**†
1.	*SLC6A4*	455	293 - 503	*1.*	5-HTTLPR	*SLC6A4*	404	235 - 487
2.	*BDNF*	154	103 - 171	*2.*	rs6265	*BDNF*	154	103 - 171
3.	*COMT*	93	58 - 112	*3.*	rs4680	*COMT*	93	58 - 112
4.	*HTR2A*	75	47 - 90	*4.*	rs6311	*HTR2A*	56	28 - 79
5.	*TPH1*	59	44 - 59	*5.*	rs1800532	*TPH1*	53	35 - 58
6.	*TPH2*	55	41 - 55	*6.*	VNTR	*DRD4*	42	28 - 26
7.	*DRD2*	50	33 - 55	*7.*	rs1800497	*DRD2*	28	14 - 42
8.	*MAOA*	50	32 - 60	*8.*	VNTR	*MAOA*	25	11 - 45
9.	*DRD4*	42	28 - 46	*9.*	ε-2/3/4	*APOE*	24	16 - 32
10.	*MTHFR*	32	24 - 32	*10.*	rs1801133	*MTHFR*	16	9 - 23
11.	*APOE*	24	16 - 32	*11.*	rs1801260	*CLOCK*	16	11 - 20
12.	*CLOCK*	23	19 - 26	*12.*	VNTR	*SLC6A3*	14	10 - 18
13.	*SLC6A3*	21	16 - 24	*13.*	in/del	*ACE*	11	9 - 14
14.	*ACE*	14	11 - 17	*14.*	rs1045642	*ABCB1*	8	8 - 9
15.	*DRD3*	11	11 - 11	*15.*	rs6280	*DRD3*	6	6 - 7
16.	*ABCB1*	11	10 - 11	*16.*	rs1611115	*DBH*	5	5 - 6
17.	*DTNBP1*	10	10 - 10					
18.	*DBH*	10	10 - 10					

†*Estimates reflect the hypergeometric parameter indicating the number of correctly identified studies among the finite population of studies identified by our algorithm (details provided below).*

The open source Biopython library [1] was used to scrape titles, abstracts, and metadata from the PubMed [2] database of published scientific journal articles. We do not claim to have exhaustively examined the entire candidate gene literature; rather, we have identified 18 highly-studied candidate genes in the context of human association studies of MDD and related endophenotypes. Our estimates of the number of studies per candidate gene reflect *lower bounds* for the true numbers of studies per candidate gene, Our procedure was as follows:

1. Titles of all meta-analyses matching the PubMed search `(psychology OR (psychiatry OR (neuroscience OR behavior))) AND topic = (allele OR (gene OR (polymorphism OR (genotype))))` were accessed
2. Regular expression matching was applied to determine potential gene names
3. Potential gene names occurred at least twice were hand-checked against gene names and aliases in the HUGO Gene Nomenclature Committee (HGNC) database of gene names [3, 4]
4. True matches were used to compile a dictionary of gene:[aliases] pairs using HGNC listed aliases E.g., `SLC6A4: [ obsessive-compulsive disorder 1, serotonin transporter, SERT1, 5-HTT, 5-HTTLPR, OCD1 ]`
5. Extracted titles and abstracts of all original research articles (as opposed to reviews/meta-analyses) published between 1991 and 2016 containing the terms `DEPRESSION, MDD, DEPRESSIVE, or DEPRESSANT` and the terms `PSYCHOLOGY, PSYCHOLOGICAL, PSYCHIATRIC, PSYCHIATRY, PSYCHOPATHOLOGY, PSYCHOPATHOLOGICAL, BEHAVIOR, BEHAVIORAL, COGNITIVE, COGNITION, or NEUROIMAGING`
6. Extensive *ad hoc* exclusion terms were applied to filter out irrelevant articles or articles not involving human subjects (e.g., "`clock drawing`", which refers to a neurocognitive assessment, was excluded as to avoid mismatches with the *CLOCK* gene)
7. Titles/abstracts containing a gene name or alias for each previously identified gene-names were tallied
8. A random subset of ten articles were checked by hand for correct identification. Then, to identify polymorphisms likely to have been studied at least ten times,
   (a) The optimal coverage exact 95% hypergeometric confidence interval [5] for the true of correctly identified articles was calculated. That is, for each gene, we had a finite population of M articles identified via the above procedure. Given a random sample of $m = 10$ identified articles containing k correctly estimated articles, we estimated the number $N: \ k \leq N \leq M$ of correctly identified articles among those identified via the fact that it is a hypergeometric random variable with mass
   $$P(N = n) = \frac{\binom{N}{k}\binom{M-N}{m-k}}{\binom{M}{m}}$$
   (b) If the 95% confidence interval $[N_{lower}, N_{upper}]$ excluded 9 or 11, i.e., if
   $$[N_{lower}, N_{upper}] \subset \{z \in \mathbb{N} : z \geq 10\} \vee [N_{lower}, N_{upper}] \subset \{z \in \mathbb{N} : z < 10\}$$
   no more samples were drawn and the genes likely to have been studied at least 10 times were retained
   (c) Else, an additional $k' := \max\{3, M - k\}$ samples were taken, $k := k + k'$, and we returned to the previous step
9. This algorithm resulted in 19 genes likely to have been studied at least 10 times. The *CACNA1C* gene, which in contrast to the other genes did not become popular until it had been implicated in a early GWAS of bipolar disorder [6], was excluded, leaving the 18 polymorphisms examined in the current investigation.
10. For each of the 18 retained genes, the individual polymorphisms studied in each of the previously sampled (correctly identified) studies were tallied (Figures S1.1, S1.2, S1.3).
11. Ad-hoc examination of the distributions of polymorphisms studied in each sample identified "top" polymorphisms in 16 of the 18 genes. There were no clear top polymorphisms in *TPH2* or in *DTNBP1*.
12. Hypergeometric parameters were estimated for each polymorphism.

We emphasize that we preregistered our analysis plan after identifying the top polymorphisms, but before running any of the association models. The top 18 candidate genes and the top 16 candidate polymorphisms are presented in table S1.1. Estimates of lower bounds for number of studies-of-genes-per-year are presented in figures S1.4 and S1.5.

Figure S1.1: Polymorphism counts in top 18 genes (1 of 3)

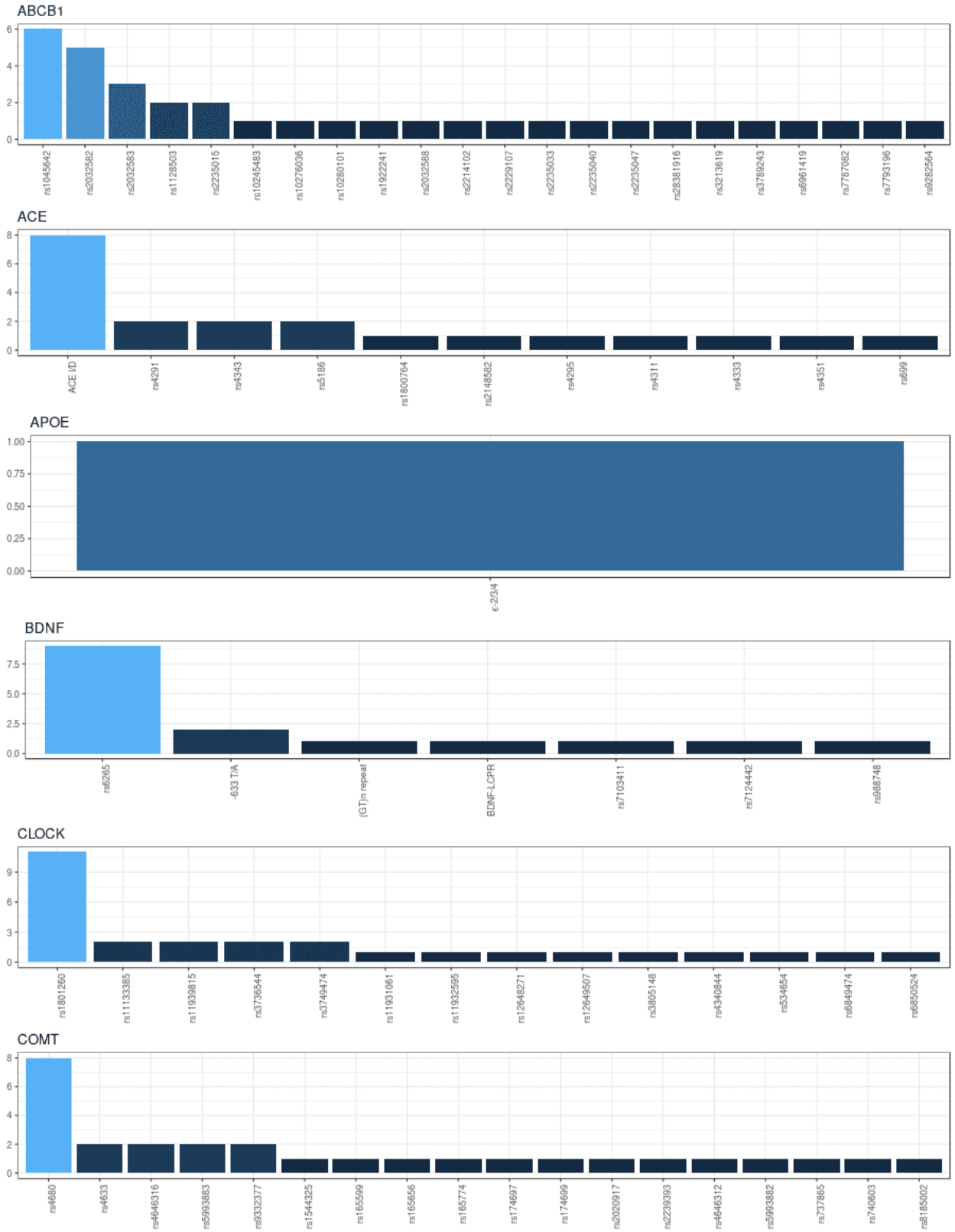


Figure S1.2: Polymorphism counts in top 18 genes (2 of 3)

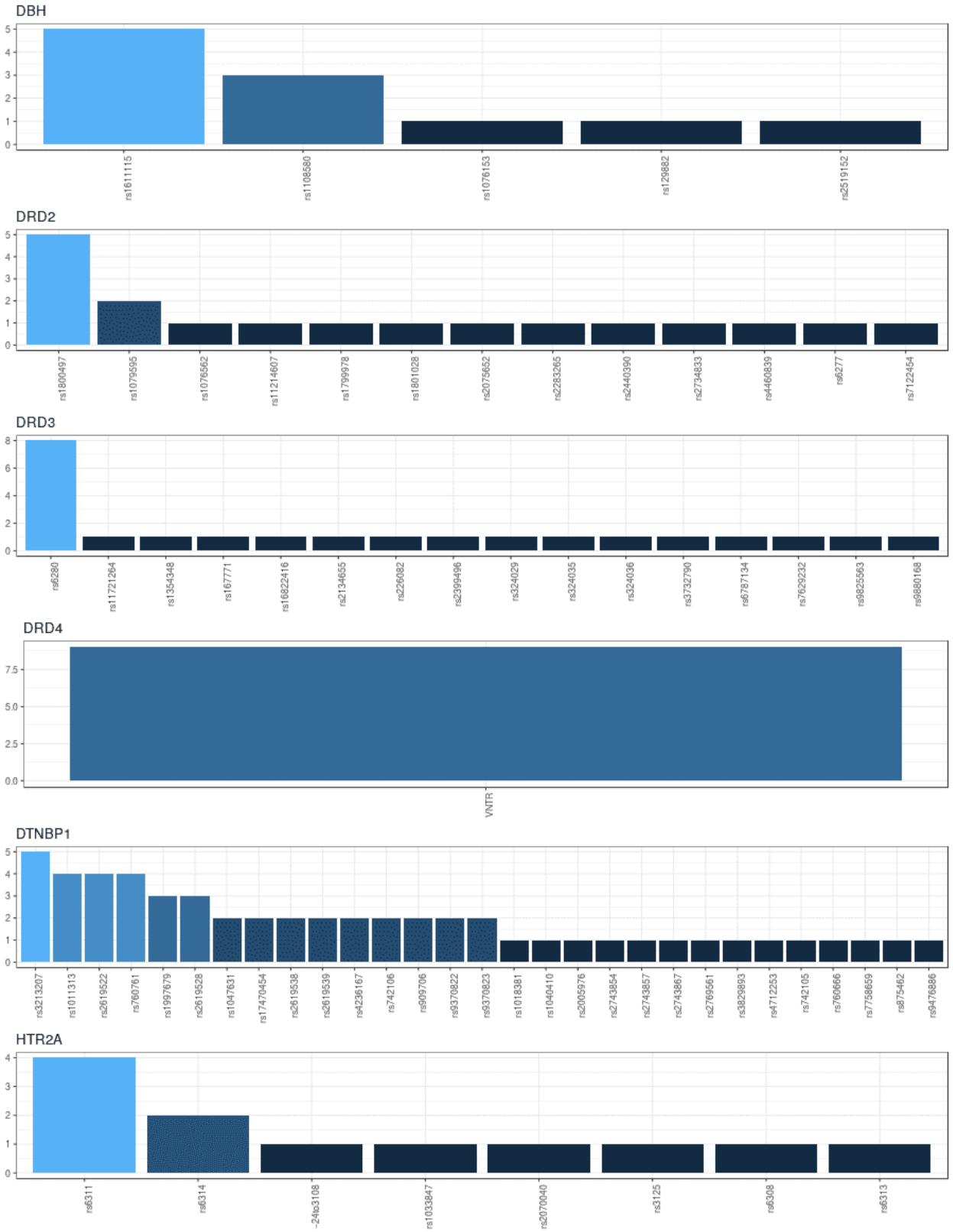
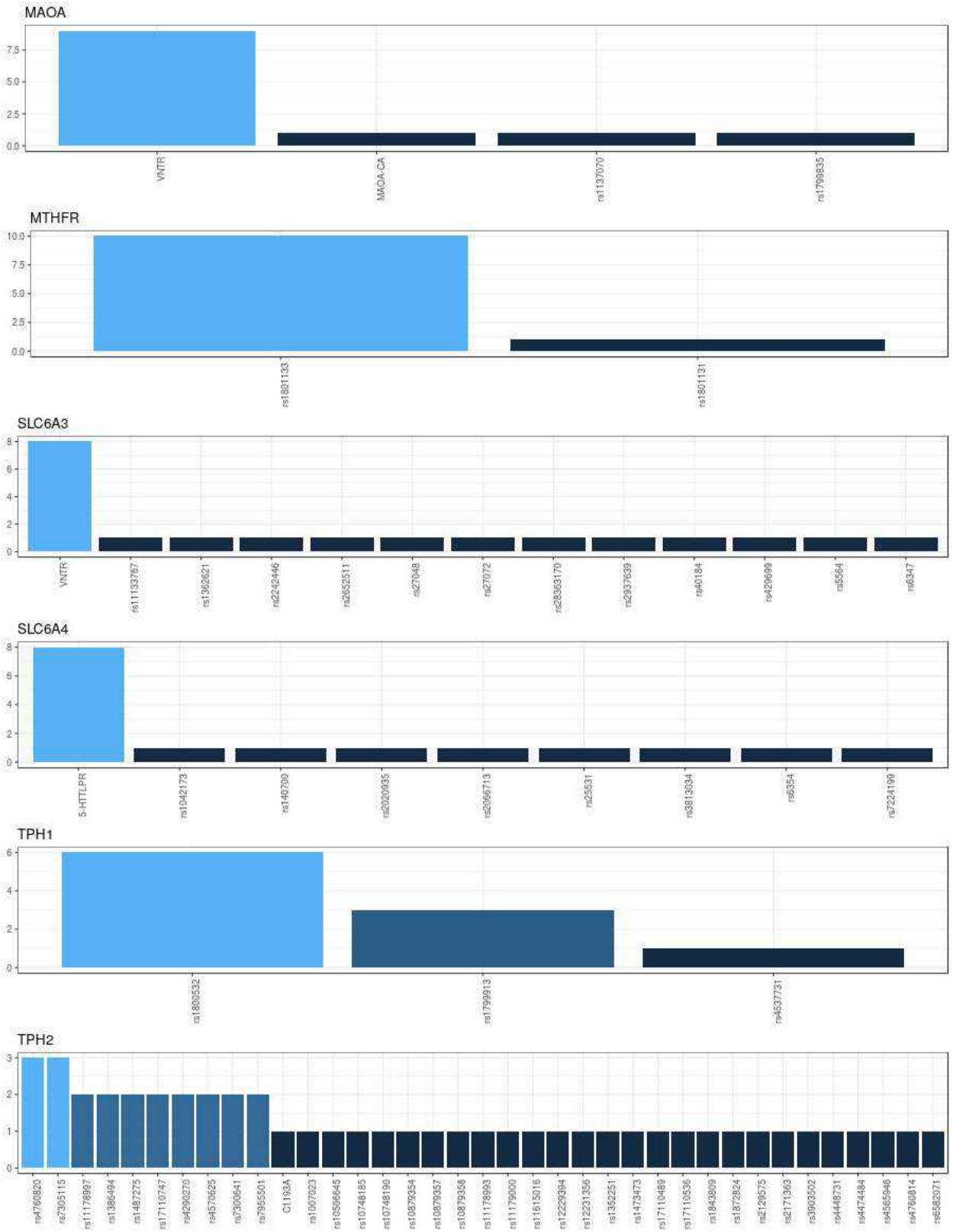


Figure S1.3: Polymorphism counts in top 18 genes (3 of 3)



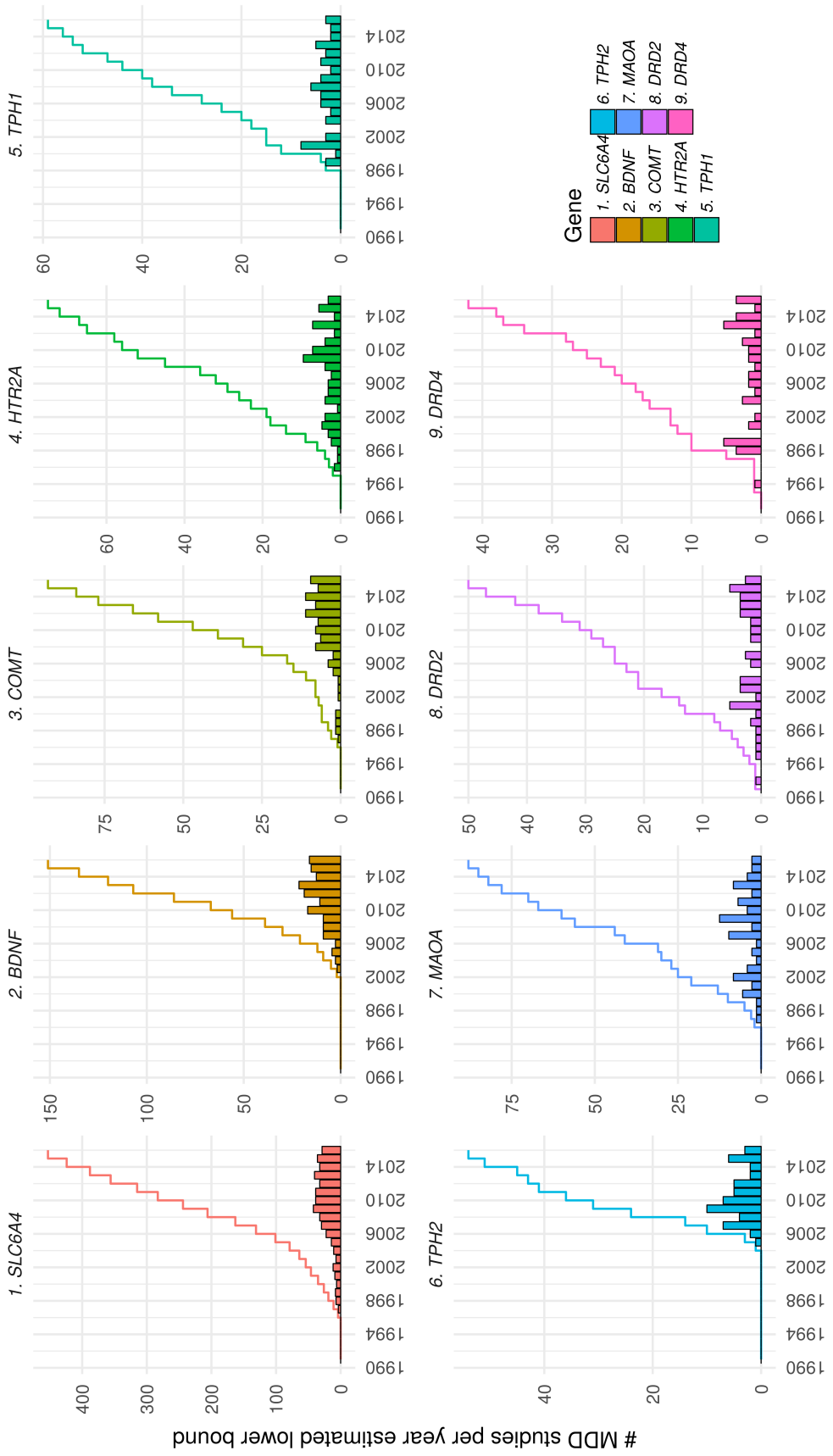


Figure S1.4: Cumulative sums of estimated lower bounds of studies-per-gene-per-year for top 18 genes (1 of 2)

Note. In contrast to the total sum, estimates of studies per individual year are highly sensitive to sampling error.

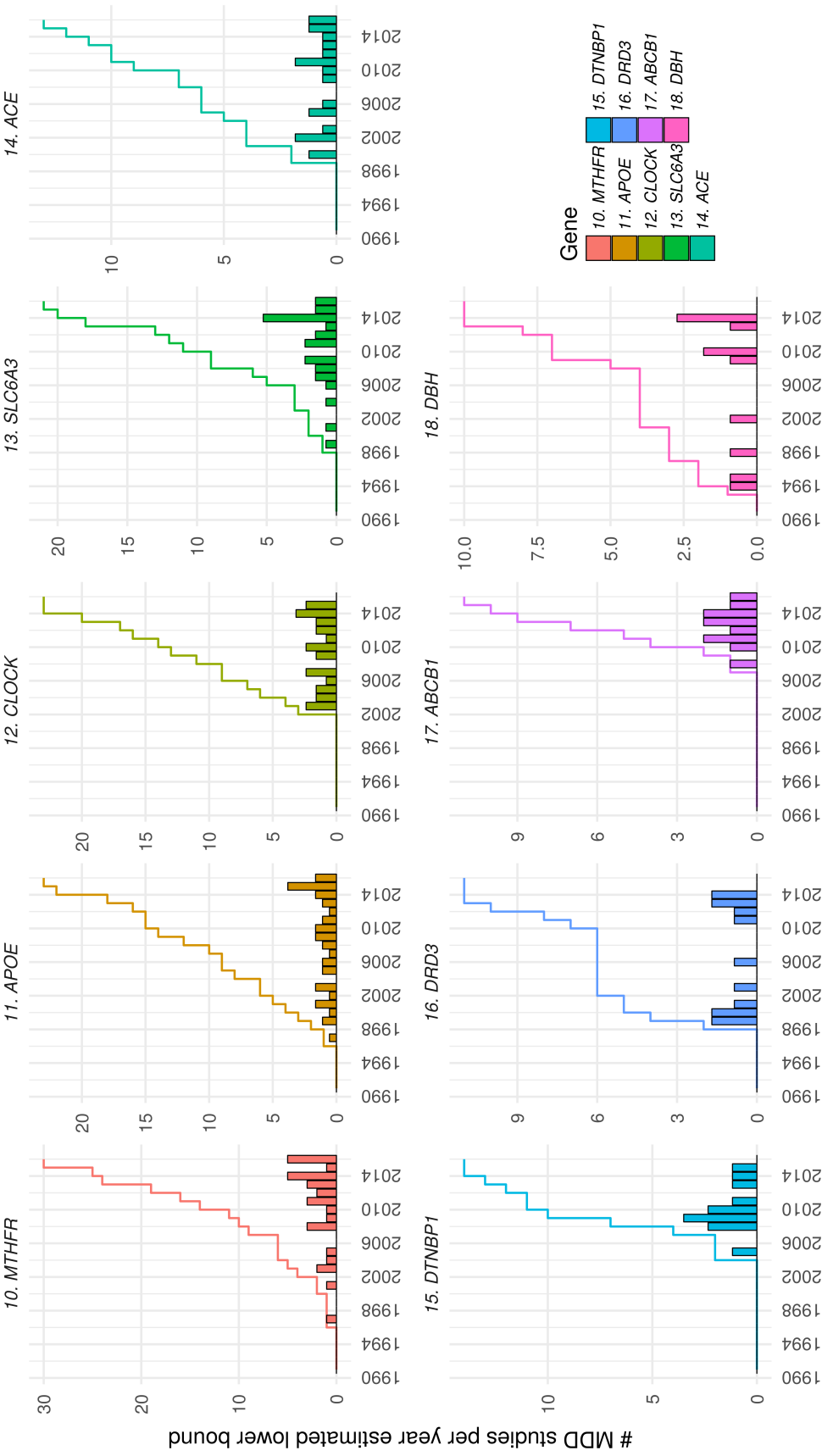


Figure S1.5: Cumulative sums of estimated lower bounds of studies-per-gene-per-year for top 18 genes (2 of 2)

Note. In contrast to the total sum, estimates of studies per individual year are highly sensitive to sampling error.

S2 Genotyping and quality control

S2.1 PGC sample

Only summary statistics (as opposed to raw genotype data) were used in the analyses of the Psychiatric Genetics Consortium (PGC) data. GWAS summary statistics from the 29 anchor cohorts and the deCODE, GERA, iPSYCH, and 23andMe expanded cohorts were meta-analyzed using the METAL software [7], using an inverse variance weighting scheme. SNPs were filtered at MAF > 0.01 and INFO > 0.8. Detailed descriptions of the genotyping and quality control procedures are provided in [8]. To prevent sample overlap, the UKBB and Generation Scotland cohorts were excluded, resulting in a total of 443,264 unrelated individuals (120,201 cases and 323,063 controls). Because raw genotype data were unavailable, only biallelic SNPs were included in polymorphism-level analyses (i.e., the triallelic *APOE ε-2/3/4* and all variable number tandem repeat [VNTR] polymorphisms were excluded). The *ACE* insertion/deletion (indel) polymorphism was determined via rs4343.

S2.2 UKBB sample

The details of the official UK Biobank genotyping, quality-control, and imputation methods in the released data can be found in Bycroft et al., 2017 [9]. We further excluded individuals with no genetic data and those whose self-reported and genetic sex conflicted (data fields `f.31.0.0` and `f.22001.0.0`), and those identified by the UK Biobank, UKBiLEVE, and Affymetrix with poor quality (`f.220010.0.0` and `f.22051.0.0`), for a total of 486,565 individuals. To reduce the influence of population stratification in our analyses, we only used individuals of primarily European ancestry. The UK Biobank identified individuals of "Caucasian" ancestry who self-identified as "British" (`f.22006.0.0`). To these individuals we added those whose first four principal component scores (from the UK Biobank-provided sample QC file) were within the range of the UK Biobank-identified "Caucasian" individuals.

In the array data we used plink v1.9 [10, 11] to LD- and MAF-prune markers with $|F_{het}| < 0.2$ in the European-ancestry sample (plink2 command: `--geno 0.05 --hwe 0.00000001 --maf 0.01 --indep-pairwise 50 5 0.2`), retaining 125,546 SNPs and 436,065 individuals.

Though UK-based cohorts of the PGC sample were excluded, further sample overlap was detected via the use of genetic checksums (software available here), resulting in the exclusion of an additional 338 individuals. We then estimated genomic relatedness matrices (GRMs; using the LD- and MAF-pruned array markers) separately for individuals for whom relevant items were measured in the initial touchscreen interview and those for whom relevant items were measured in the online mental health follow-up questionnaire, pruning the samples such that the maximum relatedness was 0.05 for any two individuals assessed on a given outcome. This resulted in two partially overlapping sub-samples (91,121 for the touchscreen interview outcomes and 115,458 for the online mental health follow-up outcomes) comprised of 177,950 unique individuals.

For gene-wise association analyses, we used the UK Biobank-imputed biallelic SNPs, applying the following thresholds: minor allele count of at least 3, Hardy-Weinberg p-value greater than 10^{-6}, no more than 2% missing calls, and imputation INFO score of at least 0.3, retaining only the HRC-imputed SNPs.

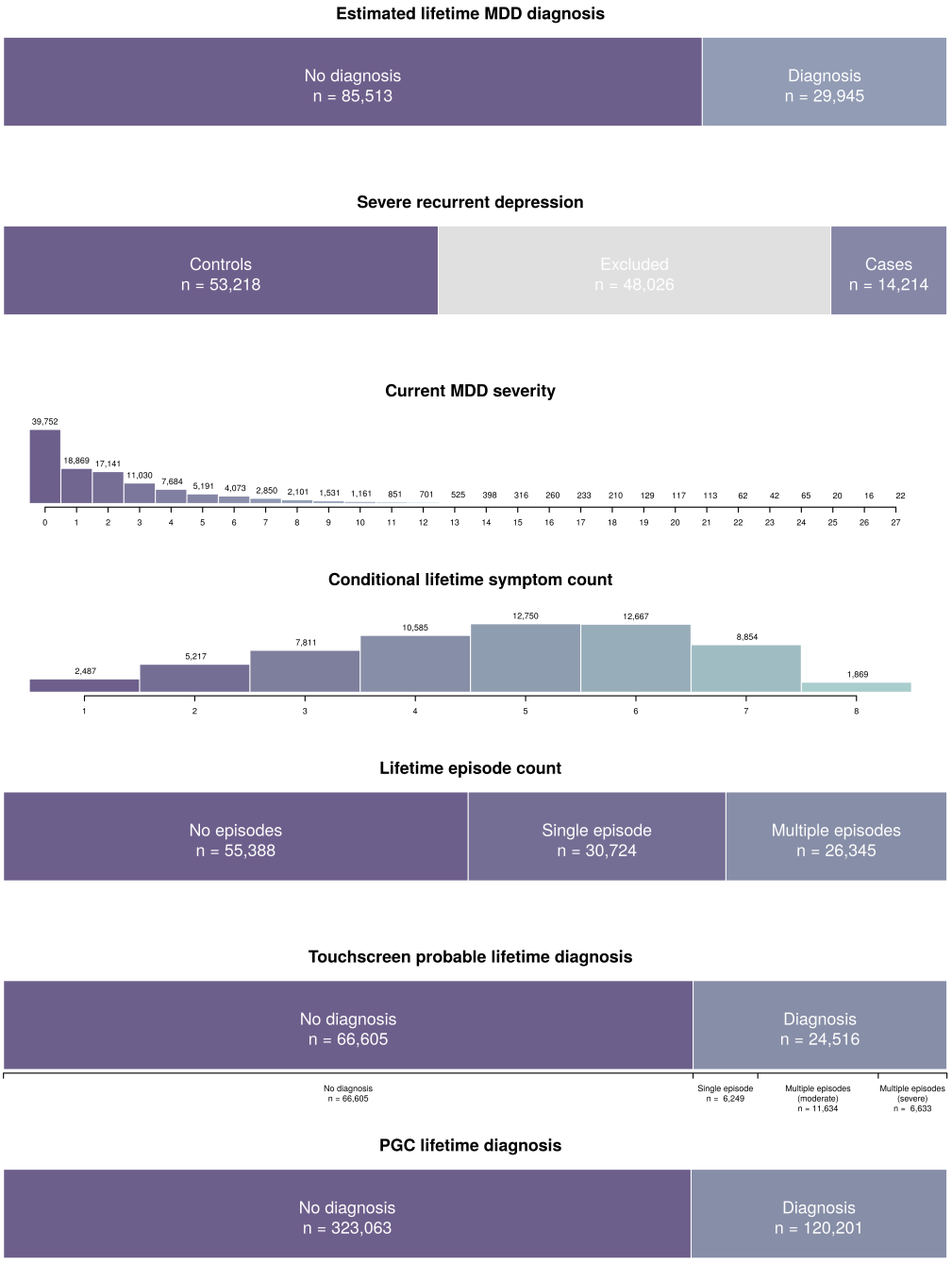
The VNTRs in *SLC6A4*, *DRD4*, *MAOA*, and *SLC6A3*, as well as a SNP in *SLC6A4* were imputed as detailed in [12]. 5-HTTLPR activity level was determined based on the number of repeats and genotype at *SLC6A4* rs25531 as described in [13] and the *ACE* insertion/deletion (in/del) polymorphism was determined via genotype at rs4343.

Hard calls, rather than dosages, were used in all association analyses.

S3 Measures

S3.1 Outcome measures

Figure S3.1: Distributions of depression phenotypes



The corresponding field number is listed next to each UKBB item. With the exception of *Probable MDD, ordinal (TSI)* and *Probable MDD diagnosis (TSI),* all psychiatric indices were derived by the authors, with details provided below.

S3.1.1 Estimated lifetime MDD Diagnosis

To receive a diagnosis, participants had to meet all of the following criteria:

1. **Anhedonia or depressed mood.** Respondents needed to have responded affirmatively to either of the following questions. Further criteria were assessed only among individuals endorsing one of the two below criteria.

   Anhedonia (20441) "Have you ever had a time in your life lasting two weeks or more when you lost interest in most things like hobbies, work, or activities that usually give you pleasure?"

   Mood (20446) "Have you ever had a time in your life when you felt sad, blue, or depressed for two weeks or more in a row?"

2. **Symptom count.** Respondents needed to endorse 4 or more of the following symptoms (unfortunately, motor agitation/retardation was not assessed) with respect to their worst period of depression:

   (a) **Anhedonia** (20441) "Have you ever had a time in your life lasting two weeks or more when you lost interest in most things like hobbies, work, or activities that usually give you pleasure?"

   (b) **Mood** (20446) "Have you ever had a time in your life when you felt sad, blue, or depressed for two weeks or more in a row?"

   (c) **Sleep** (20533, 20534, 20533) “Trouble falling asleep” or “sleeping too much” or “waking too early”

   (d) **Fatigue** (20449) "Did you feel more tired out or low on energy than is usual for you?"

   (e) **Appetite/weight** (20536) "Did you gain or lose weight without trying, or did you stay about the same weight?"

   (f) **Feelings of worthlessness** (20450) "People sometimes feel down on themselves, no good, worthless. Did you feel this way?"

   (g) **Concentration** (20435) "Did you have a lot more trouble concentrating than usual?"

   (h) **Ideation** (20437) "Did you think a lot about death - either your own, someone else's, or death in general?"

3. **Frequency.** (20439) With respect to their worst period of depression, respondents had to indicate a 2 or 3 on the following scale assessing “how often [they] felt this way”: {1: “Less often”, 2: “Almost every day”, 3: “Every day”}

4. **Fraction of day affected.** (20436) With respect to their worst period of depression, respondents had to indicate a 2 or greater on the following scale assessing “How much of the day did these feelings usually last?”: {1: “Less than half of the day”, 2: “About half of the day”, 3: “Most of the day”, 4: “All day long”}

5. **Impairment.** (20440) With respect to their worst period of depression, respondents had to indicate a 2 or greater on the following scale assessing "Think about your roles at the time of this episode, including study / employment, childcare and housework, leisure pursuits. How much did these problems interfere with your life or activities?" : {0: “Not at all”, 1: “A little”, 2: “Somewhat”, 3: “A lot”}

Individuals endorsing symptoms of mania, hallucinations, or delusions were excluded from analyses.

S3.1.2 Conditional lifetime symptom count

Estimated MDD symptom count among UKBB online mental health follow-up respondents who endorsed anhedonia and/or depressed mood.

As the remaining symptoms described above were assessed conditionally upon endorsement of anhedonia or depressed mood, we only present symptom counts among individuals endorsing either of the two “threshold criteria”. Symptom counts thus ranged from 1 to 8.

Individuals endorsing symptoms of mania, hallucinations, or delusions were excluded from analyses.

S3.1.3 Current MDD severity

Severity of depression symptoms over the past two weeks leading up to assessment among UKBB online mental health follow-up respondents.

Measurement of this outcome differs from that of estimated MDD diagnosis in the following respects:

- Respondents ranked the severity of each symptom on a 0-4 scale rather than providing a binary endorsement
- All symptoms were assessed regardless of whether or not anhedonia or depressed mood was diagnosed.

The following items were assessed as "Over the last 2 weeks, how often have you been bothered by any of the following problems?"

1. **Anhedonia** (`20510`) "Little interest or pleasure in doing things"
2. **Mood** (`20514`) "Feeling down, depressed, or hopeless"
3. **Sleep** (`20517`) "Trouble falling or staying asleep, or sleeping too much"
4. **Fatigue** (`20519`) "Did you feel more tired out or low on energy than is usual for you?"
5. **Appetite/weight** (`20511`) "Poor appetite or overeating"
6. **Feelings of worthlessness** (`20507`) "Feeling bad about yourself or that you are a failure or have let yourself or your family down"
7. **Concentration** (`20508`) "Trouble concentrating on things, such as reading the newspaper or watching television"
8. **Ideation** (`20513`) "Thoughts that you would be better off dead or of hurting yourself in some way"

Individuals endorsing symptoms of mania, hallucinations, or delusions were excluded from analyses.

S3.1.4 Lifetime episode count

Ordinal measure of lifetime number of depressive episodes among UKBB online mental health follow-up respondents who endorsed anhedonia and/or depressed mood (`20442`).

Individuals who endorsed a two week period of either anhedonia or depressed mood were asked "How many periods did you have in your life lasting two or more weeks where you felt like this?". Respondents supplied either an integer between 1 and 999 or responded "Too many to count / One episode ran into the next", rendering counts greater than one difficult to compare. We thus assigned scores as follows:

0	individuals who endorsed neither anhedonia or depressed mood
1	individuals who indicated a single depressive episode
2	individuals who indicated ≥ 2 depressive episodes

Individuals endorsing symptoms of mania, hallucinations, or delusions were excluded from analyses.

S3.1.5 Touchscreen probable lifetime diagnosis, ordinal classification

Ordinal measure of lifetime depression diagnostic status assessed as part of the UKBB initial touchscreen interview (`20126`).

This measure has been extensively studied and is described in great detail in Smith et al., 2013 [14]. Additionally, further details are provided in http://biobank.ctsu.ox.ac.uk/crystal/docs/MentalStatesDerivation.pdf. Briefly, a selection items of items from the Patient Health Questionnaire [15], the Structured Clinical Interview for DSM-IV Axis I Disorders-Research Version [16, 15], and items assessing treatment seeking behavior specific to the UKBB touchscreen interview. Response were classified as follows:

0 No bipolar or depression

1 Single probable major depression episode

2 Probable recurrent major depression (moderate)

3 Probable recurrent major depression (severe)

Individuals likely to meet criteria for bipolar disorder were excluded from analyses.

S3.1.6 Touchscreen probable lifetime diagnosis

Binary measure of probable depression diagnostic status assessed as part of the UKBB initial touchscreen interview (`20126`).

The three non-zero categories of the previous outcome were collapsed to create a dichotomous indicator of probable lifetime diagnosis. Individuals likely to meet criteria for bipolar disorder were excluded from analyses.

S3.1.7 Severe recurrent depression

Binary indicator of severe recurrent MDD versus no lifetime endorsement of depressed mood and anhedonia among individuals assessed in the UKBB online mental health follow-up. This measure was utilized in a follow-up sensitivity analysis to ensure our results would not differ dramatically if a stricter case/control assignment procedure were employed.

Scores were assigned as follows:

0 Neither the lifetime anhedonia or depressed mood items described in S3.1.1 were endorsed.

1 Lifetime estimated diagnosis criteria (S3.1.1) were met, five or more of the lifetime symptoms (one of which needed to be anhedonia or depressed mood; S3.1.2) were endorsed, and two or more lifetime episodes (S3.1.4) were endorsed

All other participants were excluded, as were individuals endorsing symptoms of mania, hallucinations, or delusions.

S3.1.8 PGC lifetime MDD diagnosis

Binary indicator of MDD diagnostic status (see [8] for further details and exclusion criteria). The current investigation utilized data from the full expanded cohort meta-analysis, excepting UK-based cohorts (UKBB and Generation Scotland).

Table S3.1: Bivariate correlations between UKBB depression phenotypes

Current MDD severity	0.381[1]					
Lifetime episode count	0.787[2]	0.395[1]				
Conditional lifetime symptom count	0.776[1]	0.328[3]	0.371[2]			
Touchscreen probable lifetime diagnosis, ordinal	0.701[2]	0.344[1]	0.720[2]	0.464[1]		
Touchscreen probable lifetime diagnosis	0.760[2]	0.316[1]	0.695[2]	0.451[1]	†	
Severe recurrent depression	†	0.660[2]	‡	‡	0.896[2]	0.714[2]
	Estimated lifetime MDD diagnosis	*Current MDD severity*	*Lifetime episode count*	*Conditional lifetime symptom count*	*Touchscreen probable lifetime diagnosis, ordinal*	*Touchscreen probable lifetime diagnosis*

[1]Polychoric, [2]polyserial, and [3]Pearson correlations between depression phenotypes in the UKBB based on pairwise complete observations. Touchscreen items were measured at a different time point than the other items. †Severe recurrent depression is a refinement of estimated lifetime MDD diagnosis, as is the ordinal classification of touchscreen probable lifetime diagnosis with respect to its binary counterpart–the agreement of these variables is necessarily perfect. ‡By definition, controls for severe recurrent depression had zero lifetime symptoms (or episodes) and cases had multiple episodes.

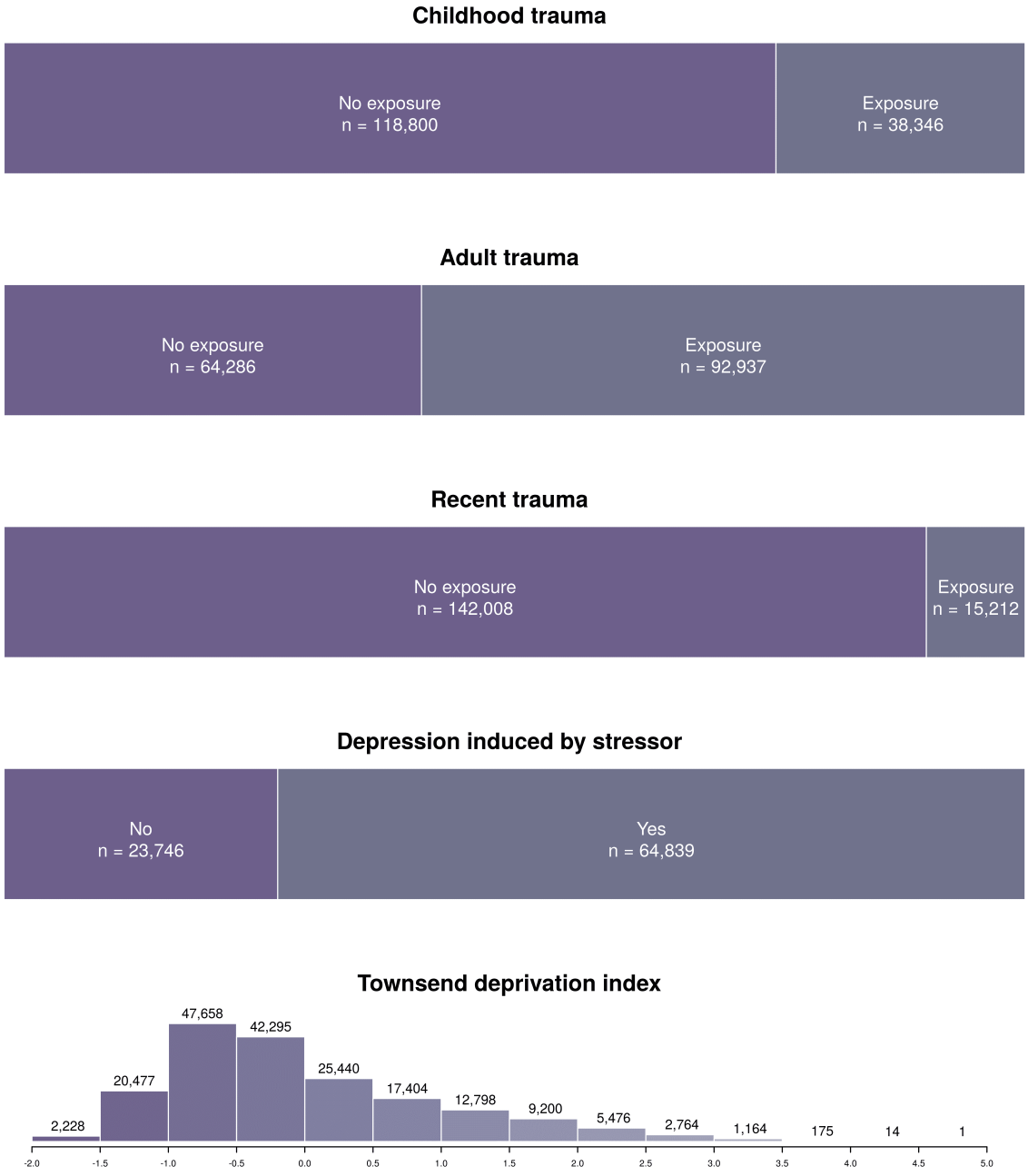
Table S3.2: Cohen's κ (inter-rater reliability) between UKBB diagnosis phenotypes

Touchscreen probable lifetime diagnosis	0.517	
Severe recurrent depression	†	0.903
	Estimated lifetime MDD diagnosis	*Touchscreen probable lifetime diagnosis*

Cohen's κ statistics for UKBB binary MDD diagnosis phenotypes based on pairwise complete observations. Note that the estimate for touchscreen probable lifetime diagnosis is likely biased upwards as the severe recurrent depression phenotype has more restrictive criteria for both cases and controls and only pairwise-complete data contributed were used in generating the estimate. †Severe recurrent depression is a refinement of estimated lifetime MDD diagnosis–agreement is necessarily perfect.

S3.2 Moderators

Figure S3.2: Distributions of environmental moderator phenotypes



Note: Graphs above reflect participants for whom any of the depression phenotypes described in S3.1 were available.

S3.2.1 Childhood trauma

Binary indicator of trauma during childhood assessed among UKBB online mental health follow-up respondents. Participants were given a positive indication if they endorsed either of the following items:

Physical abuse (20488) "When I was growing up... People in my family hit me so hard that it left me with bruises

or marks"

Sexual abuse (20490) "When I was growing up... Someone molested me (sexually)"

S3.2.2 Adult trauma

Binary indicator of trauma during adulthood assessed among UKBB online mental health follow-up respondents.

Participants were given a positive indication if they endorsed any of the following items:

Physical assault (20529) "In your life, have you...? Been attacked, mugged, robbed, or been the victim of a physically violent crime"

Sexual assault (20531) "In your life, have you...? Been a victim of a sexual assault, whether by a stranger or someone you knew"

Physical assault by partner (20523) "Since I was sixteen... A partner or ex-partner deliberately hit me or used violence in any other way"

Sexual assault by partner (20524) "Since I was sixteen... A partner or ex-partner sexually interfered with me, or forced me to have sex against my wishes"

Violence (20530) "In your life, have you...? Witnessed a sudden violent death (eg. murder, suicide, aftermath of an accident)"

Illness (20528) "In your life, have you...? Been diagnosed with a life-threatening illness"

Accident (20526) "In your life, have you...? Been in a serious accident that you believed to be life-threatening at the time"

War (20527) "In your life, have you...? Been involved in combat or exposed to a war-zone (either in the military or as a civilian)"

S3.2.3 Recent trauma

Binary indicator of trauma during the previous twelve months leading up to assessment among UKBB online mental health follow-up respondents.

Participants were given a positive indication if they answered “yes, within the last 12 months” to any of the following items (only the below were assessed for incidence in the past year):

Physical assault (20529) "In your life, have you...? Been attacked, mugged, robbed, or been the victim of a physically violent crime"

Sexual assault (20531) "In your life, have you...? Been a victim of a sexual assault, whether by a stranger or someone you knew"

Violence (20530) "In your life, have you...? Witnessed a sudden violent death (eg. murder, suicide, aftermath of an accident)"

Illness (20528) "In your life, have you...? Been diagnosed with a life-threatening illness"

Accident (20526) "In your life, have you...? Been in a serious accident that you believed to be life-threatening at the time"

War (20527) "In your life, have you...? Been involved in combat or exposed to a war-zone (either in the military or as a civilian)"

S3.2.4 Stressor-induced depression

Binary indicator of whether period of depressed mood or anhedonia was a possible consequence of a traumatic event assessed among UKBB online mental health follow-up respondents who endorsed anhedonia and/or depressed mood (`20447`).

Participants were given a positive indication if they answered “yes” to "Did this worst period start within two months of the death of someone close to you or after a stressful or traumatic event in your life?".

S3.2.5 Townsend deprivation index (TDI)

Widely-used measure of adverse socioeconomic circumstances assessed during the UKBB initial touchscreen interview (`189`) [17]. Higher values indicate greater adversity.

S4 Statistical models

S4.1 Polymorphism-wise analyses

For each polymorphism outcome Y we fit a generalized linear model (GLM) of the form

$$\textbf{Main effect model}: \qquad Y = g^{-1}\left(\alpha + \beta_G G + \sum_{C \in \text{Covariates}} [\beta_C C]\right)$$

where g is the link function implied by S4.1. Covariates for all models included age, age^2, sex, assessment center, genotyping batch, and the first ten European ancestry principle components.

Additionally, for the combinations of outcomes Y and moderators E listed in S4.1, we fit the following interaction GLMs for each polymorphism:

$$\textbf{Interaction model}: \quad Y = g^{-1}\left(\alpha + \beta_G G + \beta_E E + \beta_{G\times E} G \times E + \sum_{C \in \text{Covariates}} [\beta_C C + \beta_{C\times E} C \times E + \beta_{C\times G} C \times G]\right)$$

$$\textbf{Interaction model}_{\text{(alternate scale)}}: \quad Y = h^{-1}\left(\alpha + \beta_G G + \beta_E E + \beta_{G\times E} G \times E + \sum_{C \in \text{Covariates}} [\beta_C C + \beta_{C\times E} C \times E + \beta_{C\times G} C \times G]\right)$$

$$\textbf{Interaction model}_{\text{(improper control)}}: \quad Y = g^{-1}\left(\alpha + \beta_G G + \beta_E E + \beta_{G\times E} G \times E + \sum_{C \in \text{Covariates}} [\beta_C C]\right)$$

The *alternate scale* model assesses for interaction effects on the multiplicative scale for outcomes primarily assessed on the additive scale and vice-versa. E.g., MDD diagnosis, which is analyzed via logistic regression in our primary analyses, is reanalyzed using ordinary least squares.

The first two models control for all covariate- and variant-by-polymorphism interactions as is necessary to avoid potential confounding [18, 19], but as is rarely employed in candidate gene-by-interaction research [20]. For the latter reason, we also present the results from the *improper control* interaction models.

Table S3.3: LDSC genetic correlation estimates among depression phenotypes, height, and type 2 diabetes

Trait 1	*Trait 2*	r_g	se	p
Touchscreen probable lifetime diagnosis	Severe recurrent MDD[‡]	0.929	0.0670	1.08e-43
Estimated lifetime MDD diagnosis	Touchscreen probable lifetime diagnosis	0.939	0.0821	2.83e-30
Estimated lifetime MDD diagnosis	Severe recurrent MDD[‡]	0.940	0.0274	3.08e-258
PGC lifetime MDD diagnosis	Estimated lifetime MDD diagnosis	0.855	0.0535	2.08e-57
PGC lifetime MDD diagnosis	Touchscreen probable lifetime diagnosis	0.822	0.0490	2.81e-63
PGC lifetime MDD diagnosis	Severe recurrent MDD[‡]	0.885	0.0492	2.21e-72
Type 2 diabetes (DIAGRAM consortium)	PGC lifetime MDD diagnosis	0.036	0.0374	0.339
Type 2 diabetes (DIAGRAM consortium)	Estimated lifetime MDD diagnosis	0.077	0.0762	0.314
Type 2 diabetes (DIAGRAM consortium)	Touchscreen probable lifetime diagnosis	0.074	0.0784	0.344
Type 2 diabetes (DIAGRAM consortium)	Severe recurrent MDD[‡]	-0.052	0.0727	0.476
Height (GIANT consortium)	Type 2 diabetes (DIAGRAM consortium)	-0.010	0.0392	0.809
Height (GIANT consortium)	PGC lifetime MDD diagnosis	-0.026	0.0186	0.167
Height (GIANT consortium)	Estimated lifetime MDD diagnosis	-0.023	0.0332	0.485
Height (GIANT consortium)	Touchscreen probable lifetime diagnosis	-0.088	0.0344	0.011
Height (GIANT consortium)	Severe recurrent MDD[‡]	-0.028	0.0334	0.410
Touchscreen probable lifetime diagnosis (ordinal)[†]	Height (GIANT consortium)	-0.081	0.0326	0.013
Touchscreen probable lifetime diagnosis (ordinal)[†]	Type 2 diabetes (DIAGRAM consortium)	0.096	0.0738	0.194
Touchscreen probable lifetime diagnosis (ordinal)[†]	PGC lifetime MDD diagnosis	0.801	0.0455	1.94e-69
Touchscreen probable lifetime diagnosis (ordinal)[†]	Estimated lifetime MDD diagnosis	0.922	0.0787	1.16e-31
Touchscreen probable lifetime diagnosis (ordinal)[†]	Touchscreen probable lifetime diagnosis	1.003	0.0066	0.000
Touchscreen probable lifetime diagnosis (ordinal)[†]	Severe recurrent MDD[‡]	0.926	0.0637	5.90e-48
Lifetime episode count[†]	Touchscreen probable lifetime diagnosis (ordinal)[†]	0.867	0.0560	4.75e-54
Lifetime episode count[†]	Height (GIANT consortium)	-0.036	0.0311	0.252
Lifetime episode count[†]	Type 2 diabetes (DIAGRAM consortium)	-0.124	0.0703	0.077
Lifetime episode count[†]	PGC lifetime MDD diagnosis	0.815	0.0401	1.02e-91
Lifetime episode count[†]	Estimated lifetime MDD diagnosis	0.948	0.0350	4.74e-161
Lifetime episode count[†]	Touchscreen probable lifetime diagnosis	0.845	0.0592	3.60e-46
Lifetime episode count[†]	Severe recurrent MDD[‡]	1.017	0.0190	0.000
Current MDD severity[†]	Lifetime episode count[†]	0.681	0.0457	2.52e-50
Current MDD severity[†]	Touchscreen probable lifetime diagnosis (ordinal)[†]	0.621	0.0529	7.32e-32
Current MDD severity[†]	Height (GIANT consortium)	0.001	0.0357	0.974
Current MDD severity[†]	Type 2 diabetes (DIAGRAM consortium)	0.207	0.0677	2.24e-3
Current MDD severity[†]	PGC lifetime MDD diagnosis	0.675	0.0381	2.29e-70
Current MDD severity[†]	Estimated lifetime MDD diagnosis	0.641	0.0652	7.89e-23
Current MDD severity[†]	Touchscreen probable lifetime diagnosis	0.618	0.0572	3.53e-27
Current MDD severity[†]	Severe recurrent MDD[‡]	0.704	0.0514	1.21e-42
Conditional lifetime symptom count[†]	Current MDD severity[†]	0.689	0.0753	5.85e-20
Conditional lifetime symptom count[†]	Lifetime episode count[†]	0.646	0.0763	2.77e-17
Conditional lifetime symptom count[†]	Touchscreen probable lifetime diagnosis (ordinal)[†]	0.676	0.0830	3.85e-16
Conditional lifetime symptom count[†]	Height (GIANT consortium)	-0.010	0.0431	0.812
Conditional lifetime symptom count[†]	Type 2 diabetes (DIAGRAM consortium)	0.256	0.0861	2.93e-3
Conditional lifetime symptom count[†]	PGC lifetime MDD diagnosis	0.636	0.0594	9.36e-27
Conditional lifetime symptom count[†]	Estimated lifetime MDD diagnosis	0.695	0.0675	6.95e-25
Conditional lifetime symptom count[†]	Touchscreen probable lifetime diagnosis	0.683	0.0877	6.80e-15
Conditional lifetime symptom count[†]	Severe recurrent MDD[‡]	0.695	0.0652	1.41e-26

LD score regression genetic correlation estimates on the liability scale. Details of the estimation procedure are given in S4.4. [†]It is uncertain how to properly account for sample ascertainment for these non-binary phenotypes. These estimates should be interpreted with caution. [‡]As the severe recurrent MDD phenotype was constructed by excluding cases and controls with mild or moderate MDD symptoms, we were unable to translate these estimates to the liability scale.

Table S3.4: LDSC heritability estimates on the liability scale

Trait	$h^2_{\text{liability}}$	se
Estimated lifetime MDD diagnosis	0.057	0.007
Current MDD severity†	0.059	0.005
Conditional lifetime symptom count†	0.052	0.008
Lifetime episode count†	0.059	0.005
Touchscreen probable lifetime diagnosis	0.090	0.008
Touchscreen probable lifetime diagnosis (ordinal)†	0.065	0.006
Severe recurrent MDD‡	0.075	0.008
PGC lifetime MDD diagnosis	0.085	0.004
Type 2 diabetes (DIAGRAM consortium)	0.342	0.018
Height (GIANT consortium)	0.120	0.013

LD score regression heritability estimates on the liability scale. Details of the estimation procedure are given in S4.4. †It is uncertain how to properly account for sample ascertainment for these non-binary phenotypes. These estimates should be interpreted with caution. ‡As the severe recurrent MDD phenotype was constructed by excluding cases and controls with mild or moderate MDD symptoms, we were unable to translate these estimates to the liability scale.

S4.1.1 Design matrix lower rank approximation

Because many variables were only available for a subset of UKBB participants, including fixed effects of both genotyping batch and assessment center (which were crossed with one another) resulted in high multicollinearity such that the design matrix was no longer of full column rank. This caused difficulty in model fitting, which we avoided by using the following lower rank approximate design matrix:

In the case of the main effect models, if the complete design matrix was of the form

$$D = \left(\begin{array}{ccc|c} x_{11} & \dots & x_{1p} & g_1 \\ \vdots & & \vdots & \vdots \\ x_{n1} & \dots & x_{np} & g_n \end{array}\right) \equiv \begin{pmatrix} X & g \end{pmatrix}$$

with g the genotype vector for n participants and X the matrix of p covariates with rank $r < p$, we computed the "skinny" SVD:

$$X = \underset{n\times r}{U}\ \underset{r\times r}{\Sigma}\ \underset{r\times p}{V^*}$$

and constructed a lower rank approximation to X via

$$X' \equiv U\Sigma \in \mathbb{R}^{n\times p}$$

such that

$$\text{span col}X' = \text{span col}X$$

In model fitting, the previous design matrix was then replaced by

$$D' \equiv \begin{pmatrix} X' & g \end{pmatrix}$$

or, in the case of the interaction models, by

$$D' \equiv \begin{pmatrix} x_{11} & \dots & x_{1r} & g_1x_{11} & \dots & g_1x_{1r} & e_1x_{11} & \dots & e_1x_{1r} & g_1 & e_1 & g_1\cdot e_1 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ x_{n1} & \dots & x_{nr} & g_nx_{n1} & \dots & g_nx_{nr} & e_nx_{n1} & \dots & e_nx_{nr} & g_n & e_n & g_n\cdot e_n \end{pmatrix}$$

where e is the environmental moderator.

Figure S3.3: LDSC genetic correlation estimate heatmap

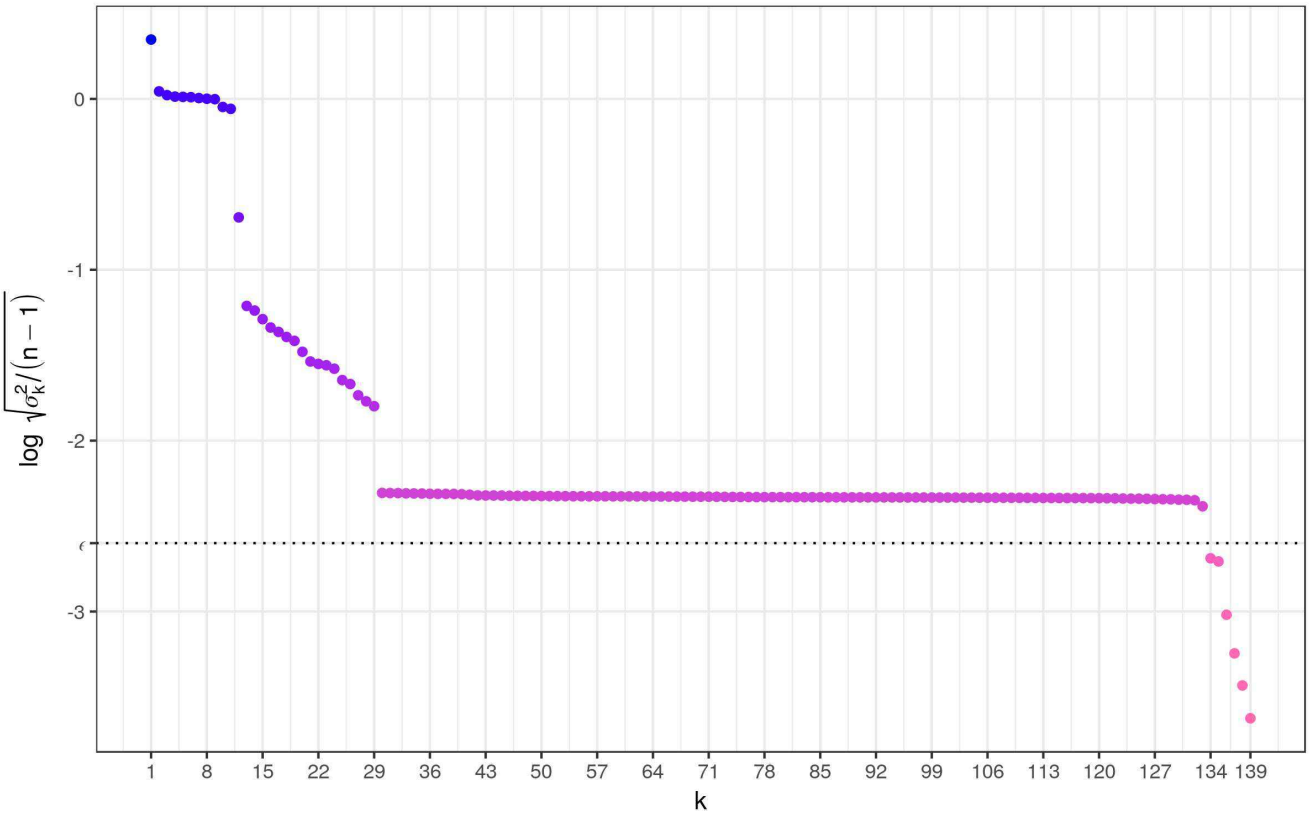
LD score regression genetic correlation estimates on the liability scale. Details of the estimation procedure are given in S4.4. Estimates for lifetime episode count, current MDD severity, conditional lifetime symptom count, and the ordinal classification of touchscreen probable lifetime diagnosis should be interpreted with caution as it's uncertain how to properly account for sample ascertainment for these non-binary phenotypes. As the severe recurrent MDD phenotype was constructed by excluding cases and controls with mild or moderate MDD symptoms, we were unable to translate these estimates to the liability scale. *Significant at $\alpha = .01$

In practice, of course, the trailing singular values $\sigma_{r+1}, \ldots, \sigma_p$ are not exactly zero but are very small, in which case, by the Eckart-Young-Mirsky theorem, X' is the optimal rank r approximation of X with respect to both the spectral and Frobenius norms [21]. In practice, we chose ϵ such that

$$\sigma_k > \epsilon \text{ only for } k = 1, \ldots, r$$

as pictured in S4.1.

Figure S4.1: Singular value threshold for design matrix



S4.2 Gene-wise and gene-set analyses

The MAGMA software [22] was used to perform gene-wise and gene-set analyses for the top eighteen candidate genes, separately in the PGC and UKBB datasets. Summary statistics from the PGC2 MDD GWAS [8] were used as input for the PGC analyses, whereas raw genotypes were available for the UKBB.

As only summary statistics were available for the PGC sample, the primary analyses used the $-\sum \log p$ method, which is MAGMA's default gene-level association statistic for summary statistics. In the UKBB, where individual-level genotype data were available, primary analyses used the principal components regression method (regressing phenotype on the principal components corresponding to nontrivial singular values of the matrix of SNP genotypes within a given gene), which is MAGMA's default gene-level association statistic for genotype data. Secondary analyses for all samples utilized the $\min p$-value method.

Paraphrasing the MAGMA manual, the $\min p$-value model is most sensitive when only a small proportion of SNPs in a gene show association, whereas both the mean SNP association and principal components regression models are more attuned to the mean SNP association. The mean SNP association model tends to skew towards associations in areas of higher linkage disequilibrium (LD) within a gene, whereas the principal components regression model has greater power to detect associations in low LD areas, but is less sensitive when only a small proportion of SNPs within a gene are associated.

The "competitive" tests (see below) are reported with one-sided p-values and "relative" tests are reported with two-sided p-values as per MAGMA guidelines. We included sex, age, age^2, genotyping batch, assessment center, and the first 10 European ancestry principal components as covariates for UKBB phenotypes. When annotating SNPs to genes, we used the NCBI Build 37 gene locations and allowed SNPs within a 25kb window of the gene start and end points to be mapped to that gene. Comparison gene sets were downloaded from the GWAS Catalog [23] and from the CTG lab's list of curated pre- and post-synapse gene sets [24, 25].

S4.2.1 Gene-wise analyses

UKBB sample (individual-level genotype data). Briefly, the default gene analysis in MAGMA for raw genotype data is based on multiple linear principal components regression, using an F-test to compute the gene p-value (although some assumptions of the F-test are violated when the outcome is a polychotomous phenotype, as in some of our analyses, comparisons of MAGMA's F-test p-values with p-values based on permutation procedures showed that the F-test remains accurate [22]). This default model first projects the SNP matrix for a gene onto its principal components (PCs), prunes away the PCs with near-zero eigenvalues, and retains the remaining PCs as predictors for the phenotype in a linear regression model, controlling for relevant covariates (e.g., gene size, density of SNPs within the gene). In secondary analyses, we used the smallest SNP p-value within each gene as the gene-level test statistic.

PGC sample (summary statistic data). Primary analyses measured gene-wise association strength by the sum of the $-\log(p)$ values for all SNPs within the gene boundary. This model tests the mean association within a gene, and is similar to models implemented in VEGAS [26] and plink v1.9 [10]. The European subset of 1000 Genomes Phase 3 [27] was used as a reference sample to account for LD between genes. Gene-level p-values are derived from this scaled χ^2 distribution and standardized via the inverse standard normal distribution function. Secondary analyses instead used the minimum SNP p-value per gene.

S4.2.2 Gene-set analyses

After calculating the strength of association for all genes across the genome, the 18 identified candidate genes were considered as a gene set in two series of analyses:

1. MAGMA's "competitive" test assesses whether genes in the candidate gene set are more associated with MDD than all other genes not in the gene set, controlling for potentially confounding gene characteristics in the model (inverse gene minor allele count (MAC), gene density, gene length, and the log of those values).
2. MAGMA's "relative" test assesses whether genes in the candidate gene set show stronger or weaker associations with MDD than control sets of genes (genes involved in type 2 diabetes, height, or synaptic processes, chosen as negative controls).

S4.3 Power analyses

S4.3.1 Logistic models

Logistic power analyses were performed using Purcell's Genetic Power Calculator [28]. Average counted allele frequency across the sixteen polymorphisms was used for the analysis presented in Figure 2, whereas the specific allele frequency of rs12552 was used in calculating its estimated minimum sample size for 80% power.

S4.3.2 Negative binomial models

To our knowledge, no closed form power function for incident rate ratios associated with polychotomous predictors (i.e., genetic polymorphisms) in the context of negative binomial GLMs is known. We therefore proceeded using a combination of simulation and numerical techniques.

1. Using the mean empirical dispersion parameter $\overline{\theta}$ across negative binomial models of current MDD severity (MHF) and the average counted allele frequency of the sixteen candidate polymorphisms $\overline{p}$, we executed 1000 Monte Carlo iterations regressing N_k observations Y on G with

$$G \sim Binom(2, \overline{p})$$
$$Y \sim NegBinom\left(\text{rate} = 1 \cdot \lambda_k^G,\ \text{dispersion} = \overline{\theta}\right)$$

   for varying incident rate ratios λ_k and sample sizes N_k to obtain empirical power estimates γ_k.

2. A logistic function of the form
$$\hat{\gamma_k} = \left(1 + \exp\left(-\kappa * (x - \delta)\right)\right)^{-1},$$
with parameters κ, δ as linear functions of the sample size and rate
$$(\kappa, \delta) = (\mathbf{1}, N, \lambda) \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \end{pmatrix},$$
was fit to the data by minimizing the loss function
$$g\left(a_0, a_1, b_0, b_1\right) = \left\|\gamma - \hat{\gamma}\right\|_2.$$

3. This resulted in a logistic approximation to the power function with a mean absolute deviation of 3.764e-03 (sd =6.629e-03) and maximal deviation of 4.596e-02, which was used to interpolate approximate power for new values of N and λ. Results are presented graphically in S4.2.

S4.3.3 Power under measurement error regimes

Main effects We employed simulation to demonstrate that even severe measurement error with respect to MDD phenotypes would not impact our ability to detect the large candidate polymorphism main effects. For simplicity, we consider the only the binary estimated lifetime MDD diagnosis phenotype, for which we observed 29,945 cases and 85,513 controls. As a lower bound on a plausibly detectable candidate polymorphism effect, we considered the minimum detectable relative risk at 50% power in a perfectly balanced study of 500 cases and 500 controls, assuming a disease prevalence of 14.6% [29] and a risk allele frequency of 0.5 (RR = 1.161; OR = 1.189; note that choosing a lower risk allele frequency would increase the corresponding RR/OR). We then simulated case/control phenotypes Y and genotypes G for 29,945 cases and 85,513 controls and subsequently corrupted phenotype observations under three severe systematic measurement error regimes:

1. **50/50 misclassification**: for each observation, there was 50% chance that we observed the true diagnosis and a 50% chance that we based the diagnosis on the outcome of a fair coin toss;
2. **Overdiagnosis**: cases were correctly identified, but controls had a 50% chance of being misclassified as cases;
3. **Underdiagnosis**: controls were correctly identified, but cases had a 50% chance of being misclassified as controls.

Monte Carlo simulation results indicated power $\approx$ 100% at α_{gwas} = 5e-08 for detecting the effect (which is small relative to effects reported in the candidate polymorphism literature and large relative to those reported in the GWAS literature) under all three regimes (Figure S4.3).

Figure S4.3: Power simulations for main effect detection under measurement error regimes

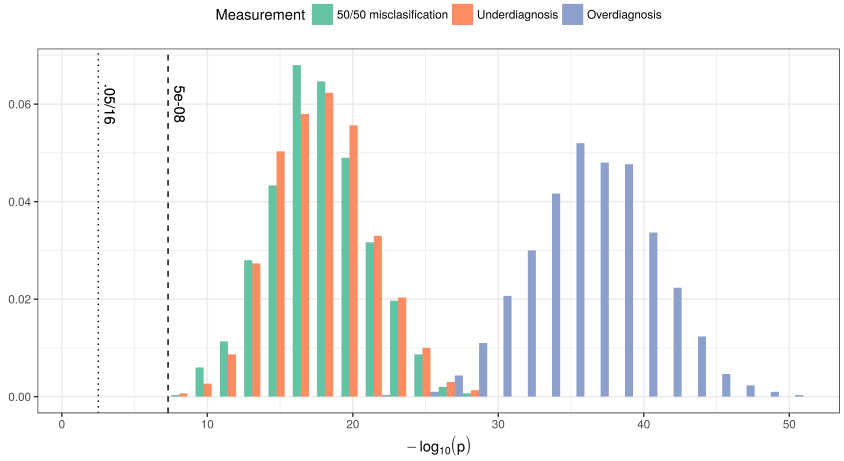
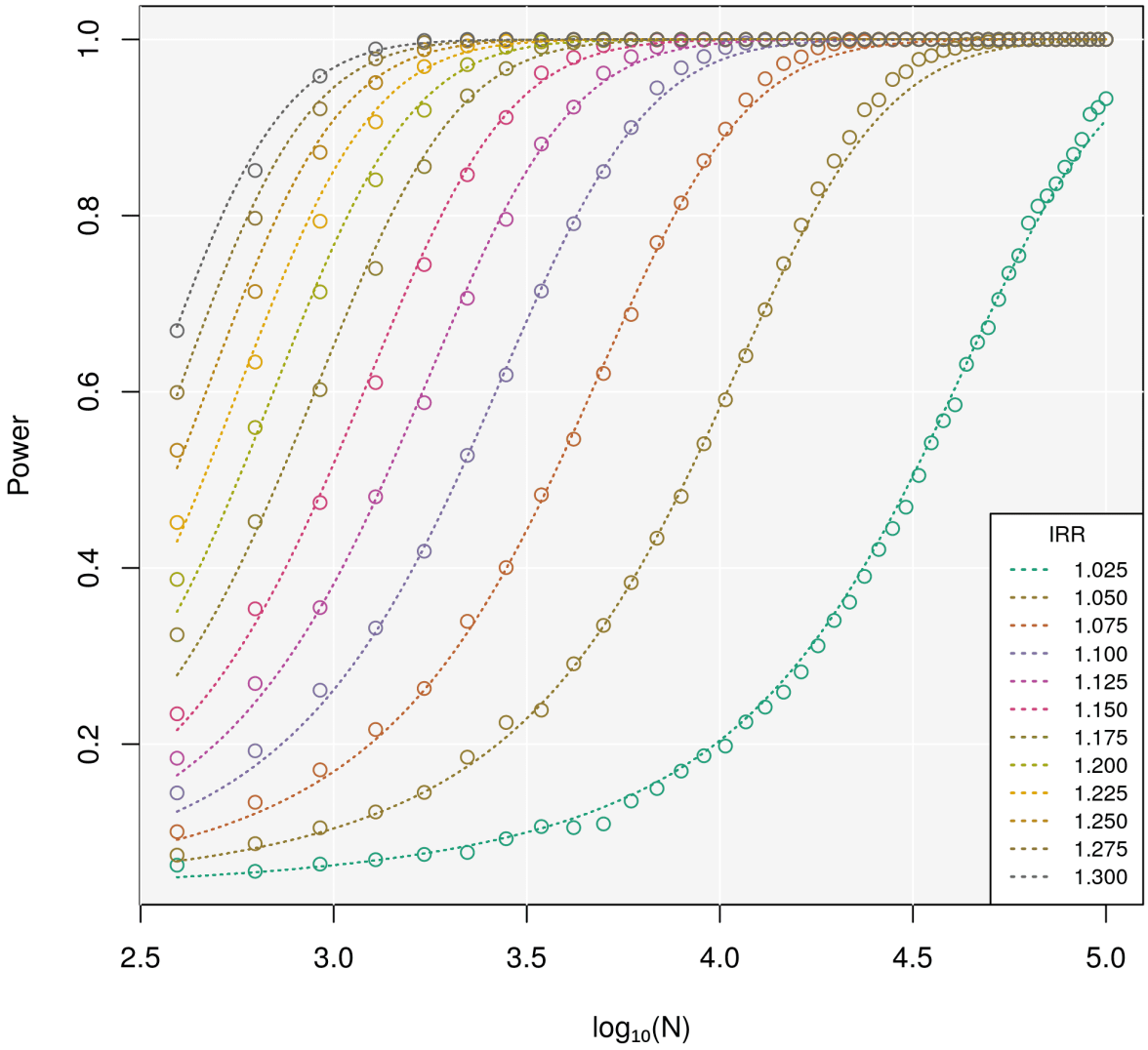


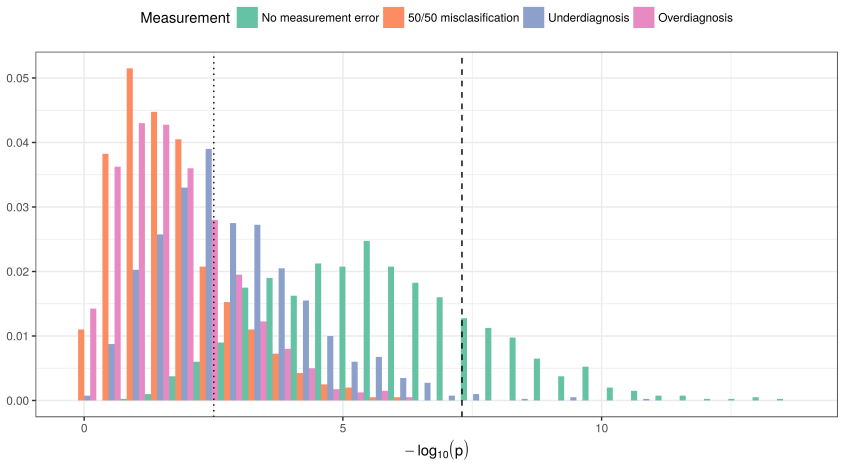
Figure S4.2: Logistic approximation of negative binomial regression power function



Dashed lines represent fitted predictions of logistic approximation whereas points represent empirical power estimates from Monte Carlo simulations.

This is not to say measurement error is a non-issue in genetic association studies. Indeed, the extreme scenarios considers above would have a catastrophic impact on our ability to detect an effect on the order of the strongest association observed in the PGC (rs12552, uncorrected OR = 1.044, p=6.093e-15), which we demonstrate via an analogous set of simulations in Figure S4.4.

Figure S4.4: Power simulations for detection of association at rs12552 under measurement error regimes



Left dotted line: $\alpha_{\text{poly}} = .05/16$; right dashed line: $\alpha_{\text{gwas}} = 5\text{e-}08$.

Interactions effects We also used simulation to demonstrate that severe measurement error with respect to both MDD phenotypes and environmental exposures would have a limited impact on our ability to detect the large candidate polymorphism × environment effects. We constructed a genotype G with a risk allele frequency of .5 and a binary exposure phenotype T with an exposure rate of .222, which matched that of our childhood traumatic event measurement in the UKBB. Additionally, we again constructed a binary diagnosis phenotype Y with 29,945 cases and 85,513 controls, via the following logistic model:

$$Pr(\text{Diagnosis}) = \text{logistic}\left\{ \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ \vec{1} & G & T & (G \circ T) \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \text{logit}(0.259) \\ \log(0.9738) \\ \log(1.677) \\ \log(1.1442) \end{pmatrix} \right\}$$

where $\circ$ is the element-wise product. We assume that G and T have been mean centered.

The above parameters, which were derived numerically and are readily verified via simulation, are such that there is near zero main effect of genotype (detectable with $\approx 0.4\%$ power at $\alpha_{\text{poly}} = .05/16$ in the context of a model excluding the interaction term in our sample of 115,458) and that the interaction would again be detectable with $\approx 50\%$ power in a sample of 500 cases and 500 controls.

We then again used a Monte Carlo procedure to determine power to detect the interaction in a sample analogous in size and prevalence of trauma exposure and MDD diagnosis prevalence to the subsample of the UKBB for which those measures were observed, corrupted under all pairwise combinations of the following severe measurement error regimes:

MDD diagnosis error regimes:

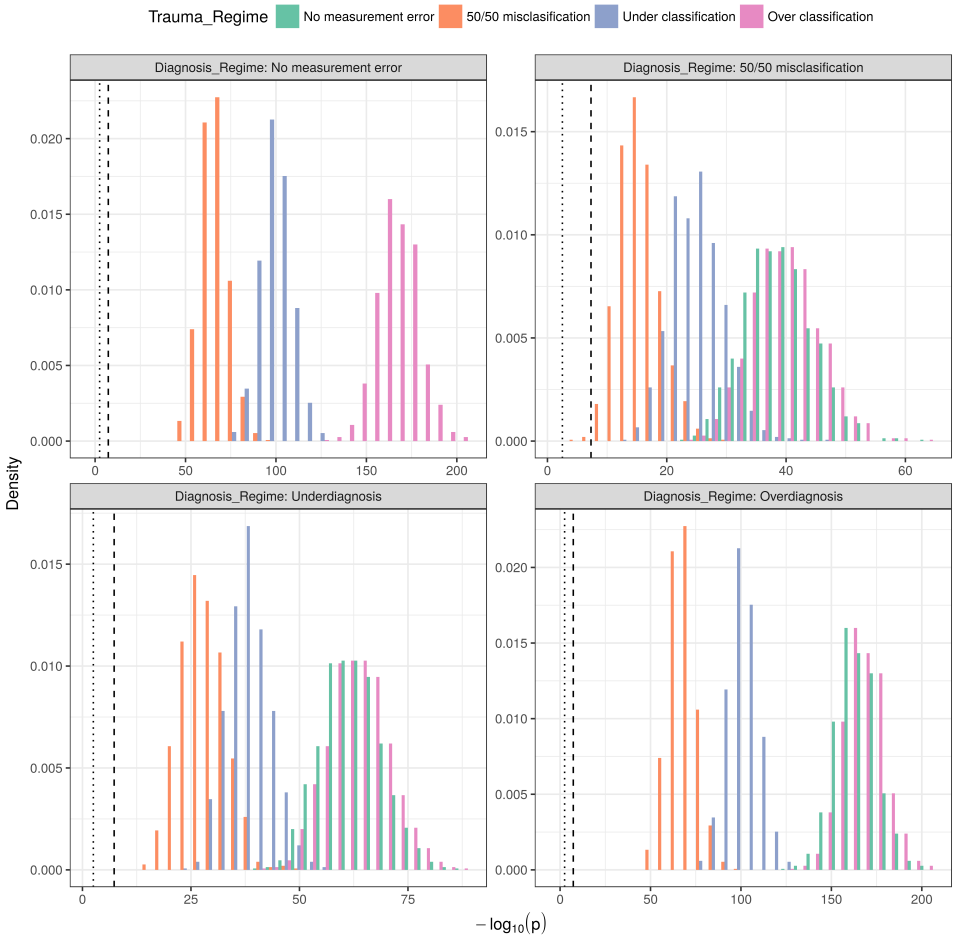
1. **No measurement error**: the correct case/control status is observed for all individuals;
2. **50/50 misclassification**: for each observation, there was 50% chance that we observed the true diagnosis and a 50% chance that we based the diagnosis on the outcome of a fair coin toss;
3. **Overdiagnosis**: cases were correctly identified, but controls had a 50% chance of being misclassified as cases;
4. **Underdiagnosis**: controls were correctly identified, but cases had a 50% chance of being misclassified as controls.

Trauma exposure error regimes:

1. **No measurement error**: the correct exposure status is observed for all individuals;
2. **50/50 misclassification**: for each observation, there was 50% chance that we observed their true exposure status and a 50% chance that we based exp on the outcome of a fair coin toss;
3. **Over classification**: exposed individuals were correctly identified, but non-exposed individuals had a 50% chance of being misclassified as exposed;
4. **Under classification**: non-exposed individuals were correctly identified, but exposed individuals had a 50% chance of being misclassified as non-exposed.

The choice of these parameters reflects a small interaction effect by candidate gene study standards (one that would only be detected half the time in a balanced case/control sample of 1000) obscured by horrendous measurement error. Nevertheless, Monte Carlo simulation results indicated power $\approx 100\%$ at $\alpha_{\text{poly}} = .05/16$ for detecting the interaction effect under every combination of measurement error regimes in a sample analogous to our own (Figure S4.5).

Figure S4.5: Power simulations for interaction effect detection under measurement error regimes



Left dotted line: $\alpha_{\text{poly}} = .05/16$; right dashed line: $\alpha_{\text{gwas}} = 5\text{e-}08$.

S4.4 Heritability and genetic correlation estimation

LD score regression (LDSC v1.0.0; [30, 31]) was used to estimate heritability and genetic correlation among depression phenotypes, as well as height and type 2 diabetes (genes associated with the latter phenotypes were used as negative controls in the relative gene-set analyses; see S4.2.2 for further details).

Whereas genome-wide summary statistics were available for PGC lifetime MDD diagnosis, linear mixed-model association tests for the ˜1.4 million SNPs utilized by LDSC were performed for each of the UKBB depression phenotypes using BOLT-LMM v2.3.2 [32], controlling for fixed effects of age, age^2, sex, assessment center, genotyping batch, and the first ten European ancestry principle components. Publicly available summary statistics for height and type 2 diabetes were downloaded from the GIANT [33] and DIAGRAM [34] consortia, respectively. Heritability estimates for binary phenotypes were translated to the liability scale via LDSC, assuming lifetime prevalences of 14.6% [29] and 5.7% [35] for MDD and type-2 diabetes, respectively. Results are presented in Tables S3.3-S3.4 and Figure S3.3.

S4.5 Replication of top PGC hits

In order to better contextualize the lack of replication of the of 16 candidate genetic polymorphisms, we sought to replicate the top 16 independent loci (among genome-wide significant loci) identified by the independent PGC meta-analysis (described in S2.1) in the UKBB, with respect to estimated lifetime MDD diagnosis. Estimated lifetime MDD diagnosis (described in S3.1.1) was chosen as we believe it most closely resembles the clinical diagnosis phenotype examined in the PGC. The choice of sixteen polymorphisms was made in the interest of parallelism with our investigation of candidate gene polymorphisms. Results are presented in S12.

S4.5.1 Identification of independent loci

Mirroring the approach taken by [8], independent genome-wide significant loci were identified via clumping within sliding 3 megabase/$R^2 < .1$ windows via the plink v1.9 command `--clump-r2 .1 --clump-kb 3000 --clump-p1 1e-04 --clump-p2 1e-04`. The 1000 Genomes Phase 3 (1KG; [27]) and UK10K ([36]) panels were combined to map linkage disequilibrium. The top 16 significant loci were selected for replication in the UKBB (S12).

S4.5.2 Association analyses

The main effects of the top sixteen genome-wide significant loci identified in the independent PGC meta-analysis on estimated lifetime MDD diagnosis in the UKBB were tested using the main effect model described in S4.1. Results are presented in S12.

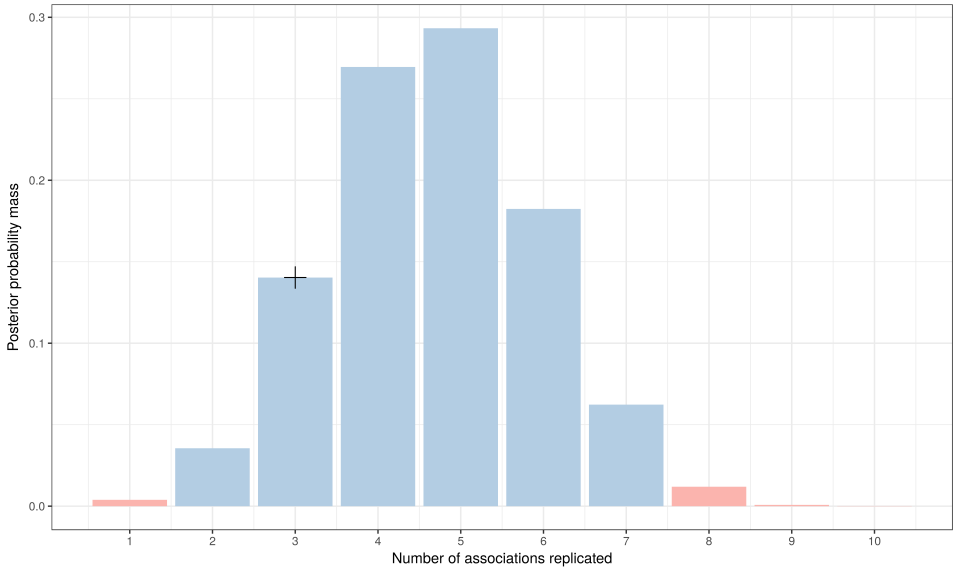
S4.5.3 Replication power analysis and correction for the winner's curse

As noted in the primary manuscript, the present study had $> 99.99\%$ power at $\alpha_{\text{gwas}} = 5\text{e-}08$ to detect an effect that could only be detected with 20% power at $\alpha = .05$ in a sample of 1,000 individuals, a sample size larger than those examined in most candidate gene studies. In contrast, the top sixteen genome-wide significant hits in the PGC were small to begin with (the largest hit, rs12552, had an estimated odds ratio of 1.044). Thus, we conducted power analyses with respect to our ability to detect the effects of the top PGC loci in the UKBB given their estimated effects. Further, because GWAS hits are only considered "hits" given significance at $\alpha_{\text{gwas}} = 5\text{e-}08$, estimated effects are subject to the *winner's curse* [37]; that is, they are biased towards extreme values. As a result, power estimates based significant GWAS effect size estimates will be commensurately biased upward.

To account for this, we implemented the weighted corrected estimator $\hat{\beta}_{\text{MSE}}$ proposed in [38], where they demonstrated unbiasedness via simulation (details provided below) For highly significant associations relative to the significance criterion (e.g., rs12552, which was significant at $p = 6.093\text{e-}15$), the winner's curse correction has little impact on the effect size distribution. In contrast, for associations only barely significant relative to the significance criterion (e.g., rs10514304, which was significant at $p = 3.627\text{e-}08$), failure to adjust for the winner's curse correction dramatically biased the estimated effect upwards (S12). We then estimated power to detect an association with estimated lifetime MDD diagnosis using allele frequencies from the UK10K reference panel and assuming a prevalence of 14.6% as reported in [29].

Finally, we calculated the expected number of replicated effects in UKBB, where "replication" was defined as attaining significance at $\alpha_{\text{poly}} = .05/16$, again for comparability with our attempts to detect associations for sixteen candidate polymorphisms in the UKBB. To do this, we modeled replication for each j^{th} locus as an independent Bernoulli trial with parameters $p_j = \widehat{\text{power}}_j$ and used a Monte Carlo procedure to obtain the distribution of the number of replications given the corrected power estimates $\widehat{\text{power}}_j$, $j = 1, \ldots, 16$ (Figure S4.6).

Figure S4.6: Distribution of number of replicated associations



Distribution of the number of the top 16 PGC associations we'd expect to replicate at $\alpha_{\text{poly}} = .05/16$ in the UKBB data given winner's curse corrected power estimates. Light blue bars indicate the 95% credible interval (exact interval: 2 - 7; smoothed interval using estimator of [39]: 2.064 - 7.321. The cross at 3 indicates the actual number of replicated associations.

Details of the estimator $\hat{\beta}_{MSE}$

We describe the corrected estimator $\hat{\beta}_{\text{MSE}}$ proposed in [38].

Let $\hat{\beta}, \hat{s}$ denote the uncorrected estimates of the log odds ratio and it's standard error, and let $\mathcal{S} = \left\{(\hat{\beta}, \hat{s}) : |\hat{\beta}/\hat{s}| > c\right\}$ denote the set of such estimates satisfying genome-wide significance, where $c = \Phi^{-1}\left(1 - \frac{\alpha_{gwas}}{2}\right)$. Then the conditional density of the uncorrected estimator $\hat{\beta}$ given the true effect β and selection is

$$f_{\hat{\beta}|(\hat{\beta},\hat{s})\in\mathcal{S}}(x;\beta) = \frac{\frac{1}{\hat{s}}\phi\left(\frac{x-\beta}{\hat{s}}\right)}{\Phi\left(\frac{x-\beta}{\hat{s}} - c\right) + \Phi\left(-\frac{\beta}{\hat{s}} - c\right)} \left[\!\left[(\hat{\beta}, \hat{s}) \in \mathcal{S}\right]\!\right].$$

The corrected quantile estimator $\hat{\beta}_{\text{Med}}$ is the solution to

$$\hat{\beta}_{\text{Med}} = \gamma \text{ s.t. } \int_{-\infty}^{\hat{\beta}} f_{\hat{\beta}|(\hat{\beta},\hat{s})\in\mathcal{S};\gamma}(x;\gamma)dx = \frac{1}{2},$$

We used the bias-reduced estimator $\hat{\beta}_{\text{MSE}}(p)$, evaluated at $p = .5$, which is given by the linear combination

$$\begin{aligned}
\hat{\beta}_{\text{MSE}}(p) &= \hat{K}(p)\hat{\beta}_{\text{uncr}}(p) + (1 - \hat{K}(p))\hat{\beta}_{\text{Q}}(p), \\
\hat{K}(p) &= \frac{\hat{s}^2}{\hat{s}^2 + \left(\hat{\beta}_{\text{uncr}}(p) - \hat{\beta}_{\text{Q}}(p)\right)^2}, \\
\hat{\beta}_{\text{uncr}}(p) &= F^{-1}_{\mathcal{N}(\hat{\beta},\hat{s})}(p).
\end{aligned}$$

$\hat{\beta}_{\text{MSE}}$ does not admit a closed form representation and thus was computed via standard root finding methods.

S5 Amendments to preregistration

Our analysis plan was preregistered through the Open Science Framework after identification of top candidate genes/polymorphisms and prior to running any analyses. The preregistration is available at https://osf.io/xrkm6/. Below we discuss changes to the preregistration reflecting 1. corrections to and 2. departures from the preregistered protocol.

S5.1 Corrections

1. Genotyping batch and assessment center are missing from the list of covariates.
2. The PubMed searches listed in the section "Documented in `geneID/geneIdentification.md`" are inaccurate. The correct searches are detailed in section S1 on page 4.
3. The estimated lower bounds of number of studies per gene are incorrect as they were based on data that included the first few months of 2017. Corrected results based on the 25 year period from 1991 to 2016 are given in Table 1 of the primary manuscript.
4. The name of the top polymorphism in *ABCB1* contains a typo.
5. In the polymorphism-based analysis plan, the GLM families for the variable `ukb_tdi` was mistakenly labeled as linear and ukb_sx_lif_cond as logit when the reverse was intended.
6. Testing center and genotyping batch were mistakenly omitted from the list of covariates to be used in the UKBB analysis protocols.

S5.2 Departures

1. The age×sex and age^2×sex covariates were omitted as high multicollinearity with the other covariates caused convergence difficulties in fitting many of the GLMs. Further, additional multicollinearity problems were addressed as described in S4.1. Both procedures were implemented for purely computational reasons did not substantively change any results.
2. Analyses using the severe recurrent depression phenotype described in S3.1.7 on page 15 were *post hoc* and intended as a sensitivity analysis to ensure results didn't change substantially when using a stricter ascertainment protocol.
3. The analyses examining the main effects of environmental exposure variables presented in S6 were not present in the preregistered analysis protocol.
4. The heritability/genetic correlation analyses detailed in S4.4 were not present in the preregistered analysis protocol.
5. The attempted replication of the top 16 loci described in S4.5 were suggested by a reviewer.
6. The simulations examining the potential impact of measurement error described in S4.3.3 were motivated by the concerns of a reviewer.

Table S4.1: Interaction models

Moderator	**Outcomes**	**Presentation**	**Family**
Childhood trauma	Estimated lifetime MDD Diagnosis	*primary*	Bernoulli
	Conditional lifetime symptom count	*secondary*	Gaussian
	Lifetime episode count	*secondary*	Binomial
	Severe recurrent depression[†]	*secondary*	Binomial
Adult trauma	Estimated lifetime MDD Diagnosis	*primary*	Gaussian
	Conditional lifetime symptom count	*secondary*	Bernoulli
	Lifetime episode count	*secondary*	Binomial
	Severe recurrent depression[†]	*secondary*	Binomial
Townsend deprivation index (TDI)	Estimated lifetime MDD Diagnosis	*secondary*	Bernoulli
	Conditional lifetime symptom count	*secondary*	Gaussian
	Lifetime episode count	*secondary*	Binomial
	Severe recurrent depression[†]	*secondary*	Binomial
	Touchscreen probable lifetime diagnosis, binary classification	*secondary*	Bernoulli
	Touchscreen probable lifetime diagnosis, ordinal classification	*secondary*	Binomial
Stressor-induced depression	Conditional lifetime symptom count	*secondary*	Gaussian
Recent trauma	Current MDD severity	*primary*	Negative binomial

Interaction models included in preregistered analysis plan. Primary analyses are presented graphically in the main body of manuscript. All plausible interactions with adequate sample sizes were tested (e.g., moderation of estimated lifetime diagnosis associations by past year trauma was not of interest).

[†]*Severe recurrent depression was added after preregistration as a post-hoc analysis.*

Part II
Supplemental results

S6 Main effects of moderator variables

Table S6.1: Main effects of environmental moderators

Outcome	*Predictor*	*Estimate*	*se*	*z*	*p*	$-\log_{10} p$	*n*
Estimated lifetime MDD diagnosis	Childhood trauma	1.655†	0.016	32.048	2.33e-225	224.633	115,405
Estimated lifetime MDD diagnosis	Adult trauma	1.670†	0.014	35.968	2.61e-283	282.583	115,450
Estimated lifetime MDD diagnosis	Townsend deprivation index	1.140†	0.008	16.318	7.31e-60	59.136	115,339
Conditional lifetime symptom count	Childhood trauma	0.408‡	0.016	26.190	2.28e-150	149.643	62,210
Conditional lifetime symptom count	Adult trauma	0.381‡	0.014	27.338	1.40e-163	162.854	62,237
Conditional lifetime symptom count	Townsend deprivation index	0.086‡	0.008	10.798	3.73e-27	26.429	62,171
Lifetime episode count	Childhood trauma	1.566†	0.014	32.930	8.12e-238	237.090	112,412
Lifetime episode count	Adult trauma	1.523†	0.012	39.707	0.00e+00	∞	112,451
Lifetime episode count	Townsend deprivation index	1.142†	0.007	19.386	1.01e-83	82.994	112,340
Touchscreen probable lifetime diagnosis	Childhood trauma	1.759†	0.031	18.493	2.33e-76	75.632	28,716
Touchscreen probable lifetime diagnosis	Adult trauma	1.796†	0.028	20.802	4.13e-96	95.385	28,727
Touchscreen probable lifetime diagnosis	Townsend deprivation index	1.180†	0.009	18.871	1.97e-79	78.706	80,083
Touchscreen probable lifetime diagnosis, ordinal	Childhood trauma	1.545†	0.031	14.145	2.00e-45	44.699	28,716
Touchscreen probable lifetime diagnosis, ordinal	Adult trauma	1.476†	0.027	14.583	3.60e-48	47.444	28,727
Touchscreen probable lifetime diagnosis, ordinal	Townsend deprivation index	1.161†	0.009	17.293	5.35e-67	66.272	80,083
Current MDD severity	Recent trauma	1.431*	0.013	27.004	1.32e-160	159.880	115,447

Main effects of environmental moderators on primary and secondary MDD outcomes in the UKBB. Only TDI is standardized.
†Odds ratios; ‡Linear regression slopes; *Incident rate ratios

S7 Polymorphism level main effects

Table S7.1: Estimated lifetime MDD diagnosis: main effect of variant

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.000	0.010	0.972 - 1.029	0.007	0.994	115,257
2	*BDNF*	*rs6265*	A	0.998	0.012	0.962 - 1.036	0.132	0.895	115,257
3	*COMT*	*rs4680*	G	0.988	0.010	0.960 - 1.017	1.249	0.212	115,257
4	*HTR2A*	*rs6311*	A	0.980	0.010	0.952 - 1.009	2.007	0.045	115,257
5	*TPH1*	*rs1800532*	A	1.000	0.010	0.971 - 1.029	0.041	0.968	115,257
6	*DRD4*	*VNTR*	7+ repeats	1.012	0.012	0.978 - 1.048	1.050	0.294	115,257
7	*DRD2*	*rs1800497*	T	0.972	0.012	0.938 - 1.007	2.352	0.019	115,257
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.998	0.009	0.973 - 1.025	0.200	0.841	115,257
9	*APOE*	*rs429358/rs7412*	ε-4	0.996	0.014	0.957 - 1.037	0.302	0.762	115,257
10	*MTHFR*	*rs1801133*	T	1.005	0.010	0.975 - 1.036	0.456	0.648	115,257
11	*CLOCK*	*rs1801260*	C	1.010	0.011	0.978 - 1.044	0.941	0.347	115,257
12	*SLC6A3*	*VNTR*	10+ repeats	1.008	0.011	0.975 - 1.042	0.721	0.471	115,257
13	*ACE*	*in/del*	deletion	1.002	0.010	0.974 - 1.031	0.229	0.819	115,257
14	*ABCB1*	*rs1045642*	C	1.005	0.010	0.976 - 1.034	0.482	0.630	115,257
15	*DRD3*	*rs6280*	C	0.988	0.010	0.958 - 1.018	1.202	0.229	115,257
16	*DBH*	*rs1611115*	T	0.986	0.012	0.951 - 1.022	1.186	0.236	113,474

Note: Logistic regression weights for estimated lifetime MDD diagnosis on variant, controling for sex, age, age^2, 10 European ancestry principle components, testing center, and batch.

Table S7.2: Current MDD severity: main effect of variant

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.008	0.005	0.992 - 1.025	1.483	0.138	115,257
2	*BDNF*	*rs6265*	A	1.004	0.007	0.983 - 1.025	0.583	0.560	115,257
3	*COMT*	*rs4680*	G	0.983	0.005	0.967 - 0.999	3.089	0.002	115,257
4	*HTR2A*	*rs6311*	A	0.998	0.006	0.981 - 1.014	0.438	0.662	115,257
5	*TPH1*	*rs1800532*	A	1.012	0.006	0.995 - 1.029	2.092	0.036	115,257
6	*DRD4*	*VNTR*	7+ repeats	1.004	0.007	0.985 - 1.024	0.626	0.531	115,257
7	*DRD2*	*rs1800497*	T	0.995	0.007	0.975 - 1.015	0.792	0.428	115,257
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.998	0.005	0.984 - 1.012	0.454	0.650	115,257
9	*APOE*	*rs429358/rs7412*	ε-4	0.990	0.008	0.968 - 1.013	1.313	0.189	115,257
10	*MTHFR*	*rs1801133*	T	1.012	0.006	0.995 - 1.030	2.116	0.034	115,257
11	*CLOCK*	*rs1801260*	C	0.999	0.006	0.981 - 1.017	0.217	0.828	115,257
12	*SLC6A3*	*VNTR*	10+ repeats	1.003	0.006	0.985 - 1.022	0.504	0.614	115,257
13	*ACE*	*in/del*	deletion	1.000	0.006	0.984 - 1.016	0.052	0.958	115,257
14	*ABCB1*	*rs1045642*	C	1.000	0.006	0.984 - 1.016	0.036	0.971	115,257
15	*DRD3*	*rs6280*	C	0.990	0.006	0.973 - 1.007	1.764	0.078	115,257
16	*DBH*	*rs1611115*	T	1.005	0.007	0.985 - 1.025	0.677	0.498	113,474

Note: Negative binomial regression weights for current MDD severity on variant, controling for sex, age, age^2, 10 European ancestry principle components, testing center, and batch.

Table S7.3: Conditional lifetime symptom count: main effect of variant

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	-0.004	0.010	-0.033 - 0.024	0.458	0.647	62,138
2	*BDNF*	*rs6265*	A	0.008	0.012	-0.028 - 0.045	0.665	0.506	62,138
3	*COMT*	*rs4680*	G	-0.002	0.010	-0.030 - 0.027	0.191	0.848	62,138
4	*HTR2A*	*rs6311*	A	-0.003	0.010	-0.032 - 0.026	0.257	0.797	62,138
5	*TPH1*	*rs1800532*	A	0.007	0.010	-0.022 - 0.037	0.740	0.459	62,138
6	*DRD4*	*VNTR*	7+ repeats	-0.002	0.012	-0.036 - 0.032	0.176	0.860	62,138
7	*DRD2*	*rs1800497*	T	-0.011	0.012	-0.047 - 0.024	0.934	0.350	62,138
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.010	0.009	-0.016 - 0.035	1.093	0.274	62,138
9	*APOE*	*rs429358/rs7412*	ε-4	-0.019	0.014	-0.059 - 0.021	1.430	0.153	62,138
10	*MTHFR*	*rs1801133*	T	-0.001	0.010	-0.031 - 0.030	0.059	0.953	62,138
11	*CLOCK*	*rs1801260*	C	0.002	0.011	-0.030 - 0.034	0.206	0.837	62,138
12	*SLC6A3*	*VNTR*	10+ repeats	-0.007	0.011	-0.039 - 0.026	0.599	0.549	62,138
13	*ACE*	*in/del*	deletion	0.004	0.010	-0.025 - 0.032	0.394	0.694	62,138
14	*ABCB1*	*rs1045642*	C	0.003	0.010	-0.026 - 0.031	0.283	0.778	62,138
15	*DRD3*	*rs6280*	C	-0.005	0.010	-0.035 - 0.025	0.507	0.612	62,138
16	*DBH*	*rs1611115*	T	-0.008	0.012	-0.044 - 0.027	0.672	0.502	61,208

Note: Linear regression weights for conditional lifetime symptom count on variant, controling for sex, age, age^2, 10 European ancestry principle components, testing center, and batch.

Table S7.4: Lifetime episode count: main effect of variant

	Gene	Polymorphism	Risk/counted allele	$\exp(\beta)$	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.998	0.008	0.975 - 1.022	0.218	0.828	112,261
2	*BDNF*	*rs6265*	A	0.999	0.010	0.969 - 1.030	0.064	0.949	112,261
3	*COMT*	*rs4680*	G	0.994	0.008	0.971 - 1.018	0.729	0.466	112,261
4	*HTR2A*	*rs6311*	A	0.999	0.008	0.975 - 1.023	0.139	0.889	112,261
5	*TPH1*	*rs1800532*	A	1.001	0.008	0.977 - 1.026	0.108	0.914	112,261
6	*DRD4*	*VNTR*	7+ repeats	1.015	0.010	0.986 - 1.044	1.488	0.137	112,261
7	*DRD2*	*rs1800497*	T	0.983	0.010	0.954 - 1.013	1.694	0.090	112,261
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.006	0.007	0.985 - 1.027	0.813	0.416	112,261
9	*APOE*	*rs429358/rs7412*	ε-4	1.019	0.011	0.986 - 1.054	1.692	0.091	112,261
10	*MTHFR*	*rs1801133*	T	1.008	0.009	0.983 - 1.034	0.912	0.362	112,261
11	*CLOCK*	*rs1801260*	C	0.996	0.009	0.970 - 1.023	0.428	0.669	112,261
12	*SLC6A3*	*VNTR*	10+ repeats	1.009	0.009	0.981 - 1.036	0.926	0.354	112,261
13	*ACE*	*in/del*	deletion	1.005	0.008	0.981 - 1.029	0.618	0.537	112,261
14	*ABCB1*	*rs1045642*	C	1.000	0.008	0.977 - 1.024	0.025	0.980	112,261
15	*DRD3*	*rs6280*	C	1.000	0.009	0.975 - 1.026	0.010	0.992	112,261
16	*DBH*	*rs1611115*	T	0.994	0.010	0.965 - 1.024	0.553	0.580	110,513

Note: Ordinal logistic regression weights for lifetime episode count on variant, controling for sex, age, age^2, 10 European ancestry principle components, testing center, and batch.

Table S7.5: Touchscreen probable lifetime diagnosis: main effect of variant

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.008	0.011	0.976 - 1.040	0.751	0.453	90,944
2	*BDNF*	*rs6265*	A	1.010	0.014	0.970 - 1.051	0.743	0.457	90,944
3	*COMT*	*rs4680*	G	1.003	0.011	0.971 - 1.034	0.238	0.812	90,944
4	*HTR2A*	*rs6311*	A	1.004	0.011	0.972 - 1.037	0.396	0.692	90,944
5	*TPH1*	*rs1800532*	A	1.000	0.011	0.968 - 1.033	0.025	0.980	90,944
6	*DRD4*	*VNTR*	7+ repeats	1.023	0.013	0.985 - 1.061	1.751	0.080	90,944
7	*DRD2*	*rs1800497*	T	1.023	0.013	0.983 - 1.062	1.663	0.096	90,944
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.001	0.010	0.973 - 1.030	0.130	0.896	90,944
9	*APOE*	*rs429358/rs7412*	ε-4	1.009	0.015	0.965 - 1.053	0.633	0.527	90,944
10	*MTHFR*	*rs1801133*	T	1.008	0.011	0.974 - 1.042	0.699	0.485	90,944
11	*CLOCK*	*rs1801260*	C	1.031	0.012	0.995 - 1.067	2.482	0.013	90,944
12	*SLC6A3*	*VNTR*	10+ repeats	1.020	0.012	0.983 - 1.056	1.578	0.114	90,944
13	*ACE*	*in/del*	deletion	1.016	0.011	0.984 - 1.048	1.463	0.143	90,944
14	*ABCB1*	*rs1045642*	C	1.005	0.011	0.973 - 1.037	0.431	0.667	90,944
15	*DRD3*	*rs6280*	C	1.013	0.011	0.980 - 1.047	1.174	0.241	90,944
16	*DBH*	*rs1611115*	T	1.002	0.013	0.963 - 1.042	0.157	0.876	89,524

Note: Logistic regression weights for touchscreen probable lifetime diagnosis on variant, controling for sex, age, age^2, 10 European ancestry principle components, testing center, and batch.

Table S7.6: Touchscreen probable lifetime diagnosis, ordinal classification: main effect of variant

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.009	0.010	0.978 - 1.040	0.817	0.414	48,190
2	*BDNF*	*rs6265*	A	1.011	0.013	0.972 - 1.051	0.848	0.396	48,190
3	*COMT*	*rs4680*	G	1.003	0.011	0.972 - 1.034	0.280	0.779	48,190
4	*HTR2A*	*rs6311*	A	1.006	0.011	0.974 - 1.037	0.522	0.602	48,190
5	*TPH1*	*rs1800532*	A	1.001	0.011	0.969 - 1.033	0.093	0.926	48,190
6	*DRD4*	*VNTR*	7+ repeats	1.023	0.013	0.985 - 1.060	1.756	0.079	48,190
7	*DRD2*	*rs1800497*	T	1.023	0.013	0.984 - 1.062	1.694	0.090	48,190
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.000	0.009	0.972 - 1.028	0.016	0.987	48,190
9	*APOE*	*rs429358/rs7412*	ε-4	1.009	0.015	0.966 - 1.052	0.605	0.545	48,190
10	*MTHFR*	*rs1801133*	T	1.013	0.011	0.980 - 1.046	1.138	0.255	48,190
11	*CLOCK*	*rs1801260*	C	1.029	0.012	0.994 - 1.065	2.426	0.015	48,190
12	*SLC6A3*	*VNTR*	10+ repeats	1.017	0.012	0.982 - 1.053	1.435	0.151	48,190
13	*ACE*	*in/del*	deletion	1.011	0.011	0.980 - 1.042	1.022	0.307	48,190
14	*ABCB1*	*rs1045642*	C	1.003	0.011	0.971 - 1.034	0.256	0.798	48,190
15	*DRD3*	*rs6280*	C	1.012	0.011	0.979 - 1.045	1.078	0.281	48,190
16	*DBH*	*rs1611115*	T	1.000	0.013	0.962 - 1.039	0.016	0.987	48,190

Note: Ordinal logistic regression weights for touchscreen probable lifetime diagnosis, ordinal classification on variant, controling for sex, age, age^2, 10 European ancestry principle components, testing center, and batch.

Table S7.7: Severe recurrent MDD (MHF): main effect of variant

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z\rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.000	0.014	0.960 - 1.041	0.025	0.980	67,304
2	*BDNF*	*rs6265*	A	1.018	0.018	0.965 - 1.070	0.985	0.325	67,304
3	*COMT*	*rs4680*	G	1.030	0.014	0.988 - 1.071	2.095	0.036	67,304
4	*HTR2A*	*rs6311*	A	1.012	0.014	0.970 - 1.054	0.852	0.394	67,304
5	*TPH1*	*rs1800532*	A	1.004	0.014	0.962 - 1.046	0.274	0.784	67,304
6	*DRD4*	*VNTR*	7+ repeats	1.020	0.017	0.970 - 1.069	1.175	0.240	67,304
7	*DRD2*	*rs1800497*	T	1.009	0.017	0.958 - 1.060	0.505	0.614	67,304
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.023	0.013	0.985 - 1.061	1.792	0.073	67,304
9	*APOE*	*rs429358/rs7412*	ε-4	1.013	0.019	0.955 - 1.070	0.648	0.517	67,304
10	*MTHFR*	*rs1801133*	T	1.017	0.015	0.973 - 1.060	1.117	0.264	67,304
11	*CLOCK*	*rs1801260*	C	1.007	0.016	0.960 - 1.053	0.427	0.669	67,304
12	*SLC6A3*	*VNTR*	10+ repeats	1.010	0.016	0.963 - 1.057	0.610	0.542	67,304
13	*ACE*	*in/del*	deletion	1.004	0.014	0.962 - 1.045	0.251	0.802	67,304
14	*ABCB1*	*rs1045642*	C	1.000	0.014	0.959 - 1.042	0.033	0.974	67,304
15	*DRD3*	*rs6280*	C	1.001	0.015	0.958 - 1.045	0.091	0.928	67,304
16	*DBH*	*rs1611115*	T	1.019	0.017	0.967 - 1.070	1.058	0.290	66,231

Note: Logistic regression weights for Severe recurrent MDD (MHF) on variant, controling for sex, age, age^2, 10 European ancestry principle components, testing center, and batch.

Table S7.8: PGC lifetime MDD diagnosis: main effect of variant

	Gene	Polymorphism	Risk/counted allele	exp β	se_β	Corrected CI	$\lvert z\rvert$	p	n
1	*ABCB1*	*rs1045642*	C	0.992	0.006	0.976-1.009	1.393	0.164	349,649
2	*DBH*	*rs1611115*	T	1.005	0.008	0.981-1.029	0.561	0.575	287,678
3	*DRD2*	*rs1800497*	T	0.981	0.007	0.961-1.001	2.743	0.006	349,491
4	*TPH1*	*rs1800532*	A	1.008	0.006	0.992-1.025	1.464	0.148	247,142
5	*MTHFR*	*rs1801133*	T	0.993	0.006	0.976-1.011	1.153	0.250	349,311
6	*CLOCK*	*rs1801260*	C	1.014	0.006	0.995-1.032	2.177	0.029	349,530
7	*ACE*	*in/del*	deletion	0.995	0.005	0.979-1.011	0.891	0.380	349,632
8	*COMT*	*rs4680*	G	1.002	0.006	0.985-1.019	0.333	0.742	349,318
9	*BDNF*	*rs6265*	A	0.995	0.007	0.974-1.016	0.743	0.462	349,649
10	*DRD3*	*rs6280*	C	1.004	0.006	0.987-1.022	0.678	0.499	349,649
11	*HTR2A*	*rs6311*	A	0.998	0.006	0.982-1.015	0.321	0.743	349,157

Note: Differing sample sizes reflect that fact that not all polymorphisms were available in across all subsamples. Only polymorphisms that didn't require raw genotype to identify data were available.

S8 Polymorphism level $G \times E$ effects

Table S8.1: Estimated lifetime MDD diagnosis on variant × childhood trauma

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.998	0.022	0.934 - 1.066	0.101	0.919	115,204
2	*BDNF*	*rs6265*	A	1.007	0.029	0.925 - 1.097	0.250	0.803	115,204
3	*COMT*	*rs4680*	G	0.988	0.022	0.924 - 1.056	0.537	0.591	115,204
4	*HTR2A*	*rs6311*	A	0.966	0.023	0.902 - 1.033	1.526	0.127	115,204
5	*TPH1*	*rs1800532*	A	1.027	0.023	0.959 - 1.100	1.156	0.248	115,204
6	*DRD4*	*VNTR*	7+ repeats	0.993	0.027	0.916 - 1.076	0.265	0.791	115,204
7	*DRD2*	*rs1800497*	T	0.977	0.028	0.900 - 1.061	0.823	0.410	115,204
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.983	0.020	0.926 - 1.044	0.857	0.392	115,204
9	*APOE*	*rs429358/rs7412*	ε-4	1.039	0.032	0.946 - 1.141	1.209	0.227	115,204
10	*MTHFR*	*rs1801133*	T	0.989	0.024	0.922 - 1.061	0.451	0.652	115,204
11	*CLOCK*	*rs1801260*	C	0.994	0.025	0.922 - 1.071	0.253	0.800	115,204
12	*SLC6A3*	*VNTR*	10+ repeats	0.959	0.026	0.888 - 1.034	1.644	0.100	115,204
13	*ACE*	*in/del*	deletion	1.006	0.023	0.941 - 1.075	0.254	0.800	115,204
14	*ABCB1*	*rs1045642*	C	0.970	0.023	0.907 - 1.037	1.356	0.175	115,204
15	*DRD3*	*rs6280*	C	0.962	0.024	0.896 - 1.032	1.648	0.099	115,204
16	*DBH*	*rs1611115*	T	0.966	0.028	0.889 - 1.050	1.221	0.222	113,422

Note: Logistic regression weights for estimated lifetime MDD diagnosis on variant × childhood trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch and variant/moderator main effects.

Table S8.2: Estimated lifetime MDD diagnosis on variant × adult trauma

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.016	0.020	0.958 - 1.079	0.803	0.422	115,249
2	*BDNF*	*rs6265*	A	0.995	0.026	0.921 - 1.074	0.210	0.833	115,249
3	*COMT*	*rs4680*	G	0.997	0.020	0.939 - 1.058	0.168	0.867	115,249
4	*HTR2A*	*rs6311*	A	1.012	0.021	0.952 - 1.076	0.593	0.553	115,249
5	*TPH1*	*rs1800532*	A	1.026	0.021	0.965 - 1.091	1.222	0.222	115,249
6	*DRD4*	*VNTR*	7+ repeats	0.973	0.024	0.905 - 1.045	1.141	0.254	115,249
7	*DRD2*	*rs1800497*	T	1.016	0.025	0.943 - 1.095	0.644	0.520	115,249
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.985	0.019	0.932 - 1.040	0.828	0.408	115,249
9	*APOE*	*rs429358/rs7412*	ε-4	0.987	0.028	0.907 - 1.073	0.463	0.644	115,249
10	*MTHFR*	*rs1801133*	T	1.040	0.022	0.976 - 1.109	1.837	0.066	115,249
11	*CLOCK*	*rs1801260*	C	0.999	0.023	0.934 - 1.069	0.036	0.971	115,249
12	*SLC6A3*	*VNTR*	10+ repeats	0.983	0.023	0.918 - 1.053	0.724	0.469	115,249
13	*ACE*	*in/del*	deletion	0.999	0.020	0.941 - 1.061	0.037	0.971	115,249
14	*ABCB1*	*rs1045642*	C	0.983	0.020	0.926 - 1.044	0.835	0.404	115,249
15	*DRD3*	*rs6280*	C	0.986	0.021	0.925 - 1.050	0.677	0.498	115,249
16	*DBH*	*rs1611115*	T	0.967	0.025	0.897 - 1.042	1.334	0.182	113,466

Note: Logistic regression weights for estimated lifetime MDD diagnosis on variant × adult trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch and variant/moderator main effects.

Table S8.3: Estimated lifetime MDD diagnosis on variant × TDI

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.014	0.011	0.982 - 1.046	1.304	0.192	115,138
2	*BDNF*	*rs6265*	A	1.035	0.014	0.995 - 1.076	2.539	0.011	115,138
3	*COMT*	*rs4680*	G	1.008	0.011	0.977 - 1.040	0.791	0.429	115,138
4	*HTR2A*	*rs6311*	A	1.009	0.011	0.977 - 1.041	0.800	0.424	115,138
5	*TPH1*	*rs1800532*	A	1.008	0.011	0.976 - 1.041	0.741	0.459	115,138
6	*DRD4*	*VNTR*	7+ repeats	1.009	0.013	0.971 - 1.048	0.715	0.474	115,138
7	*DRD2*	*rs1800497*	T	1.006	0.013	0.967 - 1.046	0.469	0.639	115,138
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.010	0.010	0.981 - 1.039	1.024	0.306	115,138
9	*APOE*	*rs429358/rs7412*	ε-4	1.001	0.015	0.956 - 1.045	0.038	0.970	115,138
10	*MTHFR*	*rs1801133*	T	1.005	0.011	0.971 - 1.039	0.439	0.660	115,138
11	*CLOCK*	*rs1801260*	C	1.021	0.012	0.985 - 1.057	1.687	0.092	115,138
12	*SLC6A3*	*VNTR*	10+ repeats	1.002	0.012	0.966 - 1.039	0.193	0.847	115,138
13	*ACE*	*in/del*	deletion	1.004	0.011	0.973 - 1.036	0.406	0.685	115,138
14	*ABCB1*	*rs1045642*	C	1.000	0.011	0.968 - 1.032	0.029	0.977	115,138
15	*DRD3*	*rs6280*	C	1.007	0.011	0.973 - 1.040	0.573	0.567	115,138
16	*DBH*	*rs1611115*	T	1.003	0.013	0.964 - 1.042	0.214	0.831	113,356

Note: Logistic regression weights for estimated lifetime MDD diagnosis on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

Table S8.4: Current MDD severity on variant × recent trauma

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.991	0.019	0.938 - 1.048	0.464	0.643	115,246
2	*BDNF*	*rs6265*	A	1.017	0.024	0.946 - 1.092	0.683	0.495	115,246
3	*COMT*	*rs4680*	G	0.982	0.019	0.929 - 1.038	0.963	0.336	115,246
4	*HTR2A*	*rs6311*	A	1.005	0.019	0.950 - 1.064	0.275	0.783	115,246
5	*TPH1*	*rs1800532*	A	1.007	0.019	0.951 - 1.067	0.369	0.712	115,246
6	*DRD4*	*VNTR*	7+ repeats	0.994	0.023	0.930 - 1.063	0.264	0.792	115,246
7	*DRD2*	*rs1800497*	T	0.974	0.023	0.909 - 1.043	1.147	0.252	115,246
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.013	0.016	0.965 - 1.064	0.806	0.420	115,246
9	*APOE*	*rs429358/rs7412*	ε-4	0.953	0.027	0.880 - 1.032	1.800	0.072	115,246
10	*MTHFR*	*rs1801133*	T	1.005	0.020	0.948 - 1.066	0.255	0.799	115,246
11	*CLOCK*	*rs1801260*	C	1.007	0.021	0.946 - 1.072	0.342	0.732	115,246
12	*SLC6A3*	*VNTR*	10+ repeats	0.999	0.022	0.938 - 1.065	0.031	0.975	115,246
13	*ACE*	*in/del*	deletion	1.001	0.019	0.946 - 1.058	0.029	0.977	115,246
14	*ABCB1*	*rs1045642*	C	1.039	0.019	0.982 - 1.099	2.018	0.044	115,246
15	*DRD3*	*rs6280*	C	0.974	0.020	0.919 - 1.033	1.315	0.189	115,246
16	*DBH*	*rs1611115*	T	0.985	0.024	0.919 - 1.057	0.627	0.531	113,463

Note: Negative binomial regression weights for current MDD severity on variant × recent trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch and variant/moderator main effects.

Table S8.5: Conditional lifetime symptom count on variant × childhood trauma

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\|z\|$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	-0.012	0.022	-0.077 - 0.054	0.524	0.601	62,108
2	*BDNF*	*rs6265*	A	0.006	0.028	-0.079 - 0.090	0.198	0.843	62,108
3	*COMT*	*rs4680*	G	0.005	0.022	-0.061 - 0.071	0.225	0.822	62,108
4	*HTR2A*	*rs6311*	A	0.009	0.023	-0.058 - 0.076	0.401	0.688	62,108
5	*TPH1*	*rs1800532*	A	0.045	0.023	-0.022 - 0.113	1.973	0.049	62,108
6	*DRD4*	*VNTR*	7+ repeats	0.015	0.027	-0.065 - 0.094	0.542	0.588	62,108
7	*DRD2*	*rs1800497*	T	0.005	0.028	-0.076 - 0.087	0.190	0.849	62,108
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	-0.019	0.020	-0.078 - 0.041	0.927	0.354	62,108
9	*APOE*	*rs429358/rs7412*	ε-4	0.004	0.031	-0.088 - 0.096	0.119	0.906	62,108
10	*MTHFR*	*rs1801133*	T	-0.023	0.024	-0.092 - 0.047	0.972	0.331	62,108
11	*CLOCK*	*rs1801260*	C	-0.005	0.025	-0.079 - 0.069	0.199	0.842	62,108
12	*SLC6A3*	*VNTR*	10+ repeats	-0.023	0.025	-0.097 - 0.052	0.889	0.374	62,108
13	*ACE*	*in/del*	deletion	-0.007	0.022	-0.072 - 0.059	0.299	0.765	62,108
14	*ABCB1*	*rs1045642*	C	0.031	0.022	-0.035 - 0.097	1.403	0.161	62,108
15	*DRD3*	*rs6280*	C	-0.049	0.024	-0.118 - 0.021	2.071	0.038	62,108
16	*DBH*	*rs1611115*	T	0.001	0.028	-0.081 - 0.083	0.036	0.972	61,178

Note: Linear regression weights for conditional lifetime symptom count on variant × childhood trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch and variant/moderator main effects.

Table S8.6: Conditional lifetime symptom count on variant × adult trauma

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\|z\|$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.001	0.020	-0.057 - 0.059	0.056	0.956	62,135
2	*BDNF*	*rs6265*	A	0.026	0.025	-0.049 - 0.101	1.041	0.298	62,135
3	*COMT*	*rs4680*	G	-0.020	0.020	-0.079 - 0.038	1.036	0.300	62,135
4	*HTR2A*	*rs6311*	A	-0.001	0.020	-0.061 - 0.059	0.044	0.965	62,135
5	*TPH1*	*rs1800532*	A	0.011	0.020	-0.049 - 0.071	0.552	0.581	62,135
6	*DRD4*	*VNTR*	7+ repeats	-0.023	0.024	-0.093 - 0.047	0.964	0.335	62,135
7	*DRD2*	*rs1800497*	T	0.008	0.025	-0.065 - 0.081	0.342	0.732	62,135
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	-0.022	0.018	-0.075 - 0.032	1.202	0.229	62,135
9	*APOE*	*rs429358/rs7412*	ε-4	-0.013	0.028	-0.096 - 0.069	0.481	0.630	62,135
10	*MTHFR*	*rs1801133*	T	0.011	0.021	-0.051 - 0.073	0.521	0.603	62,135
11	*CLOCK*	*rs1801260*	C	0.040	0.022	-0.027 - 0.106	1.768	0.077	62,135
12	*SLC6A3*	*VNTR*	10+ repeats	-0.023	0.023	-0.090 - 0.044	1.012	0.311	62,135
13	*ACE*	*in/del*	deletion	0.023	0.020	-0.036 - 0.082	1.168	0.243	62,135
14	*ABCB1*	*rs1045642*	C	-0.004	0.020	-0.062 - 0.055	0.196	0.845	62,135
15	*DRD3*	*rs6280*	C	-0.020	0.021	-0.082 - 0.042	0.941	0.347	62,135
16	*DBH*	*rs1611115*	T	-0.033	0.025	-0.106 - 0.041	1.310	0.190	61,205

Note: Linear regression weights for conditional lifetime symptom count on variant × adult trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch and variant/moderator main effects.

Table S8.7: Conditional lifetime symptom count on variant × TDI

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.008	0.011	-0.024 - 0.039	0.740	0.459	62,069
2	*BDNF*	*rs6265*	A	0.012	0.014	-0.028 - 0.052	0.890	0.374	62,069
3	*COMT*	*rs4680*	G	0.002	0.011	-0.029 - 0.034	0.217	0.828	62,069
4	*HTR2A*	*rs6311*	A	0.005	0.011	-0.027 - 0.037	0.432	0.666	62,069
5	*TPH1*	*rs1800532*	A	-0.007	0.011	-0.039 - 0.025	0.646	0.518	62,069
6	*DRD4*	*VNTR*	7+ repeats	0.002	0.013	-0.035 - 0.040	0.194	0.846	62,069
7	*DRD2*	*rs1800497*	T	-0.019	0.013	-0.058 - 0.020	1.416	0.157	62,069
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	-0.024	0.010	-0.052 - 0.005	2.452	0.014	62,069
9	*APOE*	*rs429358/rs7412*	ε-4	0.002	0.015	-0.043 - 0.046	0.111	0.912	62,069
10	*MTHFR*	*rs1801133*	T	0.004	0.011	-0.030 - 0.037	0.328	0.743	62,069
11	*CLOCK*	*rs1801260*	C	0.014	0.012	-0.021 - 0.050	1.186	0.235	62,069
12	*SLC6A3*	*VNTR*	10+ repeats	0.008	0.012	-0.028 - 0.044	0.627	0.531	62,069
13	*ACE*	*in/del*	deletion	-0.010	0.011	-0.042 - 0.021	0.974	0.330	62,069
14	*ABCB1*	*rs1045642*	C	0.002	0.011	-0.030 - 0.033	0.169	0.866	62,069
15	*DRD3*	*rs6280*	C	-0.004	0.011	-0.037 - 0.029	0.351	0.726	62,069
16	*DBH*	*rs1611115*	T	0.005	0.013	-0.034 - 0.044	0.362	0.717	61,139

Note: Linear regression weights for conditional lifetime symptom count on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

Table S8.8: Conditional lifetime symptom count on variant × stressor-induced depression

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	-0.030	0.022	-0.096 - 0.036	1.357	0.175	61,888
2	*BDNF*	*rs6265*	A	0.007	0.029	-0.078 - 0.092	0.234	0.815	61,888
3	*COMT*	*rs4680*	G	0.045	0.022	-0.021 - 0.111	2.017	0.044	61,888
4	*HTR2A*	*rs6311*	A	0.026	0.023	-0.041 - 0.093	1.138	0.255	61,888
5	*TPH1*	*rs1800532*	A	0.015	0.023	-0.053 - 0.083	0.644	0.519	61,888
6	*DRD4*	*VNTR*	7+ repeats	0.037	0.027	-0.043 - 0.116	1.353	0.176	61,888
7	*DRD2*	*rs1800497*	T	0.038	0.028	-0.045 - 0.121	1.365	0.172	61,888
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	-0.010	0.020	-0.069 - 0.049	0.507	0.612	61,888
9	*APOE*	*rs429358/rs7412*	ε-4	0.023	0.031	-0.070 - 0.116	0.735	0.462	61,888
10	*MTHFR*	*rs1801133*	T	-0.025	0.024	-0.096 - 0.045	1.072	0.284	61,888
11	*CLOCK*	*rs1801260*	C	0.023	0.025	-0.052 - 0.098	0.907	0.364	61,888
12	*SLC6A3*	*VNTR*	10+ repeats	-0.001	0.026	-0.077 - 0.075	0.027	0.979	61,888
13	*ACE*	*in/del*	deletion	0.014	0.022	-0.052 - 0.080	0.627	0.531	61,888
14	*ABCB1*	*rs1045642*	C	0.020	0.022	-0.046 - 0.086	0.884	0.377	61,888
15	*DRD3*	*rs6280*	C	-0.045	0.024	-0.115 - 0.025	1.898	0.058	61,888
16	*DBH*	*rs1611115*	T	0.028	0.028	-0.055 - 0.110	0.998	0.318	60,960

Note: Linear regression weights for conditional lifetime symptom count on variant × stressor-induced depression controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch and variant/moderator main effects.

Table S8.9: Lifetime episode count on variant × childhood trauma

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.000	0.019	0.945 - 1.059	0.019	0.985	112,216
2	*BDNF*	*rs6265*	A	1.018	0.025	0.946 - 1.096	0.718	0.473	112,216
3	*COMT*	*rs4680*	G	0.983	0.019	0.929 - 1.041	0.863	0.388	112,216
4	*HTR2A*	*rs6311*	A	0.974	0.020	0.919 - 1.033	1.321	0.186	112,216
5	*TPH1*	*rs1800532*	A	0.996	0.020	0.939 - 1.056	0.204	0.839	112,216
6	*DRD4*	*VNTR*	7+ repeats	1.002	0.023	0.935 - 1.074	0.085	0.932	112,216
7	*DRD2*	*rs1800497*	T	1.003	0.024	0.934 - 1.077	0.122	0.903	112,216
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.983	0.017	0.935 - 1.035	0.974	0.330	112,216
9	*APOE*	*rs429358/rs7412*	ε-4	1.025	0.027	0.946 - 1.111	0.913	0.361	112,216
10	*MTHFR*	*rs1801133*	T	0.970	0.021	0.913 - 1.031	1.486	0.137	112,216
11	*CLOCK*	*rs1801260*	C	1.001	0.022	0.938 - 1.068	0.050	0.960	112,216
12	*SLC6A3*	*VNTR*	10+ repeats	0.973	0.022	0.911 - 1.039	1.228	0.219	112,216
13	*ACE*	*in/del*	deletion	0.992	0.019	0.937 - 1.051	0.403	0.687	112,216
14	*ABCB1*	*rs1045642*	C	0.976	0.019	0.921 - 1.034	1.254	0.210	112,216
15	*DRD3*	*rs6280*	C	0.993	0.021	0.935 - 1.055	0.339	0.735	112,216
16	*DBH*	*rs1611115*	T	0.968	0.024	0.901 - 1.040	1.324	0.186	110,469

Note: Ordinal logistic regression weights for lifetime episode count on variant × childhood trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch and variant/moderator main effects.

Table S8.10: Lifetime episode count on variant × adult trauma

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.007	0.016	0.960 - 1.058	0.453	0.650	112,255
2	*BDNF*	*rs6265*	A	0.985	0.021	0.925 - 1.048	0.720	0.471	112,255
3	*COMT*	*rs4680*	G	0.997	0.016	0.949 - 1.047	0.192	0.848	112,255
4	*HTR2A*	*rs6311*	A	0.998	0.017	0.950 - 1.049	0.102	0.919	112,255
5	*TPH1*	*rs1800532*	A	1.011	0.017	0.961 - 1.063	0.634	0.526	112,255
6	*DRD4*	*VNTR*	7+ repeats	1.004	0.020	0.947 - 1.065	0.226	0.821	112,255
7	*DRD2*	*rs1800497*	T	1.011	0.021	0.951 - 1.074	0.518	0.605	112,255
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.007	0.015	0.964 - 1.052	0.442	0.658	112,255
9	*APOE*	*rs429358/rs7412*	ε-4	1.001	0.023	0.935 - 1.072	0.059	0.953	112,255
10	*MTHFR*	*rs1801133*	T	1.034	0.018	0.982 - 1.089	1.896	0.058	112,255
11	*CLOCK*	*rs1801260*	C	0.997	0.019	0.944 - 1.054	0.149	0.882	112,255
12	*SLC6A3*	*VNTR*	10+ repeats	1.002	0.019	0.947 - 1.060	0.102	0.919	112,255
13	*ACE*	*in/del*	deletion	0.999	0.017	0.951 - 1.049	0.083	0.934	112,255
14	*ABCB1*	*rs1045642*	C	0.999	0.017	0.952 - 1.050	0.037	0.971	112,255
15	*DRD3*	*rs6280*	C	0.978	0.017	0.929 - 1.029	1.293	0.196	112,255
16	*DBH*	*rs1611115*	T	0.987	0.021	0.928 - 1.049	0.644	0.519	110,507

Note: Ordinal logistic regression weights for lifetime episode count on variant × adult trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch and variant/moderator main effects.

Table S8.11: Lifetime episode count on variant × TDI

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.019	0.009	0.992 - 1.046	2.042	0.041	112,144
2	*BDNF*	*rs6265*	A	1.021	0.012	0.987 - 1.056	1.790	0.073	112,144
3	*COMT*	*rs4680*	G	1.012	0.009	0.985 - 1.039	1.272	0.204	112,144
4	*HTR2A*	*rs6311*	A	1.006	0.009	0.978 - 1.033	0.602	0.547	112,144
5	*TPH1*	*rs1800532*	A	1.004	0.009	0.977 - 1.032	0.467	0.641	112,144
6	*DRD4*	*VNTR*	7+ repeats	1.005	0.011	0.972 - 1.037	0.439	0.661	112,144
7	*DRD2*	*rs1800497*	T	1.009	0.011	0.975 - 1.042	0.752	0.452	112,144
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.001	0.008	0.977 - 1.025	0.173	0.862	112,144
9	*APOE*	*rs429358/rs7412*	ε-4	1.005	0.013	0.967 - 1.043	0.408	0.684	112,144
10	*MTHFR*	*rs1801133*	T	1.002	0.010	0.974 - 1.031	0.251	0.802	112,144
11	*CLOCK*	*rs1801260*	C	1.026	0.010	0.995 - 1.056	2.455	0.014	112,144
12	*SLC6A3*	*VNTR*	10+ repeats	1.005	0.010	0.974 - 1.035	0.436	0.663	112,144
13	*ACE*	*in/del*	deletion	1.015	0.009	0.988 - 1.042	1.612	0.107	112,144
14	*ABCB1*	*rs1045642*	C	1.017	0.009	0.990 - 1.044	1.868	0.062	112,144
15	*DRD3*	*rs6280*	C	1.001	0.010	0.973 - 1.030	0.125	0.901	112,144
16	*DBH*	*rs1611115*	T	1.014	0.011	0.981 - 1.048	1.255	0.210	110,397

Note: Ordinal logistic regression weights for lifetime episode count on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

Table S8.12: Touchscreen probable lifetime diagnosis on variant × TDI

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.007	0.012	0.972 - 1.042	0.592	0.554	90,812
2	*BDNF*	*rs6265*	A	1.013	0.015	0.969 - 1.058	0.873	0.383	90,812
3	*COMT*	*rs4680*	G	1.012	0.012	0.977 - 1.047	1.028	0.304	90,812
4	*HTR2A*	*rs6311*	A	1.001	0.012	0.966 - 1.037	0.120	0.904	90,812
5	*TPH1*	*rs1800532*	A	1.016	0.012	0.980 - 1.051	1.308	0.191	90,812
6	*DRD4*	*VNTR*	7+ repeats	1.002	0.014	0.960 - 1.044	0.156	0.876	90,812
7	*DRD2*	*rs1800497*	T	1.017	0.015	0.973 - 1.060	1.121	0.262	90,812
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.003	0.010	0.972 - 1.034	0.297	0.766	90,812
9	*APOE*	*rs429358/rs7412*	ε-4	1.022	0.016	0.974 - 1.070	1.339	0.180	90,812
10	*MTHFR*	*rs1801133*	T	1.011	0.012	0.974 - 1.047	0.846	0.398	90,812
11	*CLOCK*	*rs1801260*	C	1.009	0.013	0.969 - 1.048	0.661	0.509	90,812
12	*SLC6A3*	*VNTR*	10+ repeats	1.020	0.013	0.980 - 1.060	1.484	0.138	90,812
13	*ACE*	*in/del*	deletion	1.003	0.012	0.968 - 1.038	0.236	0.814	90,812
14	*ABCB1*	*rs1045642*	C	1.015	0.012	0.980 - 1.050	1.244	0.213	90,812
15	*DRD3*	*rs6280*	C	1.017	0.012	0.980 - 1.053	1.320	0.187	90,812
16	*DBH*	*rs1611115*	T	1.011	0.015	0.967 - 1.054	0.719	0.472	89,397

Note: Logistic regression weights for touchscreen probable lifetime diagnosis on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

Table S8.13: Touchscreen probable lifetime diagnosis, ordinal classification on variant × TDI

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.009	0.012	0.975 - 1.043	0.743	0.457	90,812
2	*BDNF*	*rs6265*	A	1.018	0.015	0.975 - 1.062	1.234	0.217	90,812
3	*COMT*	*rs4680*	G	1.007	0.012	0.973 - 1.041	0.600	0.548	90,812
4	*HTR2A*	*rs6311*	A	1.005	0.012	0.970 - 1.040	0.410	0.682	90,812
5	*TPH1*	*rs1800532*	A	1.016	0.012	0.981 - 1.051	1.343	0.179	90,812
6	*DRD4*	*VNTR*	7+ repeats	1.001	0.014	0.960 - 1.042	0.101	0.920	90,812
7	*DRD2*	*rs1800497*	T	1.006	0.014	0.963 - 1.049	0.406	0.685	90,812
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.005	0.010	0.975 - 1.035	0.474	0.636	90,812
9	*APOE*	*rs429358/rs7412*	ε-4	1.026	0.016	0.978 - 1.073	1.585	0.113	90,812
10	*MTHFR*	*rs1801133*	T	1.009	0.012	0.973 - 1.045	0.740	0.459	90,812
11	*CLOCK*	*rs1801260*	C	1.002	0.013	0.963 - 1.040	0.122	0.903	90,812
12	*SLC6A3*	*VNTR*	10+ repeats	1.018	0.013	0.979 - 1.057	1.346	0.178	90,812
13	*ACE*	*in/del*	deletion	1.002	0.012	0.967 - 1.036	0.147	0.883	90,812
14	*ABCB1*	*rs1045642*	C	1.013	0.012	0.979 - 1.047	1.095	0.273	90,812
15	*DRD3*	*rs6280*	C	1.014	0.012	0.978 - 1.050	1.171	0.242	90,812
16	*DBH*	*rs1611115*	T	1.011	0.014	0.969 - 1.053	0.755	0.450	89,397

Note: Ordinal logistic regression weights for touchscreen probable lifetime diagnosis, ordinal classification on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

Table S8.14: Severe recurrent MDD (MHF) on variant × adult trauma

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.016	0.014	-0.027 - 0.058	1.089	0.276	67,304
2	*BDNF*	*rs6265*	A	-0.009	0.019	-0.064 - 0.047	0.461	0.645	67,304
3	*COMT*	*rs4680*	G	0.004	0.015	-0.039 - 0.047	0.254	0.800	67,304
4	*HTR2A*	*rs6311*	A	-0.006	0.015	-0.050 - 0.037	0.437	0.662	67,304
5	*TPH1*	*rs1800532*	A	-0.021	0.015	-0.065 - 0.022	1.441	0.150	67,304
6	*DRD4*	*VNTR*	7+ repeats	-0.012	0.018	-0.064 - 0.040	0.682	0.495	67,304
7	*DRD2*	*rs1800497*	T	-0.013	0.018	-0.065 - 0.040	0.724	0.469	67,304
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.001	0.013	-0.038 - 0.040	0.063	0.950	67,304
9	*APOE*	*rs429358/rs7412*	ε-4	0.000	0.021	-0.060 - 0.061	0.022	0.982	67,304
10	*MTHFR*	*rs1801133*	T	-0.001	0.015	-0.046 - 0.044	0.070	0.944	67,304
11	*CLOCK*	*rs1801260*	C	-0.002	0.016	-0.050 - 0.046	0.099	0.921	67,304
12	*SLC6A3*	*VNTR*	10+ repeats	0.001	0.017	-0.048 - 0.050	0.070	0.944	67,304
13	*ACE*	*in/del*	deletion	0.002	0.014	-0.040 - 0.045	0.162	0.871	67,304
14	*ABCB1*	*rs1045642*	C	-0.002	0.015	-0.045 - 0.041	0.115	0.908	67,304
15	*DRD3*	*rs6280*	C	-0.025	0.015	-0.070 - 0.020	1.655	0.098	67,304
16	*DBH*	*rs1611115*	T	-0.002	0.018	-0.055 - 0.051	0.092	0.927	66,231

Note: Linear regression weights for Severe recurrent MDD (MHF) on variant × adult trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

Table S8.15: Severe recurrent MDD (MHF) on variant × childhood trauma

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.008	0.014	-0.032 - 0.048	0.586	0.558	67,304
2	*BDNF*	*rs6265*	A	0.021	0.018	-0.032 - 0.073	1.162	0.245	67,304
3	*COMT*	*rs4680*	G	0.012	0.014	-0.028 - 0.052	0.885	0.376	67,304
4	*HTR2A*	*rs6311*	A	0.018	0.014	-0.024 - 0.059	1.262	0.207	67,304
5	*TPH1*	*rs1800532*	A	-0.008	0.014	-0.050 - 0.033	0.597	0.551	67,304
6	*DRD4*	*VNTR*	7+ repeats	-0.004	0.016	-0.052 - 0.044	0.241	0.810	67,304
7	*DRD2*	*rs1800497*	T	-0.006	0.017	-0.056 - 0.044	0.364	0.716	67,304
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	-0.009	0.013	-0.046 - 0.028	0.699	0.485	67,304
9	*APOE*	*rs429358/rs7412*	ε-4	0.006	0.019	-0.052 - 0.063	0.291	0.771	67,304
10	*MTHFR*	*rs1801133*	T	0.027	0.014	-0.016 - 0.070	1.870	0.061	67,304
11	*CLOCK*	*rs1801260*	C	-0.003	0.015	-0.049 - 0.042	0.229	0.819	67,304
12	*SLC6A3*	*VNTR*	10+ repeats	-0.024	0.016	-0.070 - 0.022	1.522	0.128	67,304
13	*ACE*	*in/del*	deletion	0.011	0.014	-0.030 - 0.051	0.773	0.440	67,304
14	*ABCB1*	*rs1045642*	C	0.006	0.014	-0.034 - 0.046	0.438	0.661	67,304
15	*DRD3*	*rs6280*	C	-0.008	0.014	-0.050 - 0.035	0.528	0.597	67,304
16	*DBH*	*rs1611115*	T	-0.024	0.017	-0.075 - 0.026	1.430	0.153	66,231

Note: Logistic regression weights for Severe recurrent MDD (MHF) on variant × childhood trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

Table S8.16: Severe recurrent MDD (MHF) on variant × TDI

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.009	0.005	0.994 - 1.024	1.768	0.077	67,237
2	*BDNF*	*rs6265*	A	1.012	0.007	0.992 - 1.031	1.755	0.079	67,237
3	*COMT*	*rs4680*	G	1.005	0.005	0.990 - 1.020	1.044	0.296	67,237
4	*HTR2A*	*rs6311*	A	1.001	0.005	0.985 - 1.016	0.147	0.883	67,237
5	*TPH1*	*rs1800532*	A	1.001	0.005	0.985 - 1.016	0.135	0.893	67,237
6	*DRD4*	*VNTR*	7+ repeats	1.009	0.006	0.991 - 1.027	1.418	0.156	67,237
7	*DRD2*	*rs1800497*	T	1.002	0.006	0.983 - 1.021	0.368	0.713	67,237
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.004	0.005	0.990 - 1.017	0.757	0.449	67,237
9	*APOE*	*rs429358/rs7412*	ε-4	1.005	0.007	0.983 - 1.026	0.628	0.530	67,237
10	*MTHFR*	*rs1801133*	T	1.001	0.005	0.985 - 1.017	0.179	0.858	67,237
11	*CLOCK*	*rs1801260*	C	1.015	0.006	0.997 - 1.032	2.496	0.013	67,237
12	*SLC6A3*	*VNTR*	10+ repeats	1.007	0.006	0.989 - 1.024	1.127	0.260	67,237
13	*ACE*	*in/del*	deletion	1.007	0.005	0.992 - 1.022	1.316	0.188	67,237
14	*ABCB1*	*rs1045642*	C	1.002	0.005	0.987 - 1.017	0.351	0.726	67,237
15	*DRD3*	*rs6280*	C	1.003	0.005	0.987 - 1.019	0.602	0.547	67,237
16	*DBH*	*rs1611115*	T	1.008	0.006	0.989 - 1.027	1.222	0.222	66,165

Note: Logistic regression weights for Severe recurrent MDD (MHF) on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

S9 Polymorphism level $G \times E$ effects (alternate scale)

Table S9.1: Estimated lifetime MDD diagnosis on variant × childhood trauma, alternate scale

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	-0.000	0.004	-0.013 - 0.012	0.104	0.917	115,204
2	*BDNF*	*rs6265*	A	0.002	0.006	-0.015 - 0.018	0.298	0.766	115,204
3	*COMT*	*rs4680*	G	-0.003	0.004	-0.016 - 0.010	0.701	0.483	115,204
4	*HTR2A*	*rs6311*	A	-0.008	0.004	-0.021 - 0.005	1.798	0.072	115,204
5	*TPH1*	*rs1800532*	A	0.005	0.004	-0.008 - 0.018	1.176	0.240	115,204
6	*DRD4*	*VNTR*	7+ repeats	-0.001	0.005	-0.016 - 0.014	0.174	0.862	115,204
7	*DRD2*	*rs1800497*	T	-0.006	0.005	-0.021 - 0.010	1.030	0.303	115,204
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	-0.003	0.004	-0.014 - 0.008	0.851	0.395	115,204
9	*APOE*	*rs429358/rs7412*	ε-4	0.007	0.006	-0.011 - 0.025	1.129	0.259	115,204
10	*MTHFR*	*rs1801133*	T	-0.002	0.005	-0.015 - 0.012	0.345	0.730	115,204
11	*CLOCK*	*rs1801260*	C	-0.001	0.005	-0.015 - 0.014	0.157	0.875	115,204
12	*SLC6A3*	*VNTR*	10+ repeats	-0.008	0.005	-0.023 - 0.006	1.650	0.099	115,204
13	*ACE*	*in/del*	deletion	0.001	0.004	-0.012 - 0.014	0.269	0.788	115,204
14	*ABCB1*	*rs1045642*	C	-0.006	0.004	-0.019 - 0.007	1.403	0.160	115,204
15	*DRD3*	*rs6280*	C	-0.009	0.005	-0.022 - 0.005	1.894	0.058	115,204
16	*DBH*	*rs1611115*	T	-0.008	0.005	-0.024 - 0.008	1.440	0.150	113,422

Note: Linear regression weights for estimated lifetime MDD diagnosis on variant × childhood trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

Table S9.2: Estimated lifetime MDD diagnosis on variant × adult trauma, alternate scale

	Gene	Polymorphism	Risk/counted allele	$\exp(\beta)$	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.003	0.004	0.992 - 1.014	0.780	0.435	115,249
2	*BDNF*	*rs6265*	A	0.999	0.005	0.986 - 1.013	0.152	0.880	115,249
3	*COMT*	*rs4680*	G	0.999	0.004	0.988 - 1.010	0.325	0.745	115,249
4	*HTR2A*	*rs6311*	A	1.001	0.004	0.990 - 1.012	0.274	0.784	115,249
5	*TPH1*	*rs1800532*	A	1.004	0.004	0.993 - 1.016	1.180	0.238	115,249
6	*DRD4*	*VNTR*	7+ repeats	0.996	0.004	0.983 - 1.009	0.871	0.384	115,249
7	*DRD2*	*rs1800497*	T	1.002	0.005	0.989 - 1.015	0.432	0.666	115,249
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.998	0.003	0.988 - 1.007	0.754	0.451	115,249
9	*APOE*	*rs429358/rs7412*	ε-4	0.997	0.005	0.982 - 1.012	0.567	0.571	115,249
10	*MTHFR*	*rs1801133*	T	1.007	0.004	0.996 - 1.019	1.930	0.054	115,249
11	*CLOCK*	*rs1801260*	C	1.001	0.004	0.989 - 1.013	0.147	0.883	115,249
12	*SLC6A3*	*VNTR*	10+ repeats	0.998	0.004	0.986 - 1.010	0.512	0.609	115,249
13	*ACE*	*in/del*	deletion	1.000	0.004	0.989 - 1.011	0.004	0.997	115,249
14	*ABCB1*	*rs1045642*	C	0.997	0.004	0.986 - 1.008	0.790	0.430	115,249
15	*DRD3*	*rs6280*	C	0.997	0.004	0.985 - 1.008	0.877	0.380	115,249
16	*DBH*	*rs1611115*	T	0.993	0.005	0.980 - 1.007	1.449	0.147	113,466

Note: Linear regression weights for estimated lifetime MDD diagnosis on variant × adult trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

Table S9.3: Estimated lifetime MDD diagnosis on variant × TDI, alternate scale

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.003	0.002	-0.003 - 0.009	1.358	0.174	115,138
2	*BDNF*	*rs6265*	A	0.007	0.003	-0.001 - 0.015	2.635	0.008	115,138
3	*COMT*	*rs4680*	G	-0.002	0.002	-0.008 - 0.004	0.931	0.352	115,138
4	*HTR2A*	*rs6311*	A	0.002	0.002	-0.004 - 0.008	0.905	0.366	115,138
5	*TPH1*	*rs1800532*	A	-0.001	0.002	-0.008 - 0.005	0.703	0.482	115,138
6	*DRD4*	*VNTR*	7+ repeats	-0.001	0.002	-0.009 - 0.006	0.593	0.553	115,138
7	*DRD2*	*rs1800497*	T	-0.001	0.003	-0.008 - 0.007	0.393	0.694	115,138
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	-0.002	0.002	-0.007 - 0.004	0.936	0.349	115,138
9	*APOE*	*rs429358/rs7412*	ε-4	-0.000	0.003	-0.009 - 0.008	0.103	0.918	115,138
10	*MTHFR*	*rs1801133*	T	-0.001	0.002	-0.007 - 0.005	0.432	0.666	115,138
11	*CLOCK*	*rs1801260*	C	0.004	0.002	-0.003 - 0.011	1.698	0.090	115,138
12	*SLC6A3*	*VNTR*	10+ repeats	-0.000	0.002	-0.007 - 0.007	0.160	0.873	115,138
13	*ACE*	*in/del*	deletion	-0.001	0.002	-0.007 - 0.005	0.459	0.646	115,138
14	*ABCB1*	*rs1045642*	C	-0.000	0.002	-0.006 - 0.006	0.099	0.922	115,138
15	*DRD3*	*rs6280*	C	0.001	0.002	-0.005 - 0.007	0.394	0.694	115,138
16	*DBH*	*rs1611115*	T	0.001	0.003	-0.007 - 0.008	0.250	0.802	113,356

Note: Linear regression weights for estimated lifetime MDD diagnosis on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

Table S9.4: Current MDD severity (MHF) on variant × recent trauma, alternate scale

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	-0.035	0.049	-0.178 - 0.108	0.721	0.471	115,246
2	*BDNF*	*rs6265*	A	0.057	0.063	-0.129 - 0.243	0.910	0.363	115,246
3	*COMT*	*rs4680*	G	-0.087	0.049	-0.232 - 0.058	1.774	0.076	115,246
4	*HTR2A*	*rs6311*	A	0.003	0.050	-0.143 - 0.150	0.066	0.947	115,246
5	*TPH1*	*rs1800532*	A	0.011	0.050	-0.138 - 0.160	0.225	0.822	115,246
6	*DRD4*	*VNTR*	7+ repeats	-0.000	0.059	-0.173 - 0.173	0.006	0.996	115,246
7	*DRD2*	*rs1800497*	T	-0.079	0.061	-0.258 - 0.101	1.295	0.195	115,246
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.026	0.042	-0.099 - 0.152	0.612	0.540	115,246
9	*APOE*	*rs429358/rs7412*	ε-4	-0.182	0.069	-0.388 - 0.023	2.628	0.009	115,246
10	*MTHFR*	*rs1801133*	T	0.031	0.052	-0.122 - 0.183	0.597	0.551	115,246
11	*CLOCK*	*rs1801260*	C	0.012	0.055	-0.150 - 0.175	0.224	0.822	115,246
12	*SLC6A3*	*VNTR*	10+ repeats	-0.009	0.056	-0.175 - 0.156	0.165	0.869	115,246
13	*ACE*	*in/del*	deletion	-0.007	0.049	-0.153 - 0.138	0.150	0.881	115,246
14	*ABCB1*	*rs1045642*	C	0.108	0.049	-0.037 - 0.254	2.209	0.027	115,246
15	*DRD3*	*rs6280*	C	-0.111	0.051	-0.262 - 0.041	2.162	0.031	115,246
16	*DBH*	*rs1611115*	T	-0.038	0.061	-0.219 - 0.143	0.617	0.537	113,463

Note: Linear regression weights for current MDD severity (MHF) on variant × recent trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

Table S9.5: Conditional lifetime symptom count on variant × childhood trauma, alternate scale

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.999	0.023	0.933 - 1.070	0.026	0.979	62,108
2	*BDNF*	*rs6265*	A	1.018	0.030	0.932 - 1.112	0.598	0.550	62,108
3	*COMT*	*rs4680*	G	1.001	0.023	0.934 - 1.072	0.030	0.976	62,108
4	*HTR2A*	*rs6311*	A	1.016	0.024	0.947 - 1.089	0.652	0.514	62,108
5	*TPH1*	*rs1800532*	A	1.048	0.024	0.976 - 1.124	1.947	0.052	62,108
6	*DRD4*	*VNTR*	7+ repeats	1.023	0.028	0.941 - 1.111	0.799	0.424	62,108
7	*DRD2*	*rs1800497*	T	1.006	0.029	0.923 - 1.096	0.212	0.832	62,108
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.983	0.021	0.924 - 1.047	0.792	0.428	62,108
9	*APOE*	*rs429358/rs7412*	ε-4	0.994	0.033	0.903 - 1.094	0.186	0.853	62,108
10	*MTHFR*	*rs1801133*	T	0.975	0.025	0.906 - 1.049	1.028	0.304	62,108
11	*CLOCK*	*rs1801260*	C	0.996	0.026	0.922 - 1.076	0.142	0.887	62,108
12	*SLC6A3*	*VNTR*	10+ repeats	0.970	0.027	0.897 - 1.050	1.134	0.257	62,108
13	*ACE*	*in/del*	deletion	0.998	0.023	0.932 - 1.069	0.067	0.947	62,108
14	*ABCB1*	*rs1045642*	C	1.034	0.023	0.965 - 1.108	1.429	0.153	62,108
15	*DRD3*	*rs6280*	C	0.957	0.025	0.890 - 1.030	1.768	0.077	62,108
16	*DBH*	*rs1611115*	T	1.011	0.029	0.927 - 1.102	0.373	0.709	61,178

Note: Ordinal logistic regression weights for conditional lifetime symptom count on variant × childhood trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

Table S9.6: Conditional lifetime symptom count on variant × adult trauma, alternate scale

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.009	0.021	0.949 - 1.072	0.413	0.679	62,135
2	*BDNF*	*rs6265*	A	1.036	0.026	0.958 - 1.120	1.333	0.182	62,135
3	*COMT*	*rs4680*	G	0.971	0.021	0.914 - 1.032	1.410	0.159	62,135
4	*HTR2A*	*rs6311*	A	0.998	0.021	0.938 - 1.062	0.081	0.935	62,135
5	*TPH1*	*rs1800532*	A	1.013	0.021	0.952 - 1.079	0.627	0.531	62,135
6	*DRD4*	*VNTR*	7+ repeats	0.980	0.025	0.910 - 1.055	0.818	0.413	62,135
7	*DRD2*	*rs1800497*	T	1.008	0.026	0.934 - 1.088	0.308	0.758	62,135
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.981	0.019	0.928 - 1.037	1.010	0.313	62,135
9	*APOE*	*rs429358/rs7412*	ε-4	0.981	0.029	0.900 - 1.069	0.658	0.511	62,135
10	*MTHFR*	*rs1801133*	T	1.011	0.022	0.947 - 1.078	0.482	0.630	62,135
11	*CLOCK*	*rs1801260*	C	1.038	0.023	0.969 - 1.112	1.591	0.112	62,135
12	*SLC6A3*	*VNTR*	10+ repeats	0.968	0.024	0.903 - 1.039	1.350	0.177	62,135
13	*ACE*	*in/del*	deletion	1.030	0.021	0.969 - 1.095	1.419	0.156	62,135
14	*ABCB1*	*rs1045642*	C	0.998	0.021	0.939 - 1.061	0.082	0.935	62,135
15	*DRD3*	*rs6280*	C	0.982	0.022	0.921 - 1.048	0.813	0.416	62,135
16	*DBH*	*rs1611115*	T	0.967	0.026	0.895 - 1.044	1.305	0.192	61,205

Note: Ordinal logistic regression weights for conditional lifetime symptom count on variant × adult trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

Table S9.7: Conditional lifetime symptom count on variant × TDI, alternate scale

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.013	0.011	0.980 - 1.046	1.182	0.237	62,069
2	*BDNF*	*rs6265*	A	1.011	0.014	0.969 - 1.053	0.789	0.430	62,069
3	*COMT*	*rs4680*	G	1.004	0.011	0.971 - 1.036	0.334	0.738	62,069
4	*HTR2A*	*rs6311*	A	1.000	0.011	0.967 - 1.034	0.039	0.969	62,069
5	*TPH1*	*rs1800532*	A	1.008	0.011	0.974 - 1.042	0.691	0.489	62,069
6	*DRD4*	*VNTR*	7+ repeats	1.001	0.013	0.961 - 1.040	0.059	0.953	62,069
7	*DRD2*	*rs1800497*	T	1.021	0.014	0.980 - 1.061	1.479	0.139	62,069
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.021	0.010	0.991 - 1.050	2.060	0.039	62,069
9	*APOE*	*rs429358/rs7412*	ε-4	1.002	0.016	0.956 - 1.047	0.098	0.922	62,069
10	*MTHFR*	*rs1801133*	T	1.004	0.012	0.969 - 1.039	0.349	0.727	62,069
11	*CLOCK*	*rs1801260*	C	1.018	0.013	0.981 - 1.055	1.453	0.146	62,069
12	*SLC6A3*	*VNTR*	10+ repeats	1.005	0.013	0.967 - 1.043	0.391	0.696	62,069
13	*ACE*	*in/del*	deletion	1.011	0.011	0.979 - 1.044	1.029	0.303	62,069
14	*ABCB1*	*rs1045642*	C	1.003	0.011	0.970 - 1.036	0.247	0.805	62,069
15	*DRD3*	*rs6280*	C	1.007	0.012	0.973 - 1.042	0.615	0.539	62,069
16	*DBH*	*rs1611115*	T	1.006	0.014	0.966 - 1.047	0.462	0.644	61,139

Note: Ordinal logistic regression weights for conditional lifetime symptom count on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

Table S9.8: Conditional lifetime symptom count on variant × stressor-induced depression, alternate scale

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.971	0.023	0.906 - 1.040	1.267	0.205	61,888
2	*BDNF*	*rs6265*	A	1.011	0.030	0.925 - 1.105	0.360	0.719	61,888
3	*COMT*	*rs4680*	G	1.049	0.023	0.979 - 1.124	2.052	0.040	61,888
4	*HTR2A*	*rs6311*	A	1.028	0.024	0.959 - 1.103	1.176	0.240	61,888
5	*TPH1*	*rs1800532*	A	1.009	0.024	0.940 - 1.083	0.387	0.699	61,888
6	*DRD4*	*VNTR*	7+ repeats	1.031	0.028	0.949 - 1.121	1.093	0.274	61,888
7	*DRD2*	*rs1800497*	T	1.045	0.029	0.958 - 1.140	1.497	0.134	61,888
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.990	0.021	0.931 - 1.052	0.509	0.611	61,888
9	*APOE*	*rs429358/rs7412*	ε-4	1.019	0.033	0.925 - 1.122	0.564	0.573	61,888
10	*MTHFR*	*rs1801133*	T	0.972	0.025	0.904 - 1.046	1.132	0.258	61,888
11	*CLOCK*	*rs1801260*	C	1.031	0.026	0.953 - 1.114	1.146	0.252	61,888
12	*SLC6A3*	*VNTR*	10+ repeats	1.008	0.027	0.931 - 1.092	0.307	0.759	61,888
13	*ACE*	*in/del*	deletion	1.011	0.023	0.944 - 1.084	0.487	0.626	61,888
14	*ABCB1*	*rs1045642*	C	1.019	0.023	0.951 - 1.092	0.792	0.428	61,888
15	*DRD3*	*rs6280*	C	0.955	0.025	0.888 - 1.027	1.871	0.061	61,888
16	*DBH*	*rs1611115*	T	1.028	0.029	0.943 - 1.121	0.943	0.345	60,960

Note: Ordinal logistic regression weights for conditional lifetime symptom count on variant × stressor-induced depression controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

Table S9.9: Lifetime episode count on variant × childhood trauma, alternate scale

	Gene	Polymorphism	Risk/counted allele	$\exp(\beta)$	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.001	0.008	0.977 - 1.025	0.071	0.943	112,216
2	*BDNF*	*rs6265*	A	1.006	0.010	0.976 - 1.038	0.621	0.534	112,216
3	*COMT*	*rs4680*	G	0.995	0.008	0.971 - 1.019	0.673	0.501	112,216
4	*HTR2A*	*rs6311*	A	0.990	0.008	0.966 - 1.014	1.240	0.215	112,216
5	*TPH1*	*rs1800532*	A	0.999	0.008	0.974 - 1.023	0.161	0.872	112,216
6	*DRD4*	*VNTR*	7+ repeats	1.001	0.010	0.973 - 1.030	0.096	0.924	112,216
7	*DRD2*	*rs1800497*	T	1.001	0.010	0.972 - 1.031	0.111	0.912	112,216
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.995	0.007	0.974 - 1.016	0.762	0.446	112,216
9	*APOE*	*rs429358/rs7412*	ε-4	1.008	0.011	0.975 - 1.043	0.717	0.473	112,216
10	*MTHFR*	*rs1801133*	T	0.987	0.009	0.962 - 1.012	1.551	0.121	112,216
11	*CLOCK*	*rs1801260*	C	1.000	0.009	0.974 - 1.027	0.028	0.977	112,216
12	*SLC6A3*	*VNTR*	10+ repeats	0.990	0.009	0.963 - 1.017	1.114	0.265	112,216
13	*ACE*	*in/del*	deletion	0.997	0.008	0.973 - 1.021	0.420	0.675	112,216
14	*ABCB1*	*rs1045642*	C	0.990	0.008	0.967 - 1.014	1.187	0.235	112,216
15	*DRD3*	*rs6280*	C	0.997	0.009	0.972 - 1.022	0.373	0.709	112,216
16	*DBH*	*rs1611115*	T	0.986	0.010	0.957 - 1.016	1.433	0.152	110,469

Note: Linear regression weights for lifetime episode count on variant × childhood trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

Table S9.10: Lifetime episode count on variant × adult trauma, alternate scale

	Gene	Polymorphism	Risk/counted allele	$\exp(\beta)$	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.003	0.007	0.983 - 1.023	0.469	0.639	112,255
2	*BDNF*	*rs6265*	A	0.994	0.009	0.969 - 1.020	0.683	0.495	112,255
3	*COMT*	*rs4680*	G	1.000	0.007	0.980 - 1.020	0.011	0.992	112,255
4	*HTR2A*	*rs6311*	A	1.000	0.007	0.980 - 1.021	0.003	0.997	112,255
5	*TPH1*	*rs1800532*	A	1.005	0.007	0.985 - 1.026	0.724	0.469	112,255
6	*DRD4*	*VNTR*	7+ repeats	1.002	0.008	0.978 - 1.026	0.192	0.847	112,255
7	*DRD2*	*rs1800497*	T	1.005	0.008	0.980 - 1.030	0.603	0.547	112,255
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.003	0.006	0.985 - 1.020	0.433	0.665	112,255
9	*APOE*	*rs429358/rs7412*	ε-4	0.999	0.009	0.972 - 1.028	0.080	0.936	112,255
10	*MTHFR*	*rs1801133*	T	1.014	0.007	0.992 - 1.035	1.888	0.059	112,255
11	*CLOCK*	*rs1801260*	C	1.000	0.008	0.977 - 1.022	0.050	0.960	112,255
12	*SLC6A3*	*VNTR*	10+ repeats	1.001	0.008	0.978 - 1.024	0.117	0.907	112,255
13	*ACE*	*in/del*	deletion	1.001	0.007	0.981 - 1.021	0.164	0.870	112,255
14	*ABCB1*	*rs1045642*	C	1.000	0.007	0.981 - 1.021	0.067	0.947	112,255
15	*DRD3*	*rs6280*	C	0.989	0.007	0.969 - 1.011	1.488	0.137	112,255
16	*DBH*	*rs1611115*	T	0.995	0.008	0.971 - 1.020	0.580	0.562	110,507

Note: Linear regression weights for lifetime episode count on variant × adult trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

Table S9.11: Lifetime episode count on variant × TDI, alternate scale

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.007	0.004	-0.004 - 0.019	1.934	0.053	112,144
2	*BDNF*	*rs6265*	A	0.009	0.005	-0.005 - 0.023	1.857	0.063	112,144
3	*COMT*	*rs4680*	G	0.005	0.004	-0.006 - 0.016	1.357	0.175	112,144
4	*HTR2A*	*rs6311*	A	0.003	0.004	-0.009 - 0.014	0.724	0.469	112,144
5	*TPH1*	*rs1800532*	A	0.001	0.004	-0.010 - 0.013	0.315	0.753	112,144
6	*DRD4*	*VNTR*	7+ repeats	-0.002	0.005	-0.016 - 0.011	0.542	0.588	112,144
7	*DRD2*	*rs1800497*	T	-0.004	0.005	-0.018 - 0.010	0.817	0.414	112,144
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	-0.001	0.003	-0.011 - 0.009	0.285	0.776	112,144
9	*APOE*	*rs429358/rs7412*	ε-4	-0.003	0.005	-0.018 - 0.013	0.479	0.632	112,144
10	*MTHFR*	*rs1801133*	T	-0.001	0.004	-0.013 - 0.011	0.189	0.850	112,144
11	*CLOCK*	*rs1801260*	C	0.010	0.004	-0.003 - 0.023	2.297	0.022	112,144
12	*SLC6A3*	*VNTR*	10+ repeats	0.002	0.004	-0.010 - 0.015	0.576	0.565	112,144
13	*ACE*	*in/del*	deletion	0.006	0.004	-0.005 - 0.017	1.612	0.107	112,144
14	*ABCB1*	*rs1045642*	C	-0.007	0.004	-0.018 - 0.004	1.808	0.071	112,144
15	*DRD3*	*rs6280*	C	0.001	0.004	-0.011 - 0.012	0.140	0.889	112,144
16	*DBH*	*rs1611115*	T	0.007	0.005	-0.007 - 0.021	1.418	0.156	110,397

Note: Linear regression weights for lifetime episode count on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

Table S9.12: Touchscreen probable lifetime diagnosis on variant × TDI, alternate scale

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.001	0.002	-0.005 - 0.008	0.556	0.578	90,812
2	*BDNF*	*rs6265*	A	-0.003	0.003	-0.011 - 0.006	0.967	0.334	90,812
3	*COMT*	*rs4680*	G	-0.002	0.002	-0.009 - 0.005	0.975	0.329	90,812
4	*HTR2A*	*rs6311*	A	0.000	0.002	-0.007 - 0.007	0.128	0.898	90,812
5	*TPH1*	*rs1800532*	A	-0.003	0.002	-0.010 - 0.004	1.357	0.175	90,812
6	*DRD4*	*VNTR*	7+ repeats	0.001	0.003	-0.007 - 0.009	0.286	0.775	90,812
7	*DRD2*	*rs1800497*	T	-0.003	0.003	-0.011 - 0.005	1.064	0.287	90,812
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	-0.001	0.002	-0.006 - 0.005	0.318	0.750	90,812
9	*APOE*	*rs429358/rs7412*	ε-4	0.004	0.003	-0.005 - 0.013	1.292	0.196	90,812
10	*MTHFR*	*rs1801133*	T	-0.002	0.002	-0.009 - 0.005	0.826	0.409	90,812
11	*CLOCK*	*rs1801260*	C	-0.002	0.003	-0.010 - 0.005	0.853	0.394	90,812
12	*SLC6A3*	*VNTR*	10+ repeats	0.004	0.003	-0.004 - 0.012	1.462	0.144	90,812
13	*ACE*	*in/del*	deletion	-0.000	0.002	-0.007 - 0.007	0.094	0.925	90,812
14	*ABCB1*	*rs1045642*	C	0.003	0.002	-0.004 - 0.010	1.236	0.216	90,812
15	*DRD3*	*rs6280*	C	0.003	0.002	-0.004 - 0.010	1.276	0.202	90,812
16	*DBH*	*rs1611115*	T	0.002	0.003	-0.006 - 0.010	0.740	0.460	89,397

Note: Linear regression weights for touchscreen probable lifetime diagnosis on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

Table S9.13: Touchscreen probable lifetime diagnosis, ordinal classification on variant × TDI, alternate scale

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.004	0.005	-0.011 - 0.018	0.705	0.481	90,812
2	*BDNF*	*rs6265*	A	-0.009	0.006	-0.028 - 0.010	1.449	0.147	90,812
3	*COMT*	*rs4680*	G	-0.001	0.005	-0.016 - 0.014	0.226	0.821	90,812
4	*HTR2A*	*rs6311*	A	0.001	0.005	-0.014 - 0.016	0.259	0.796	90,812
5	*TPH1*	*rs1800532*	A	-0.007	0.005	-0.022 - 0.008	1.402	0.161	90,812
6	*DRD4*	*VNTR*	7+ repeats	-0.001	0.006	-0.018 - 0.017	0.087	0.931	90,812
7	*DRD2*	*rs1800497*	T	0.001	0.006	-0.018 - 0.019	0.140	0.889	90,812
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	-0.003	0.004	-0.016 - 0.010	0.717	0.473	90,812
9	*APOE*	*rs429358/rs7412*	ε-4	0.012	0.007	-0.009 - 0.032	1.695	0.090	90,812
10	*MTHFR*	*rs1801133*	T	-0.002	0.005	-0.018 - 0.013	0.472	0.637	90,812
11	*CLOCK*	*rs1801260*	C	0.001	0.006	-0.016 - 0.018	0.152	0.879	90,812
12	*SLC6A3*	*VNTR*	10+ repeats	0.007	0.006	-0.010 - 0.023	1.151	0.250	90,812
13	*ACE*	*in/del*	deletion	0.002	0.005	-0.013 - 0.016	0.328	0.743	90,812
14	*ABCB1*	*rs1045642*	C	0.005	0.005	-0.010 - 0.020	0.986	0.324	90,812
15	*DRD3*	*rs6280*	C	0.004	0.005	-0.011 - 0.020	0.809	0.418	90,812
16	*DBH*	*rs1611115*	T	0.005	0.006	-0.013 - 0.024	0.867	0.386	89,397

Note: Linear regression weights for touchscreen probable lifetime diagnosis, ordinal classification on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

Table S9.14: Severe recurrent MDD (MHF) on variant × childhood trauma, alternate scale

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.001	0.002	-0.005 - 0.008	0.626	0.532	67,304
2	*BDNF*	*rs6265*	A	0.003	0.003	-0.006 - 0.012	1.002	0.316	67,304
3	*COMT*	*rs4680*	G	0.002	0.002	-0.005 - 0.008	0.695	0.487	67,304
4	*HTR2A*	*rs6311*	A	0.004	0.002	-0.003 - 0.011	1.610	0.107	67,304
5	*TPH1*	*rs1800532*	A	-0.001	0.002	-0.008 - 0.005	0.596	0.551	67,304
6	*DRD4*	*VNTR*	7+ repeats	-0.000	0.003	-0.008 - 0.008	0.001	0.999	67,304
7	*DRD2*	*rs1800497*	T	-0.001	0.003	-0.010 - 0.007	0.466	0.642	67,304
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	-0.001	0.002	-0.007 - 0.005	0.351	0.726	67,304
9	*APOE*	*rs429358/rs7412*	ε-4	0.001	0.003	-0.009 - 0.010	0.172	0.863	67,304
10	*MTHFR*	*rs1801133*	T	0.005	0.002	-0.003 - 0.012	1.880	0.060	67,304
11	*CLOCK*	*rs1801260*	C	-0.001	0.003	-0.008 - 0.007	0.269	0.788	67,304
12	*SLC6A3*	*VNTR*	10+ repeats	-0.004	0.003	-0.012 - 0.004	1.529	0.126	67,304
13	*ACE*	*in/del*	deletion	0.002	0.002	-0.005 - 0.009	0.882	0.378	67,304
14	*ABCB1*	*rs1045642*	C	0.001	0.002	-0.006 - 0.008	0.522	0.602	67,304
15	*DRD3*	*rs6280*	C	-0.002	0.002	-0.009 - 0.006	0.631	0.528	67,304
16	*DBH*	*rs1611115*	T	-0.005	0.003	-0.013 - 0.004	1.712	0.087	66,231

Note: Logistic regression weights for Severe recurrent MDD (MHF) on variant × childhood trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

Table S9.15: Severe recurrent MDD (MHF) on variant × childhood trauma, alternate scale

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.001	0.002	-0.005 - 0.008	0.626	0.532	67,304
2	*BDNF*	*rs6265*	A	0.003	0.003	-0.006 - 0.012	1.002	0.316	67,304
3	*COMT*	*rs4680*	G	0.002	0.002	-0.005 - 0.008	0.695	0.487	67,304
4	*HTR2A*	*rs6311*	A	0.004	0.002	-0.003 - 0.011	1.610	0.107	67,304
5	*TPH1*	*rs1800532*	A	-0.001	0.002	-0.008 - 0.005	0.596	0.551	67,304
6	*DRD4*	*VNTR*	7+ repeats	-0.000	0.003	-0.008 - 0.008	0.001	0.999	67,304
7	*DRD2*	*rs1800497*	T	-0.001	0.003	-0.010 - 0.007	0.466	0.642	67,304
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	-0.001	0.002	-0.007 - 0.005	0.351	0.726	67,304
9	*APOE*	*rs429358/rs7412*	ε-4	0.001	0.003	-0.009 - 0.010	0.172	0.863	67,304
10	*MTHFR*	*rs1801133*	T	0.005	0.002	-0.003 - 0.012	1.880	0.060	67,304
11	*CLOCK*	*rs1801260*	C	-0.001	0.003	-0.008 - 0.007	0.269	0.788	67,304
12	*SLC6A3*	*VNTR*	10+ repeats	-0.004	0.003	-0.012 - 0.004	1.529	0.126	67,304
13	*ACE*	*in/del*	deletion	0.002	0.002	-0.005 - 0.009	0.882	0.378	67,304
14	*ABCB1*	*rs1045642*	C	0.001	0.002	-0.006 - 0.008	0.522	0.602	67,304
15	*DRD3*	*rs6280*	C	-0.002	0.002	-0.009 - 0.006	0.631	0.528	67,304
16	*DBH*	*rs1611115*	T	-0.005	0.003	-0.013 - 0.004	1.712	0.087	66,231

Note: Logistic regression weights for Severe recurrent MDD (MHF) on variant × childhood trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

Table S9.16: Severe recurrent MDD (MHF) on variant × TDI, alternate scale

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.002	0.001	0.999 - 1.004	1.880	0.060	67,237
2	*BDNF*	*rs6265*	A	1.002	0.001	0.999 - 1.005	1.621	0.105	67,237
3	*COMT*	*rs4680*	G	1.001	0.001	0.998 - 1.003	0.799	0.424	67,237
4	*HTR2A*	*rs6311*	A	1.000	0.001	0.998 - 1.003	0.300	0.764	67,237
5	*TPH1*	*rs1800532*	A	1.000	0.001	0.998 - 1.003	0.113	0.910	67,237
6	*DRD4*	*VNTR*	7+ repeats	1.001	0.001	0.998 - 1.004	1.240	0.215	67,237
7	*DRD2*	*rs1800497*	T	1.000	0.001	0.997 - 1.003	0.340	0.734	67,237
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.000	0.001	0.998 - 1.002	0.427	0.669	67,237
9	*APOE*	*rs429358/rs7412*	ε-4	1.001	0.001	0.997 - 1.004	0.597	0.551	67,237
10	*MTHFR*	*rs1801133*	T	1.000	0.001	0.997 - 1.003	0.042	0.967	67,237
11	*CLOCK*	*rs1801260*	C	1.002	0.001	1.000 - 1.005	2.461	0.014	67,237
12	*SLC6A3*	*VNTR*	10+ repeats	1.001	0.001	0.998 - 1.004	1.266	0.205	67,237
13	*ACE*	*in/del*	deletion	1.001	0.001	0.999 - 1.004	1.369	0.171	67,237
14	*ABCB1*	*rs1045642*	C	1.000	0.001	0.998 - 1.003	0.390	0.697	67,237
15	*DRD3*	*rs6280*	C	1.000	0.001	0.998 - 1.003	0.428	0.668	67,237
16	*DBH*	*rs1611115*	T	1.001	0.001	0.998 - 1.004	1.384	0.166	66,165

Note: Linear regression weights for Severe recurrent MDD (MHF) on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

S10 Polymorphism level $G \times E$ effects (improper control)

Table S10.1: Estimated lifetime MDD diagnosis on variant × childhood trauma, improper control

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	-0.001	0.022	-0.066 - 0.064	0.052	0.959	115,204
2	*BDNF*	*rs6265*	A	0.006	0.029	-0.079 - 0.090	0.201	0.841	115,204
3	*COMT*	*rs4680*	G	-0.012	0.022	-0.078 - 0.053	0.549	0.583	115,204
4	*HTR2A*	*rs6311*	A	-0.031	0.023	-0.098 - 0.036	1.385	0.166	115,204
5	*TPH1*	*rs1800532*	A	0.022	0.023	-0.045 - 0.089	0.968	0.333	115,204
6	*DRD4*	*VNTR*	7+ repeats	-0.006	0.027	-0.084 - 0.073	0.206	0.837	115,204
7	*DRD2*	*rs1800497*	T	-0.024	0.028	-0.106 - 0.057	0.889	0.374	115,204
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	-0.016	0.020	-0.075 - 0.043	0.803	0.422	115,204
9	*APOE*	*rs429358/rs7412*	ε-4	0.038	0.031	-0.054 - 0.130	1.217	0.224	115,204
10	*MTHFR*	*rs1801133*	T	-0.010	0.023	-0.079 - 0.060	0.408	0.684	115,204
11	*CLOCK*	*rs1801260*	C	-0.007	0.025	-0.081 - 0.067	0.287	0.774	115,204
12	*SLC6A3*	*VNTR*	10+ repeats	-0.047	0.025	-0.122 - 0.028	1.842	0.065	115,204
13	*ACE*	*in/del*	deletion	0.005	0.022	-0.061 - 0.070	0.211	0.833	115,204
14	*ABCB1*	*rs1045642*	C	-0.025	0.022	-0.091 - 0.041	1.130	0.258	115,204
15	*DRD3*	*rs6280*	C	-0.038	0.023	-0.107 - 0.032	1.599	0.110	115,204
16	*DBH*	*rs1611115*	T	-0.034	0.028	-0.116 - 0.049	1.206	0.228	113,422

Note: Logistic regression weights for estimated lifetime MDD diagnosis on variant × childhood trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

Table S10.2: Estimated lifetime MDD diagnosis on variant × adult trauma, improper control

	Gene	Polymorphism	Risk/counted allele	$\exp(\beta)$	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.019	0.020	0.961 - 1.081	0.938	0.348	115,249
2	*BDNF*	*rs6265*	A	0.995	0.026	0.923 - 1.074	0.183	0.854	115,249
3	*COMT*	*rs4680*	G	0.993	0.020	0.936 - 1.054	0.357	0.721	115,249
4	*HTR2A*	*rs6311*	A	1.011	0.020	0.951 - 1.074	0.518	0.604	115,249
5	*TPH1*	*rs1800532*	A	1.024	0.021	0.963 - 1.088	1.145	0.252	115,249
6	*DRD4*	*VNTR*	7+ repeats	0.977	0.024	0.909 - 1.049	0.967	0.333	115,249
7	*DRD2*	*rs1800497*	T	1.017	0.025	0.945 - 1.095	0.679	0.497	115,249
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.988	0.018	0.936 - 1.043	0.664	0.506	115,249
9	*APOE*	*rs429358/rs7412*	ε-4	0.993	0.028	0.913 - 1.079	0.266	0.790	115,249
10	*MTHFR*	*rs1801133*	T	1.040	0.021	0.977 - 1.108	1.854	0.064	115,249
11	*CLOCK*	*rs1801260*	C	1.000	0.023	0.935 - 1.070	0.012	0.991	115,249
12	*SLC6A3*	*VNTR*	10+ repeats	0.979	0.023	0.914 - 1.048	0.923	0.356	115,249
13	*ACE*	*in/del*	deletion	1.000	0.020	0.942 - 1.061	0.008	0.994	115,249
14	*ABCB1*	*rs1045642*	C	0.989	0.020	0.932 - 1.050	0.540	0.589	115,249
15	*DRD3*	*rs6280*	C	0.985	0.021	0.925 - 1.049	0.704	0.482	115,249
16	*DBH*	*rs1611115*	T	0.969	0.025	0.900 - 1.044	1.236	0.216	113,466

Note: Logistic regression weights for estimated lifetime MDD diagnosis on variant × adult trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

Table S10.3: Estimated lifetime MDD diagnosis on variant × TDI, improper control

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.011	0.010	0.981 - 1.042	1.088	0.277	115,138
2	*BDNF*	*rs6265*	A	1.040	0.013	1.001 - 1.079	2.955	0.003	115,138
3	*COMT*	*rs4680*	G	1.012	0.010	0.981 - 1.042	1.117	0.264	115,138
4	*HTR2A*	*rs6311*	A	1.008	0.010	0.977 - 1.039	0.754	0.451	115,138
5	*TPH1*	*rs1800532*	A	1.009	0.011	0.977 - 1.040	0.810	0.418	115,138
6	*DRD4*	*VNTR*	7+ repeats	1.004	0.012	0.967 - 1.040	0.297	0.766	115,138
7	*DRD2*	*rs1800497*	T	1.006	0.013	0.968 - 1.044	0.458	0.647	115,138
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.009	0.009	0.982 - 1.036	0.969	0.332	115,138
9	*APOE*	*rs429358/rs7412*	ε-4	1.001	0.014	0.958 - 1.043	0.059	0.953	115,138
10	*MTHFR*	*rs1801133*	T	1.003	0.011	0.971 - 1.036	0.304	0.761	115,138
11	*CLOCK*	*rs1801260*	C	1.022	0.012	0.988 - 1.057	1.905	0.057	115,138
12	*SLC6A3*	*VNTR*	10+ repeats	1.006	0.012	0.971 - 1.041	0.482	0.630	115,138
13	*ACE*	*in/del*	deletion	1.006	0.010	0.975 - 1.036	0.557	0.577	115,138
14	*ABCB1*	*rs1045642*	C	1.002	0.010	0.971 - 1.032	0.158	0.875	115,138
15	*DRD3*	*rs6280*	C	1.004	0.011	0.972 - 1.036	0.346	0.730	115,138
16	*DBH*	*rs1611115*	T	1.007	0.013	0.970 - 1.045	0.564	0.572	113,356

Note: Logistic regression weights for estimated lifetime MDD diagnosis on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch and variant/moderator main effects.

Table S10.4: Current MDD severity (MHF) on variant × recent trauma, improper control

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	-0.014	0.019	-0.069 - 0.041	0.741	0.458	115,246
2	*BDNF*	*rs6265*	A	0.011	0.024	-0.061 - 0.082	0.440	0.660	115,246
3	*COMT*	*rs4680*	G	-0.018	0.019	-0.074 - 0.037	0.963	0.336	115,246
4	*HTR2A*	*rs6311*	A	0.003	0.019	-0.053 - 0.059	0.168	0.867	115,246
5	*TPH1*	*rs1800532*	A	0.007	0.019	-0.050 - 0.064	0.355	0.722	115,246
6	*DRD4*	*VNTR*	7+ repeats	-0.002	0.022	-0.068 - 0.064	0.096	0.924	115,246
7	*DRD2*	*rs1800497*	T	-0.021	0.023	-0.089 - 0.048	0.895	0.371	115,246
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.011	0.016	-0.037 - 0.060	0.699	0.484	115,246
9	*APOE*	*rs429358/rs7412*	ε-4	-0.051	0.027	-0.129 - 0.028	1.906	0.057	115,246
10	*MTHFR*	*rs1801133*	T	0.003	0.020	-0.055 - 0.062	0.175	0.861	115,246
11	*CLOCK*	*rs1801260*	C	0.008	0.021	-0.054 - 0.070	0.383	0.702	115,246
12	*SLC6A3*	*VNTR*	10+ repeats	-0.005	0.021	-0.068 - 0.058	0.233	0.816	115,246
13	*ACE*	*in/del*	deletion	0.001	0.019	-0.055 - 0.056	0.049	0.961	115,246
14	*ABCB1*	*rs1045642*	C	0.035	0.019	-0.021 - 0.090	1.863	0.062	115,246
15	*DRD3*	*rs6280*	C	-0.032	0.020	-0.090 - 0.026	1.646	0.100	115,246
16	*DBH*	*rs1611115*	T	-0.015	0.023	-0.084 - 0.054	0.651	0.515	113,463

Note: Negative binomial regression weights for current MDD severity (MHF) on variant × recent trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

Table S10.5: Conditional lifetime symptom count on variant × childhood trauma, improper control

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.990	0.022	0.928 - 1.056	0.449	0.653	62,108
2	*BDNF*	*rs6265*	A	1.002	0.028	0.921 - 1.088	0.054	0.957	62,108
3	*COMT*	*rs4680*	G	1.008	0.022	0.945 - 1.076	0.374	0.708	62,108
4	*HTR2A*	*rs6311*	A	1.007	0.022	0.943 - 1.076	0.325	0.745	62,108
5	*TPH1*	*rs1800532*	A	1.045	0.023	0.977 - 1.117	1.949	0.051	62,108
6	*DRD4*	*VNTR*	7+ repeats	1.011	0.027	0.935 - 1.094	0.417	0.677	62,108
7	*DRD2*	*rs1800497*	T	1.002	0.027	0.924 - 1.086	0.068	0.946	62,108
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.982	0.020	0.925 - 1.041	0.934	0.350	62,108
9	*APOE*	*rs429358/rs7412*	ε-4	1.004	0.031	0.916 - 1.099	0.121	0.904	62,108
10	*MTHFR*	*rs1801133*	T	0.983	0.023	0.917 - 1.053	0.747	0.455	62,108
11	*CLOCK*	*rs1801260*	C	0.992	0.025	0.922 - 1.067	0.323	0.746	62,108
12	*SLC6A3*	*VNTR*	10+ repeats	0.970	0.025	0.901 - 1.045	1.210	0.226	62,108
13	*ACE*	*in/del*	deletion	0.993	0.022	0.930 - 1.059	0.330	0.742	62,108
14	*ABCB1*	*rs1045642*	C	1.030	0.022	0.965 - 1.099	1.347	0.178	62,108
15	*DRD3*	*rs6280*	C	0.953	0.023	0.890 - 1.021	2.052	0.040	62,108
16	*DBH*	*rs1611115*	T	1.001	0.028	0.923 - 1.086	0.041	0.967	61,178

Note: Linear regression weights for conditional lifetime symptom count on variant × childhood trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

Table S10.6: Conditional lifetime symptom count on variant × adult trauma, improper control

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.006	0.020	0.949 - 1.065	0.282	0.778	62,135
2	*BDNF*	*rs6265*	A	1.028	0.025	0.954 - 1.107	1.099	0.272	62,135
3	*COMT*	*rs4680*	G	0.980	0.020	0.925 - 1.038	1.048	0.294	62,135
4	*HTR2A*	*rs6311*	A	0.996	0.020	0.939 - 1.057	0.187	0.852	62,135
5	*TPH1*	*rs1800532*	A	1.010	0.020	0.951 - 1.072	0.473	0.636	62,135
6	*DRD4*	*VNTR*	7+ repeats	0.978	0.024	0.912 - 1.049	0.940	0.347	62,135
7	*DRD2*	*rs1800497*	T	1.008	0.024	0.937 - 1.083	0.306	0.760	62,135
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.978	0.018	0.928 - 1.031	1.235	0.217	62,135
9	*APOE*	*rs429358/rs7412*	ε-4	0.988	0.028	0.911 - 1.072	0.426	0.670	62,135
10	*MTHFR*	*rs1801133*	T	1.009	0.021	0.949 - 1.073	0.443	0.658	62,135
11	*CLOCK*	*rs1801260*	C	1.039	0.022	0.973 - 1.109	1.708	0.088	62,135
12	*SLC6A3*	*VNTR*	10+ repeats	0.974	0.023	0.911 - 1.041	1.190	0.234	62,135
13	*ACE*	*in/del*	deletion	1.025	0.020	0.967 - 1.087	1.256	0.209	62,135
14	*ABCB1*	*rs1045642*	C	0.997	0.020	0.941 - 1.056	0.165	0.869	62,135
15	*DRD3*	*rs6280*	C	0.977	0.021	0.919 - 1.039	1.112	0.266	62,135
16	*DBH*	*rs1611115*	T	0.969	0.025	0.901 - 1.042	1.275	0.202	61,205

Note: Linear regression weights for conditional lifetime symptom count on variant × adult trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

Table S10.7: Conditional lifetime symptom count on variant × TDI, improper control

	Gene	Polymorphism	Risk/counted allele	β	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.003	0.010	-0.028 - 0.033	0.250	0.802	62,069
2	*BDNF*	*rs6265*	A	0.013	0.013	-0.026 - 0.051	0.995	0.320	62,069
3	*COMT*	*rs4680*	G	0.000	0.010	-0.030 - 0.030	0.039	0.969	62,069
4	*HTR2A*	*rs6311*	A	0.005	0.010	-0.026 - 0.035	0.463	0.643	62,069
5	*TPH1*	*rs1800532*	A	-0.006	0.010	-0.037 - 0.025	0.602	0.547	62,069
6	*DRD4*	*VNTR*	7+ repeats	0.007	0.012	-0.029 - 0.043	0.563	0.573	62,069
7	*DRD2*	*rs1800497*	T	-0.014	0.013	-0.051 - 0.024	1.087	0.277	62,069
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	-0.025	0.009	-0.052 - 0.003	2.665	0.008	62,069
9	*APOE*	*rs429358/rs7412*	ε-4	-0.001	0.014	-0.043 - 0.041	0.071	0.944	62,069
10	*MTHFR*	*rs1801133*	T	0.006	0.011	-0.026 - 0.037	0.510	0.610	62,069
11	*CLOCK*	*rs1801260*	C	0.008	0.011	-0.026 - 0.042	0.703	0.482	62,069
12	*SLC6A3*	*VNTR*	10+ repeats	0.003	0.012	-0.031 - 0.037	0.259	0.796	62,069
13	*ACE*	*in/del*	deletion	-0.011	0.010	-0.041 - 0.019	1.112	0.266	62,069
14	*ABCB1*	*rs1045642*	C	0.001	0.010	-0.029 - 0.031	0.093	0.926	62,069
15	*DRD3*	*rs6280*	C	-0.007	0.011	-0.039 - 0.025	0.638	0.523	62,069
16	*DBH*	*rs1611115*	T	-0.000	0.013	-0.037 - 0.037	0.024	0.981	61,139

Note: Linear regression weights for conditional lifetime symptom count on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch and variant/moderator main effects.

Table S10.8: Conditional lifetime symptom count on variant × stressor-induced depression, improper control

	Gene	Polymorphism	Risk/counted allele	$\exp(\beta)$	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	0.966	0.022	0.906 - 1.031	1.575	0.115	61,888
2	*BDNF*	*rs6265*	A	1.002	0.028	0.922 - 1.089	0.077	0.939	61,888
3	*COMT*	*rs4680*	G	1.047	0.022	0.981 - 1.117	2.091	0.036	61,888
4	*HTR2A*	*rs6311*	A	1.026	0.022	0.961 - 1.096	1.173	0.241	61,888
5	*TPH1*	*rs1800532*	A	1.013	0.023	0.948 - 1.083	0.584	0.559	61,888
6	*DRD4*	*VNTR*	7+ repeats	1.039	0.026	0.961 - 1.124	1.450	0.147	61,888
7	*DRD2*	*rs1800497*	T	1.051	0.027	0.970 - 1.140	1.823	0.068	61,888
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.988	0.019	0.933 - 1.046	0.640	0.522	61,888
9	*APOE*	*rs429358/rs7412*	ε-4	1.011	0.031	0.923 - 1.107	0.356	0.722	61,888
10	*MTHFR*	*rs1801133*	T	0.979	0.023	0.914 - 1.048	0.920	0.358	61,888
11	*CLOCK*	*rs1801260*	C	1.019	0.025	0.948 - 1.097	0.777	0.437	61,888
12	*SLC6A3*	*VNTR*	10+ repeats	1.011	0.025	0.939 - 1.089	0.447	0.655	61,888
13	*ACE*	*in/del*	deletion	1.014	0.022	0.951 - 1.082	0.648	0.517	61,888
14	*ABCB1*	*rs1045642*	C	1.018	0.022	0.954 - 1.086	0.813	0.416	61,888
15	*DRD3*	*rs6280*	C	0.958	0.023	0.895 - 1.026	1.853	0.064	61,888
16	*DBH*	*rs1611115*	T	1.027	0.027	0.947 - 1.113	0.969	0.332	60,960

Note: Linear regression weights for conditional lifetime symptom count on variant × stressor-induced depression controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the additive scale.

Table S10.9: Lifetime episode count on variant × childhood trauma, improper control

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.002	0.019	0.947 - 1.060	0.108	0.914	112,216
2	*BDNF*	*rs6265*	A	1.018	0.025	0.946 - 1.095	0.705	0.481	112,216
3	*COMT*	*rs4680*	G	0.986	0.019	0.932 - 1.044	0.723	0.470	112,216
4	*HTR2A*	*rs6311*	A	0.977	0.020	0.922 - 1.035	1.202	0.229	112,216
5	*TPH1*	*rs1800532*	A	0.994	0.020	0.938 - 1.054	0.294	0.769	112,216
6	*DRD4*	*VNTR*	7+ repeats	1.003	0.023	0.937 - 1.074	0.129	0.897	112,216
7	*DRD2*	*rs1800497*	T	1.003	0.024	0.934 - 1.076	0.106	0.916	112,216
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	0.983	0.017	0.935 - 1.033	1.024	0.306	112,216
9	*APOE*	*rs429358/rs7412*	ε-4	1.025	0.027	0.947 - 1.110	0.928	0.354	112,216
10	*MTHFR*	*rs1801133*	T	0.969	0.020	0.913 - 1.029	1.545	0.122	112,216
11	*CLOCK*	*rs1801260*	C	1.004	0.022	0.942 - 1.071	0.208	0.836	112,216
12	*SLC6A3*	*VNTR*	10+ repeats	0.976	0.022	0.914 - 1.041	1.121	0.262	112,216
13	*ACE*	*in/del*	deletion	0.990	0.019	0.936 - 1.048	0.499	0.618	112,216
14	*ABCB1*	*rs1045642*	C	0.976	0.019	0.922 - 1.033	1.275	0.202	112,216
15	*DRD3*	*rs6280*	C	0.995	0.020	0.937 - 1.057	0.232	0.817	112,216
16	*DBH*	*rs1611115*	T	0.969	0.024	0.903 - 1.040	1.313	0.189	110,469

Note: Ordinal logistic regression weights for lifetime episode count on variant × childhood trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

Table S10.10: Lifetime episode count on variant × adult trauma, improper control

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.011	0.016	0.964 - 1.061	0.687	0.492	112,255
2	*BDNF*	*rs6265*	A	0.986	0.021	0.927 - 1.049	0.676	0.499	112,255
3	*COMT*	*rs4680*	G	0.999	0.016	0.951 - 1.048	0.084	0.933	112,255
4	*HTR2A*	*rs6311*	A	0.998	0.017	0.950 - 1.049	0.117	0.907	112,255
5	*TPH1*	*rs1800532*	A	1.010	0.017	0.961 - 1.062	0.598	0.550	112,255
6	*DRD4*	*VNTR*	7+ repeats	1.006	0.020	0.949 - 1.066	0.294	0.769	112,255
7	*DRD2*	*rs1800497*	T	1.009	0.020	0.950 - 1.072	0.458	0.647	112,255
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.008	0.015	0.965 - 1.053	0.549	0.583	112,255
9	*APOE*	*rs429358/rs7412*	ε-4	1.005	0.023	0.939 - 1.075	0.212	0.832	112,255
10	*MTHFR*	*rs1801133*	T	1.031	0.017	0.980 - 1.086	1.768	0.077	112,255
11	*CLOCK*	*rs1801260*	C	0.999	0.019	0.946 - 1.055	0.047	0.962	112,255
12	*SLC6A3*	*VNTR*	10+ repeats	1.000	0.019	0.946 - 1.057	0.009	0.993	112,255
13	*ACE*	*in/del*	deletion	0.997	0.016	0.950 - 1.047	0.155	0.877	112,255
14	*ABCB1*	*rs1045642*	C	1.003	0.016	0.956 - 1.053	0.190	0.849	112,255
15	*DRD3*	*rs6280*	C	0.979	0.017	0.930 - 1.030	1.247	0.213	112,255
16	*DBH*	*rs1611115*	T	0.986	0.020	0.929 - 1.048	0.666	0.506	110,507

Note: Ordinal logistic regression weights for lifetime episode count on variant × adult trauma controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch, variant/moderator main effects, and all variant × covariate and moderator × covariate effects, on the multiplicative scale.

Table S10.11: Lifetime episode count on variant × TDI, improper control

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.019	0.009	0.993 - 1.045	2.168	0.030	112,144
2	*BDNF*	*rs6265*	A	1.022	0.011	0.989 - 1.055	1.946	0.052	112,144
3	*COMT*	*rs4680*	G	1.009	0.009	0.984 - 1.035	1.074	0.283	112,144
4	*HTR2A*	*rs6311*	A	1.005	0.009	0.978 - 1.031	0.522	0.601	112,144
5	*TPH1*	*rs1800532*	A	1.003	0.009	0.976 - 1.029	0.315	0.752	112,144
6	*DRD4*	*VNTR*	7+ repeats	1.006	0.011	0.975 - 1.038	0.602	0.547	112,144
7	*DRD2*	*rs1800497*	T	1.010	0.011	0.977 - 1.042	0.896	0.370	112,144
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.001	0.008	0.978 - 1.024	0.177	0.859	112,144
9	*APOE*	*rs429358/rs7412*	ε-4	1.007	0.012	0.971 - 1.043	0.556	0.578	112,144
10	*MTHFR*	*rs1801133*	T	1.004	0.009	0.977 - 1.032	0.445	0.656	112,144
11	*CLOCK*	*rs1801260*	C	1.019	0.010	0.990 - 1.048	1.906	0.057	112,144
12	*SLC6A3*	*VNTR*	10+ repeats	1.004	0.010	0.974 - 1.033	0.355	0.723	112,144
13	*ACE*	*in/del*	deletion	1.017	0.009	0.991 - 1.043	1.929	0.054	112,144
14	*ABCB1*	*rs1045642*	C	1.016	0.009	0.990 - 1.042	1.754	0.080	112,144
15	*DRD3*	*rs6280*	C	1.001	0.009	0.973 - 1.028	0.075	0.940	112,144
16	*DBH*	*rs1611115*	T	1.016	0.011	0.984 - 1.048	1.474	0.140	110,397

Note: Ordinal logistic regression weights for lifetime episode count on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch and variant/moderator main effects.

Table S10.12: Touchscreen probable lifetime diagnosis on variant × TDI, improper control

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.005	0.011	0.972 - 1.038	0.461	0.644	90,812
2	*BDNF*	*rs6265*	A	1.015	0.014	0.973 - 1.058	1.064	0.287	90,812
3	*COMT*	*rs4680*	G	1.008	0.011	0.974 - 1.041	0.686	0.493	90,812
4	*HTR2A*	*rs6311*	A	1.000	0.012	0.966 - 1.034	0.021	0.983	90,812
5	*TPH1*	*rs1800532*	A	1.022	0.012	0.988 - 1.056	1.904	0.057	90,812
6	*DRD4*	*VNTR*	7+ repeats	1.003	0.014	0.963 - 1.043	0.218	0.828	90,812
7	*DRD2*	*rs1800497*	T	1.014	0.014	0.972 - 1.056	0.978	0.328	90,812
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.002	0.010	0.972 - 1.031	0.162	0.872	90,812
9	*APOE*	*rs429358/rs7412*	ε-4	1.023	0.016	0.977 - 1.069	1.469	0.142	90,812
10	*MTHFR*	*rs1801133*	T	1.007	0.012	0.972 - 1.043	0.611	0.541	90,812
11	*CLOCK*	*rs1801260*	C	1.013	0.013	0.975 - 1.050	0.973	0.330	90,812
12	*SLC6A3*	*VNTR*	10+ repeats	1.024	0.013	0.985 - 1.062	1.808	0.071	90,812
13	*ACE*	*in/del*	deletion	1.010	0.011	0.976 - 1.043	0.834	0.404	90,812
14	*ABCB1*	*rs1045642*	C	1.011	0.011	0.977 - 1.044	0.922	0.356	90,812
15	*DRD3*	*rs6280*	C	1.017	0.012	0.982 - 1.052	1.430	0.153	90,812
16	*DBH*	*rs1611115*	T	1.015	0.014	0.974 - 1.057	1.094	0.274	89,397

Note: Logistic regression weights for touchscreen probable lifetime diagnosis on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch and variant/moderator main effects.

Table S10.13: Touchscreen probable lifetime diagnosis, ordinal classification on variant × TDI, improper control

	Gene	Polymorphism	Risk/counted allele	exp(β)	se_β	Corrected CI	$\lvert z \rvert$	p	n
1	*SLC6A4*	*5-HTTLPR*	Low activity	1.007	0.011	0.974 - 1.039	0.594	0.553	90,812
2	*BDNF*	*rs6265*	A	1.020	0.014	0.979 - 1.062	1.426	0.154	90,812
3	*COMT*	*rs4680*	G	1.004	0.011	0.971 - 1.037	0.338	0.736	90,812
4	*HTR2A*	*rs6311*	A	1.003	0.011	0.970 - 1.036	0.262	0.794	90,812
5	*TPH1*	*rs1800532*	A	1.022	0.011	0.988 - 1.055	1.888	0.059	90,812
6	*DRD4*	*VNTR*	7+ repeats	1.004	0.013	0.964 - 1.043	0.263	0.793	90,812
7	*DRD2*	*rs1800497*	T	1.004	0.014	0.963 - 1.045	0.300	0.764	90,812
8	*MAOA*	*VNTR*	2, 3, or 5 repeats	1.002	0.010	0.973 - 1.031	0.223	0.823	90,812
9	*APOE*	*rs429358/rs7412*	ε-4	1.027	0.015	0.981 - 1.072	1.717	0.086	90,812
10	*MTHFR*	*rs1801133*	T	1.005	0.012	0.970 - 1.039	0.403	0.687	90,812
11	*CLOCK*	*rs1801260*	C	1.007	0.013	0.970 - 1.045	0.587	0.557	90,812
12	*SLC6A3*	*VNTR*	10+ repeats	1.022	0.013	0.985 - 1.059	1.726	0.084	90,812
13	*ACE*	*in/del*	deletion	1.008	0.011	0.975 - 1.041	0.704	0.482	90,812
14	*ABCB1*	*rs1045642*	C	1.008	0.011	0.975 - 1.041	0.679	0.497	90,812
15	*DRD3*	*rs6280*	C	1.015	0.012	0.980 - 1.049	1.266	0.206	90,812
16	*DBH*	*rs1611115*	T	1.016	0.014	0.975 - 1.056	1.132	0.258	89,397

Note: Ordinal logistic regression weights for touchscreen probable lifetime diagnosis, ordinal classification on variant × TDI controling for sex, age, age^2, 10 European ancestry principle components, testing center, batch and variant/moderator main effects.

S11 Gene-level results

S11.1 Gene-wise models

Table S11.1: Gene-wise p-values (primary analyses)

	SLC6A4	*BDNF*	*COMT*	*HTR2A*	*TPH1*	*TPH2*
Estimated lifetime MDD diagnosis	.316	.132	.442	.311	.509	.868
Current MDD severity	.539	.066	.027	.593	.318	.541
Severe recurrent depression	.282	.360	.459	.362	.394	.969
Conditional lifetime symptom count	.365	.692	.288	.594	.867	.952
Lifetime episode count	.323	.593	.987	.839	.826	.775
Touchscreen probable diagnosis	.164	.614	.300	.678	.904	.107
Touchscreen probable diagnosis, ordinal classification	.253	.498	.203	.478	.995	.037
PGC lifetime MDD diagnosis	.100	.600	.700	.800	.100	.700
	MAOA	*DRD2*	*DRD4*	*MTHFR*	*APOE*	*CLOCK*
Estimated lifetime MDD diagnosis	.297	.180	.509	.351	.509	.531
Current MDD severity	.374	.969	.647	.014	.081	.079
Severe recurrent depression	.160	.460	.558	.249	.328	.911
Conditional lifetime symptom count	.134	.154	.284	.805	.231	.569
Lifetime episode count	.573	.428	.204	.713	.254	.918
Touchscreen probable diagnosis	.293	.112	.133	.727	.500	.143
Touchscreen probable diagnosis, ordinal classification	.223	.252	.161	.624	.682	.077
PGC lifetime MDD diagnosis	.268	5.142e-07*	.791	.524	.781	.015
	SLC6A3	*ACE*	*DTNBP1*	*DRD3*	*ABCB1*	*DBH*
Estimated lifetime MDD diagnosis	.264	.295	.114	.360	.239	.093
Current MDD severity	.051	.529	.134	.314	.114	.519
Severe recurrent depression	.152	.710	.133	.892	.633	.165
Conditional lifetime symptom count	.323	.789	.865	.373	.246	.367
Lifetime episode count	.201	.934	.151	.885	.425	.200
Touchscreen probable diagnosis	.078	.014	.956	.245	.181	.114
Touchscreen probable diagnosis, ordinal classification	.068	.057	.973	.058	.316	.045
PGC lifetime MDD diagnosis	.758	.221	.690	.637	.888	.765

Table S11.2: Gene-wise p-values (secondary analyses)

	SLC6A4	*BDNF*	*COMT*	*HTR2A*	*TPH1*	*TPH2*
Estimated lifetime MDD diagnosis	.535	.722	.233	.215	.584	.091
Current MDD severity	.379	.548	.144	.475	.801	.179
Severe recurrent depression	.174	.392	.488	.462	.505	.845
Conditional lifetime symptom count	.573	.824	.392	.493	.788	.995
Lifetime episode count	.023	.638	.927	.837	.874	.921
Touchscreen probable diagnosis	.678	.019	.012	.842	.210	.250
Touchscreen probable diagnosis, ordinal classification	.262	.857	.277	.852	.844	.639
PGC lifetime MDD diagnosis	.275	.886	.708	.837	.281	.358
	MAOA	*DRD2*	*DRD4*	*MTHFR*	*APOE*	*CLOCK*
Estimated lifetime MDD diagnosis	.463	.006	.654	.563	.306	.100
Current MDD severity	.252	.043	.405	.624	.536	.336
Severe recurrent depression	.529	.380	.287	.388	.572	.658
Conditional lifetime symptom count	.599	.226	.098	.425	.370	.684
Lifetime episode count	.598	.155	.043	.851	.623	.567
Touchscreen probable diagnosis	.084	.908	.375	.006	.585	.227
Touchscreen probable diagnosis, ordinal classification	.089	.028	.023	.879	.465	.629
PGC lifetime MDD diagnosis	.714	.003	.504	.587	.267	.131
	SLC6A3	*ACE*	*DTNBP1*	*DRD3*	*ABCB1*	*DBH*
Estimated lifetime MDD diagnosis	.435	.082	.955	.072	.099	.043
Current MDD severity	.379	.013	.835	.201	.077	.133
Severe recurrent depression	.047	.529	.115	.852	.735	.004
Conditional lifetime symptom count	.297	.877	.699	.648	.111	.092
Lifetime episode count	.690	.804	.045	.927	.700	.091
Touchscreen probable diagnosis	.083	.830	.628	.176	.552	.341
Touchscreen probable diagnosis, ordinal classification	.455	.384	.607	.713	.436	.188
PGC lifetime MDD diagnosis	.925	.092	.580	.307	.427	.880

*Genome-wide significant at $\alpha_{\text{gw}} = 2.61\text{e-}06$.

Figure S11.1: Estimated lifetime MDD diagnosis: gene-wise tests

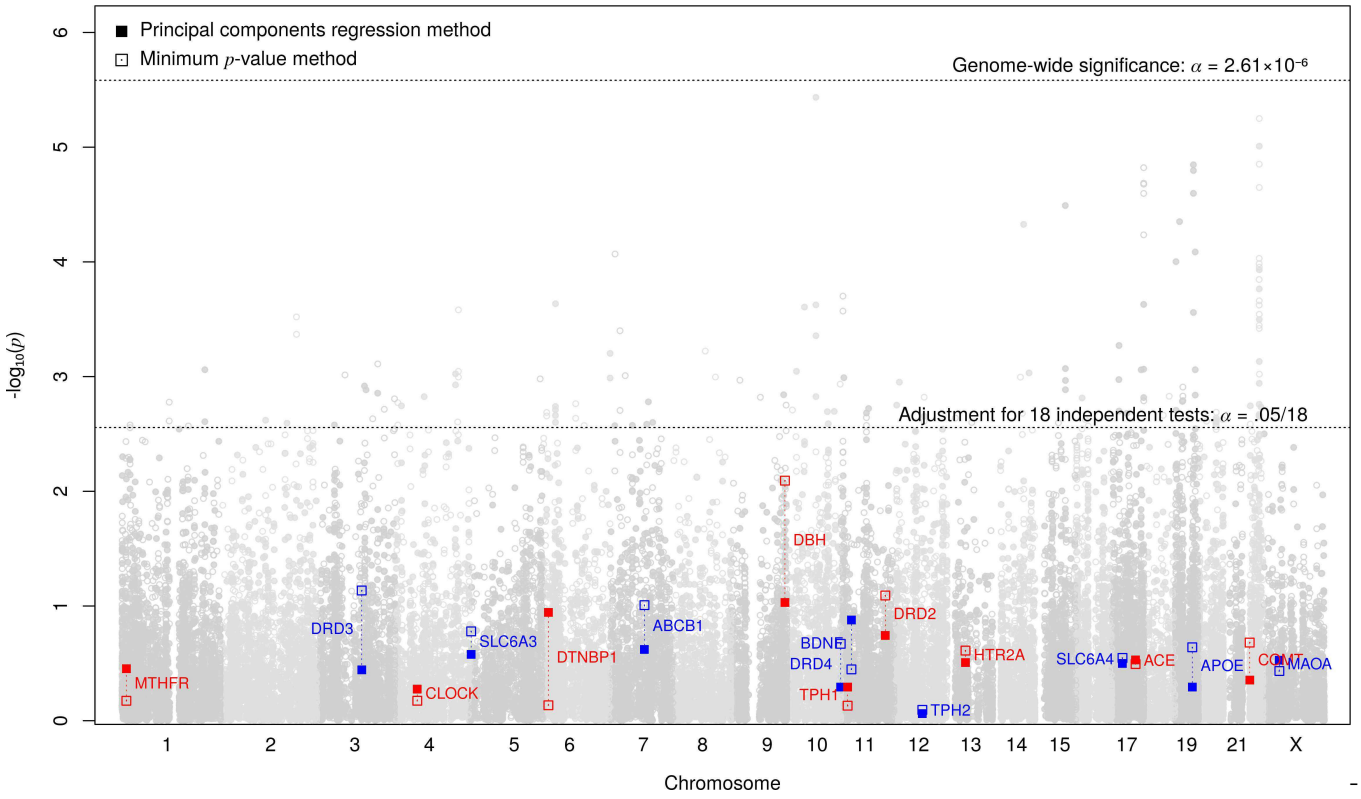


Figure S11.2: Current MDD severity: gene-wise tests

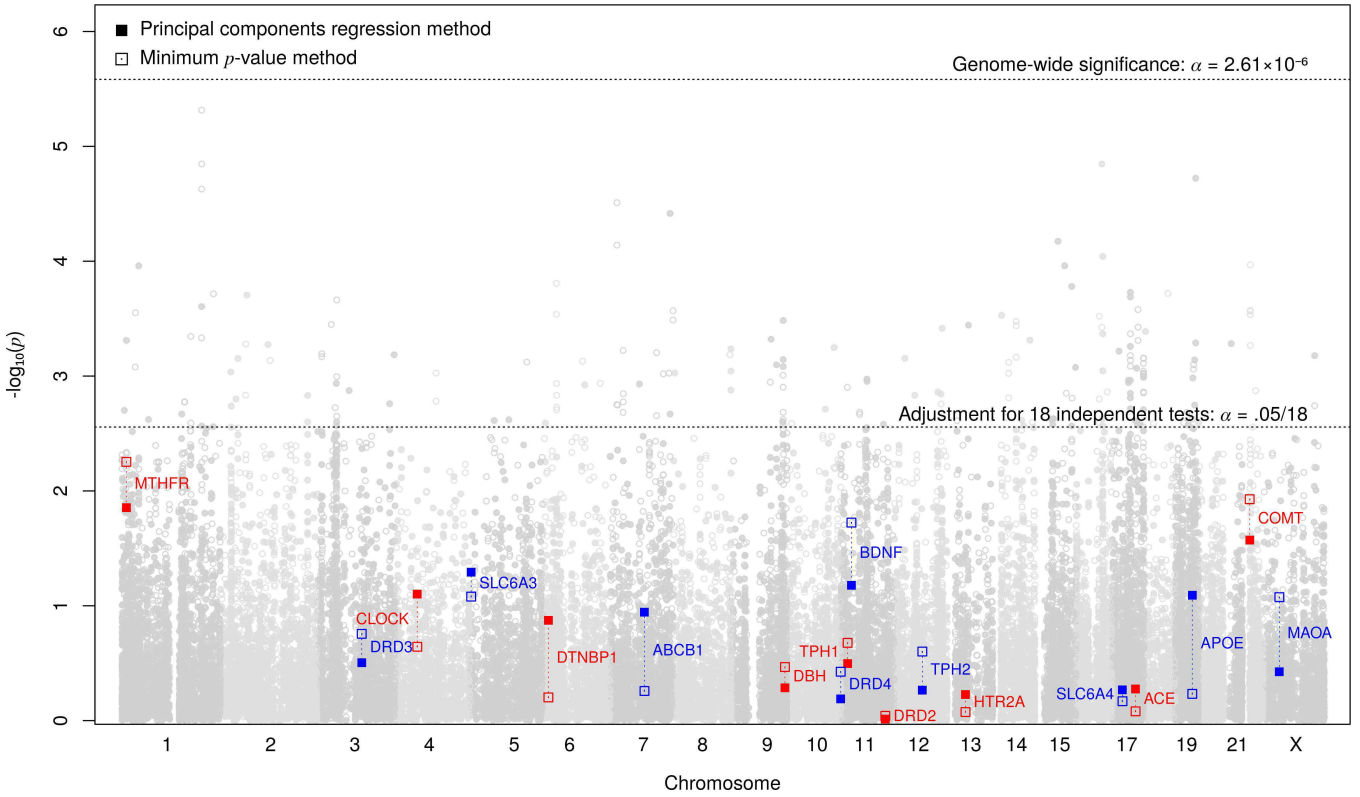


Figure S11.3: Conditional lifetime symptom count: gene-wise tests

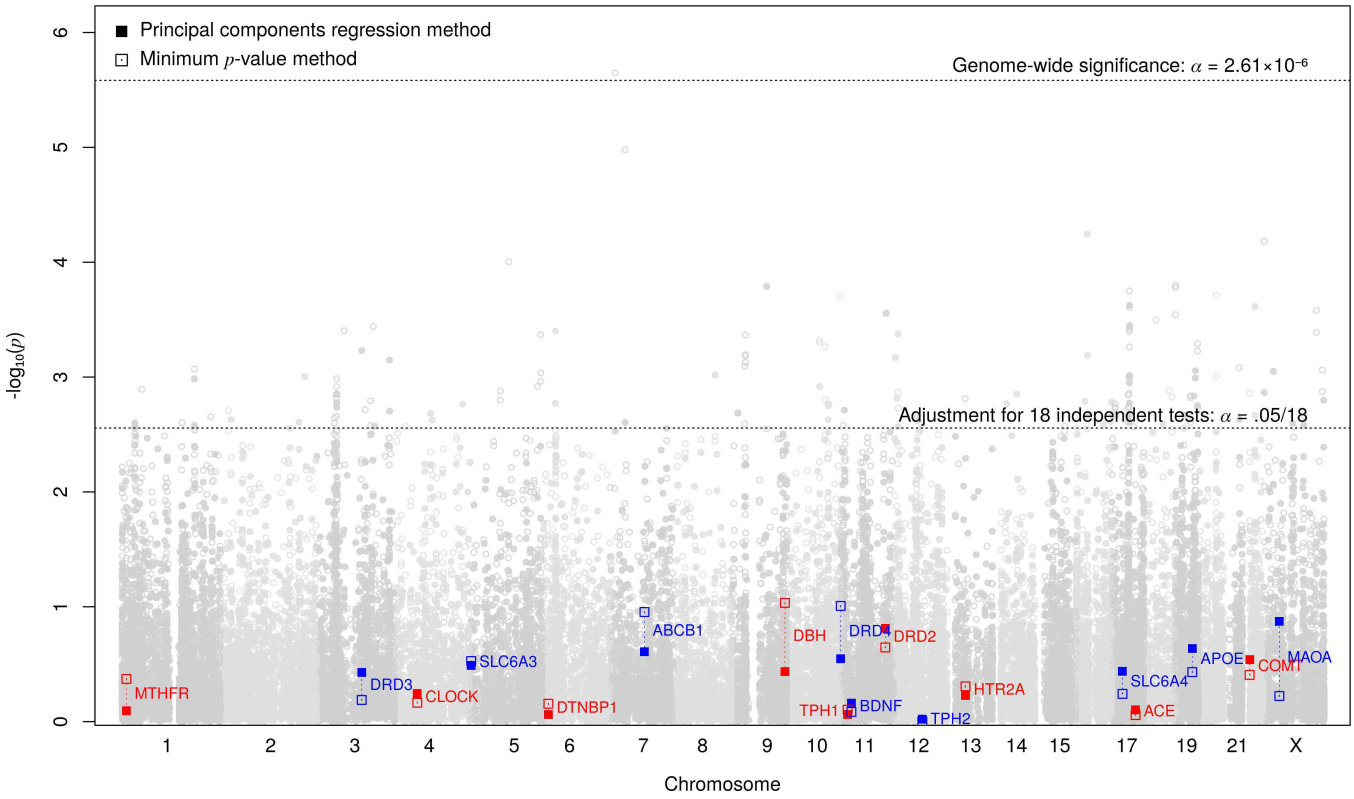


Figure S11.4: Lifetime episode count: gene-wise tests

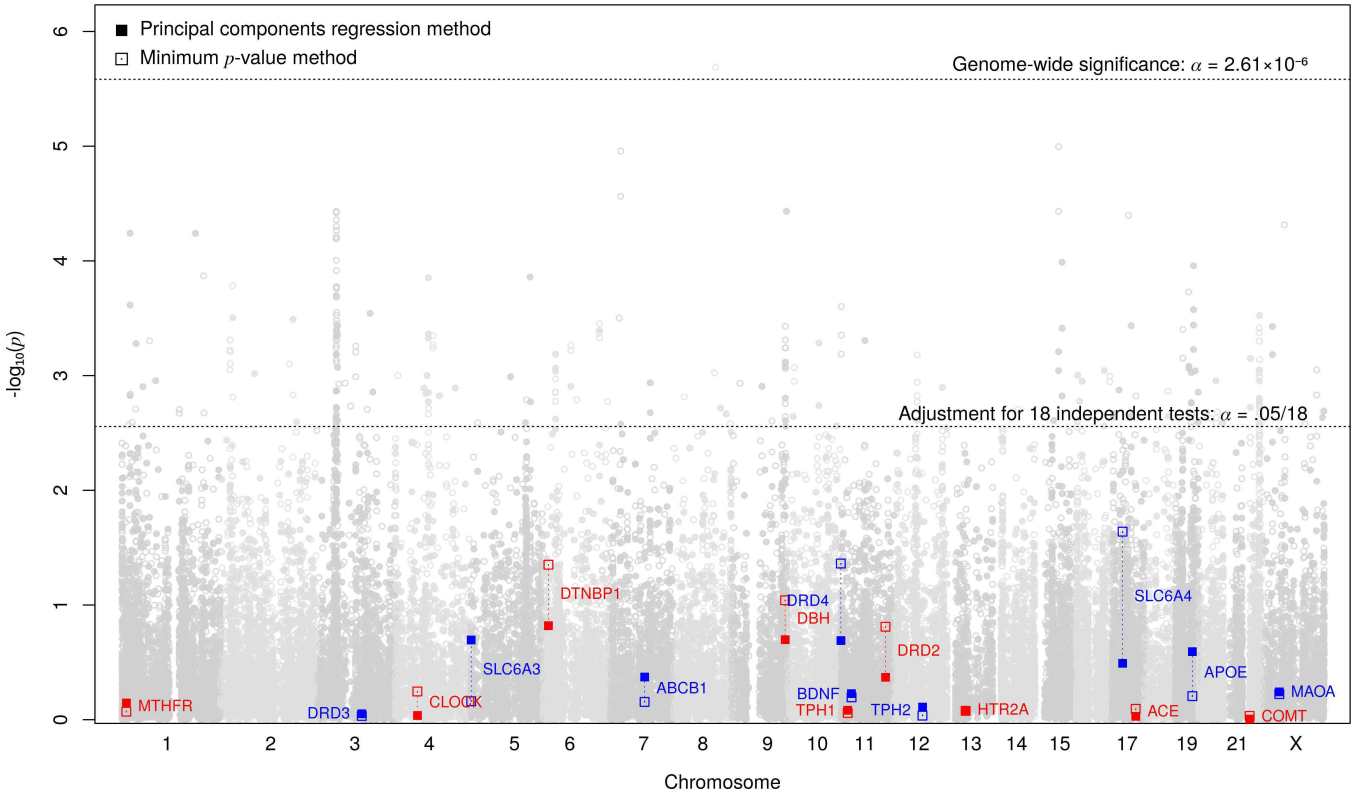


Figure S11.5: Touchscreen probable lifetime diagnosis, ordinal classification: gene-wise tests

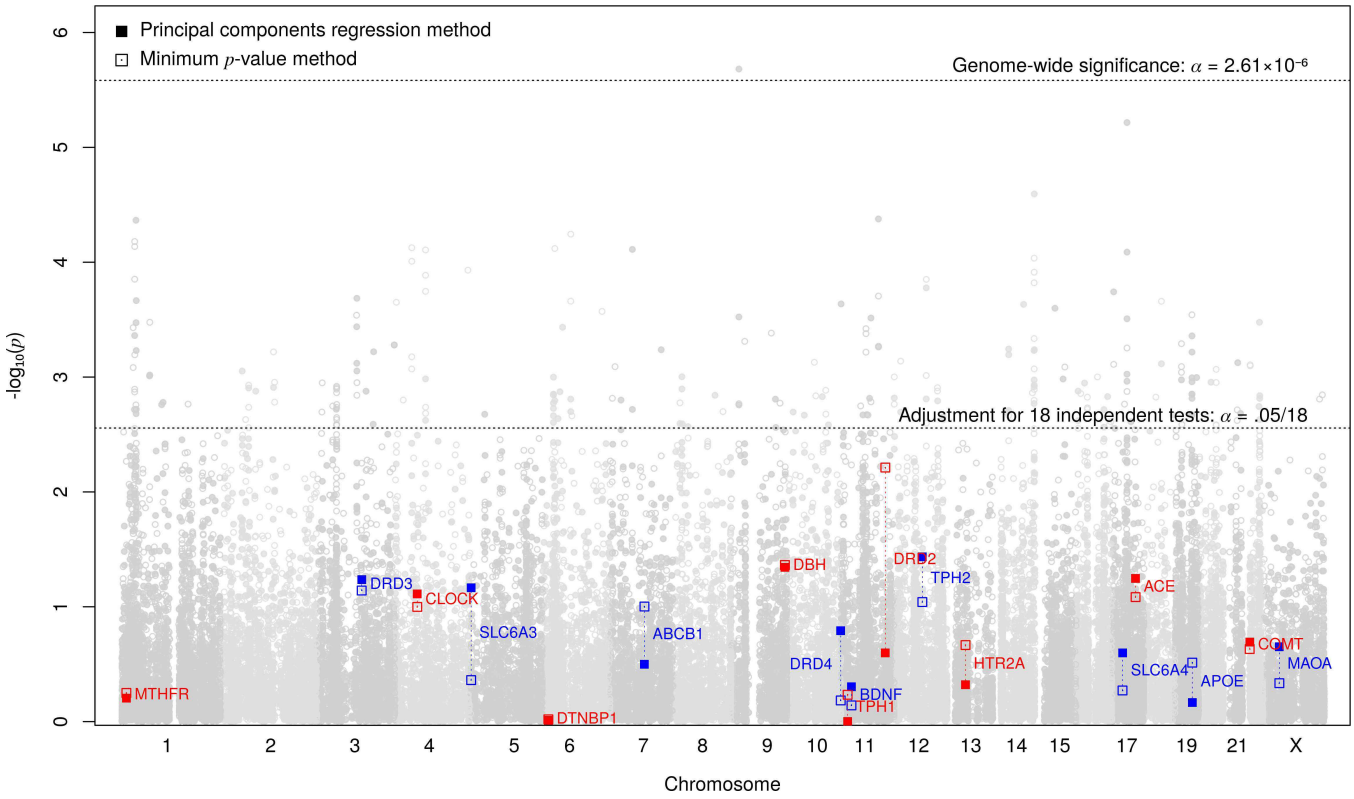


Figure S11.6: Touchscreen probable lifetime diagnosis: gene-wise tests

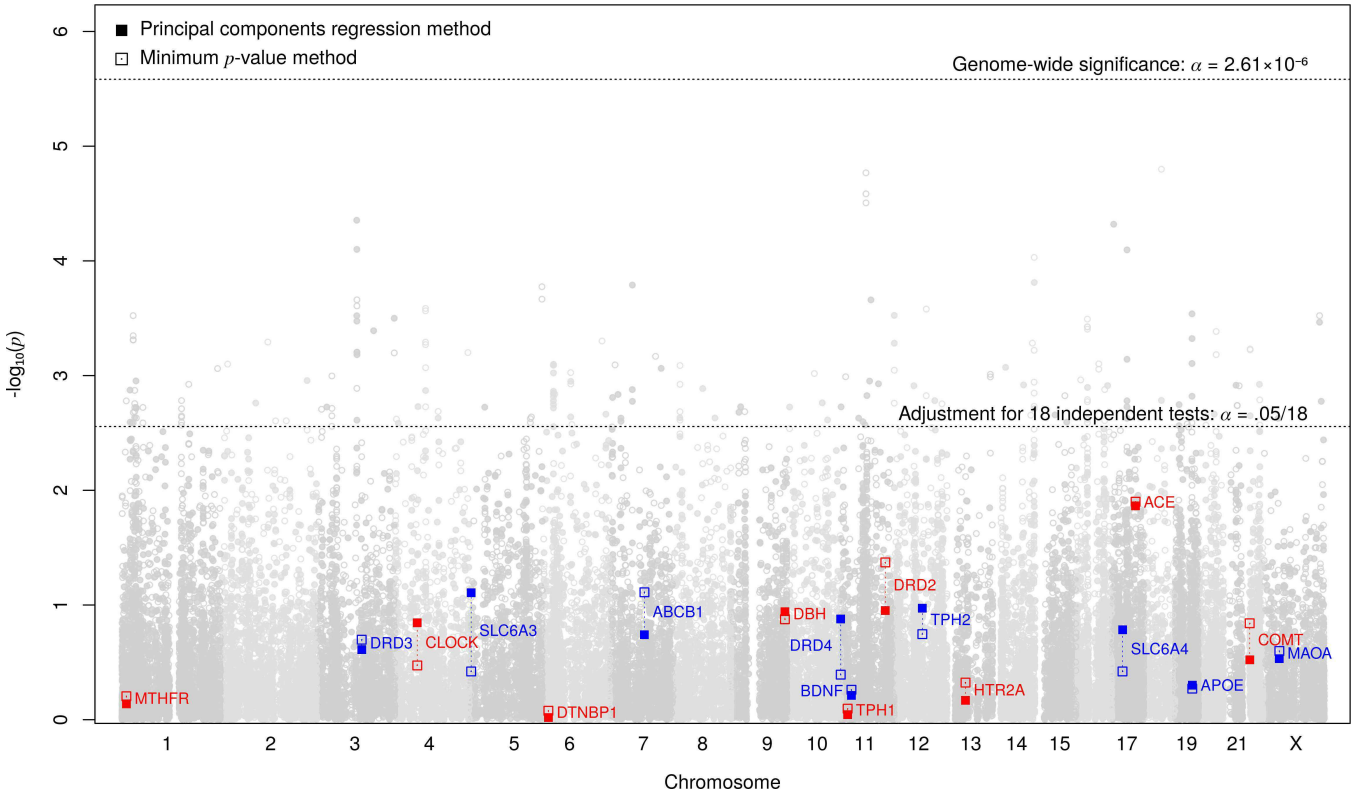


Figure S11.7: PGC lifetime diagnosis: gene-wise tests

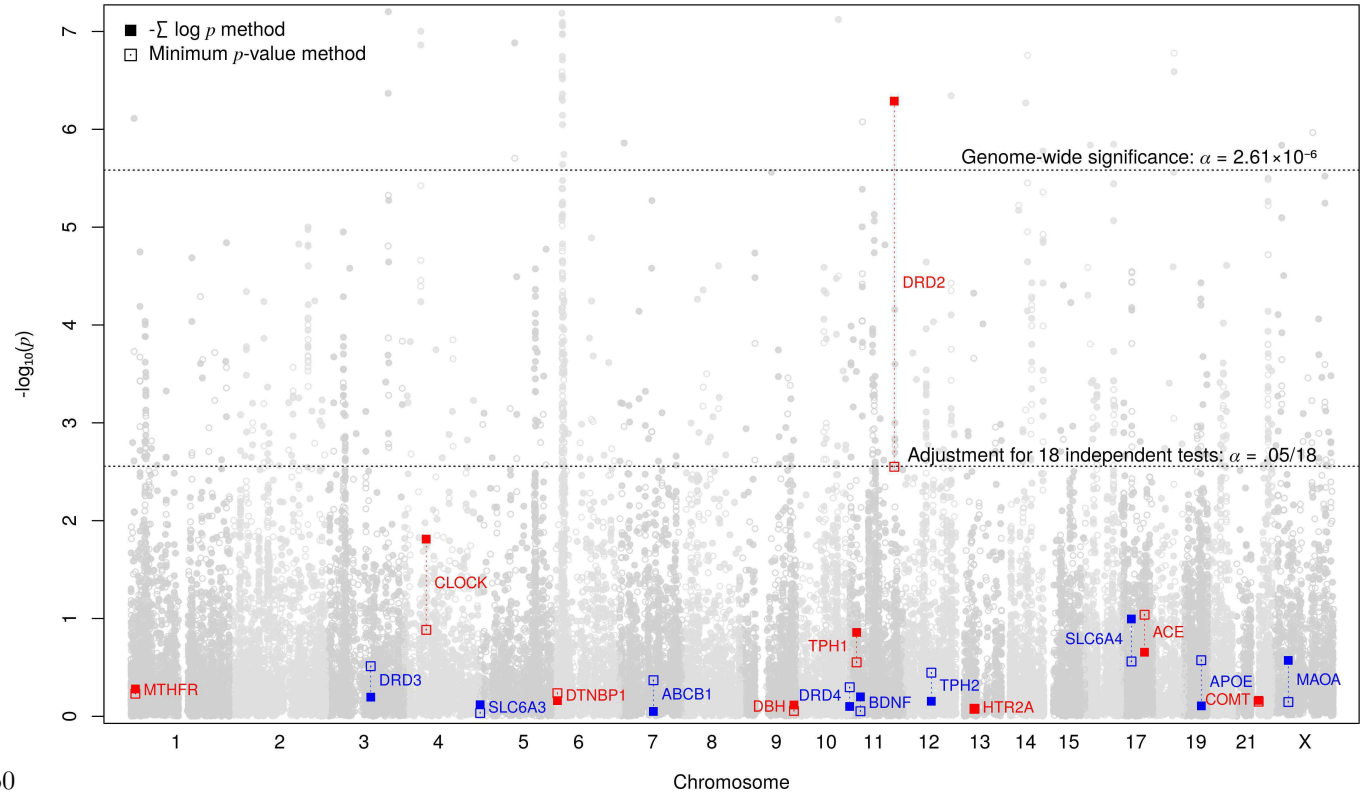
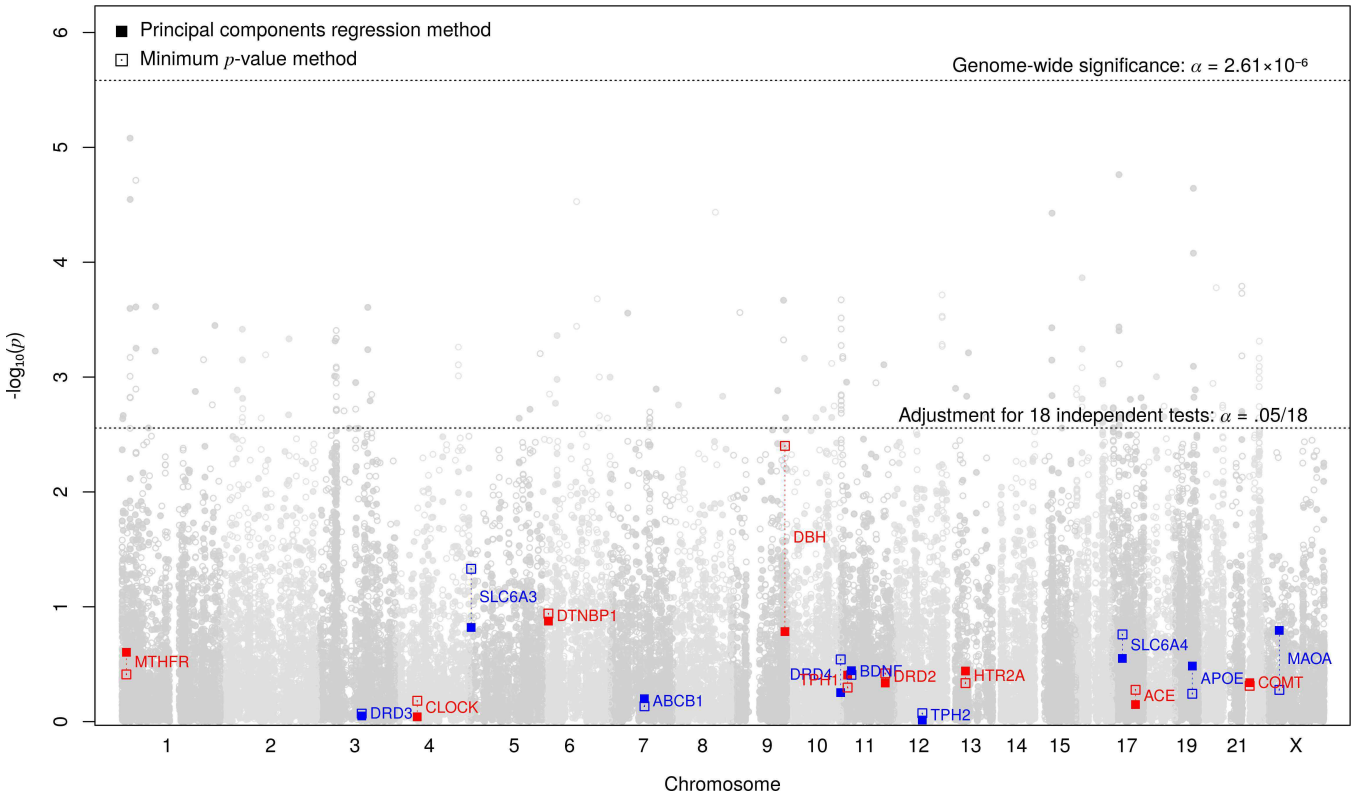


Figure S11.8: Severe recurrent depression: gene-wise tests



S11.2 Gene-set models

S11.2.1 Competitive tests

Table S11.3: Competitive gene-set tests (primary analyses)

Phenotype	β	SE	p	n
Estimated lifetime MDD diagnosis	0.106	0.163	0.26	115,257
Conditional lifetime symptom count	-0.014	0.162	0.53	62,138
Lifetime episode count	-0.163	0.162	0.84	112,261
Touchscreen probable diagnosis	-0.046	0.165	0.61	90,944
Touchscreen probable diagnosis, ordinal classification	0.018	0.166	0.46	90,944
Current MDD severity	0.129	0.162	0.21	115,257
PGC lifetime MDD diagnosis	0.213	0.216	0.16	329,462
Severe recurrent depression	-0.018	0.163	0.55	67,304

Table S11.4: Competitive gene-set tests (secondary analyses)

Phenotype	β	SE	p	n
Estimated lifetime MDD diagnosis	0.205	0.170	0.23	115,258
Conditional lifetime symptom count	-0.050	0.170	0.77	62,138
Lifetime episode count	-0.117	0.170	0.49	112,262
Touchscreen probable diagnosis	-0.014	0.173	0.93	90,945
Touchscreen probable diagnosis, ordinal classification	0.091	0.174	0.60	90,945
Current MDD severity	0.125	0.170	0.46	115,258
PGC lifetime MDD diagnosis	0.117	0.225	0.60	329,463
Severe recurrent depression	0.053	0.170	0.76	67,304

S11.2.2 Relative tests

Table S11.5: Relative gene-set tests (primary analyses)

Phenotype	Comparison	β	SE	p	n
Estimated lifetime MDD diagnosis	Height	0.20	0.170	0.226	115,258
Conditional lifetime symptom count	Height	-0.05	0.170	0.767	62,138
Lifetime episode count	Height	-0.12	0.170	0.492	112,262
Touchscreen probable diagnosis	Height	-0.01	0.173	0.934	90,945
Touchscreen probable diagnosis, ordinal classification	Height	0.09	0.174	0.599	90,945
Current MDD severity	Height	0.12	0.170	0.463	115,258
PGC lifetime MDD diagnosis	Height	0.12	0.225	0.604	329,463
Severe recurrent depression	Height	0.05	0.170	0.755	67,304
Estimated lifetime MDD diagnosis	Synaptic Processes	0.10	0.165	0.537	115,260
Conditional lifetime symptom count	Synaptic Processes	0.01	0.166	0.961	62,138
Lifetime episode count	Synaptic Processes	-0.19	0.166	0.246	112,264
Touchscreen probable diagnosis	Synaptic Processes	-0.10	0.168	0.536	90,947
Touchscreen probable diagnosis, ordinal classification	Synaptic Processes	-0.03	0.169	0.842	90,947
Current MDD severity	Synaptic Processes	0.15	0.166	0.356	115,260
PGC lifetime MDD diagnosis	Synaptic Processes	0.15	0.220	0.490	329,465
Severe recurrent depression	Synaptic Processes	-0.04	0.166	0.814	67,304
Estimated lifetime MDD diagnosis	Type 2 Diabetes	0.13	0.176	0.473	115,259
Conditional lifetime symptom count	Type 2 Diabetes	-0.11	0.177	0.518	62,138
Lifetime episode count	Type 2 Diabetes	-0.06	0.176	0.748	112,263
Touchscreen probable diagnosis	Type 2 Diabetes	0.01	0.179	0.958	90,946
Touchscreen probable diagnosis, ordinal classification	Type 2 Diabetes	0.09	0.181	0.628	90,946
Current MDD severity	Type 2 Diabetes	0.18	0.177	0.315	115,259
PGC lifetime MDD diagnosis	Type 2 Diabetes	0.18	0.231	0.428	329,464
Severe recurrent depression	Type 2 Diabetes	0.14	0.177	0.432	67,304

Table S11.6: Relative gene-set tests (secondary analyses)

Phenotype	Comparison	β	SE	p	n
Estimated lifetime MDD diagnosis	Height	0.341	0.201	0.089	115,262
Conditional lifetime symptom count	Height	-0.449	0.199	0.024	62,138
Lifetime episode count	Height	-0.316	0.204	0.122	112,266
Touchscreen probable diagnosis	Height	0.251	0.219	0.252	90,949
Touchscreen probable diagnosis, ordinal classification	Height	0.223	0.221	0.313	90,949
Current MDD severity	Height	0.357	0.203	0.079	115,262
PGC lifetime MDD diagnosis	Height	0.048	0.206	0.814	329,467
Severe recurrent depression	Height	0.115	0.205	0.574	67,304
Estimated lifetime MDD diagnosis	Synaptic Processes	0.408	0.196	0.038	115,264
Conditional lifetime symptom count	Synaptic Processes	-0.479	0.194	0.013	62,138
Lifetime episode count	Synaptic Processes	-0.350	0.200	0.080	112,268
Touchscreen probable diagnosis	Synaptic Processes	0.212	0.215	0.326	90,951
Touchscreen probable diagnosis, ordinal classification	Synaptic Processes	0.237	0.217	0.276	90,951
Current MDD severity	Synaptic Processes	0.285	0.198	0.151	115,264
PGC lifetime MDD diagnosis	Synaptic Processes	0.006	0.203	0.975	329,469
Severe recurrent depression	Synaptic Processes	0.158	0.201	0.430	67,304
Estimated lifetime MDD diagnosis	Type 2 Diabetes	0.292	0.209	0.162	115,263
Conditional lifetime symptom count	Type 2 Diabetes	-0.477	0.207	0.021	62,138
Lifetime episode count	Type 2 Diabetes	-0.223	0.213	0.295	112,267
Touchscreen probable diagnosis	Type 2 Diabetes	0.343	0.227	0.131	90,950
Touchscreen probable diagnosis, ordinal classification	Type 2 Diabetes	0.291	0.229	0.203	90,950
Current MDD severity	Type 2 Diabetes	0.381	0.211	0.071	115,263
PGC lifetime MDD diagnosis	Type 2 Diabetes	0.111	0.212	0.601	329,468
Severe recurrent depression	Type 2 Diabetes	0.254	0.214	0.236	67,304

S12 Attempted replication of top 16 independent PGC associations

Table S12.1: Attempted replication of top 16 independent PGC genome-wide significant loci in UKBB

	Variant	*Chr.*	*BP*	*Risk allele*	*Freq.*	p_{PGC}	$OR_{\text{PGC}}^{\text{Uncorrected}}$	$OR_{\text{PGC}}^{\text{Corrected}\dagger}$	$\text{Power}_{\text{UKBB}}^{\ddagger}$	OR_{UKBB}	se_{UKBB}	p_{UKBB}	n_{UKBB}
1.	rs12552	13	53625781	A	0.421	6.093e-15	1.044	1.044	0.941	1.037	0.010	1.868e-04*	114,324
2.	rs1432639	1	72813218	C	0.601	1.065e-11	1.040	1.040	0.861	0.982	0.010	6.383e-02	114,852
3.	rs10044618	5	87781168	T	0.550	2.416e-10	1.036	1.036	0.767	1.007	0.010	5.037e-01	112,796
4.	rs12129573	1	73768366	A	0.322	4.556e-09	1.034	1.029	0.445	1.018	0.010	7.401e-02	115,192
5.	rs834629	15	37678862	C	0.581	5.059e-09	0.967	0.972	0.506	1.010	0.010	3.348e-01	114,830
6.	chr15_37664874_D	15	37664874	del	0.371	5.775e-09	1.034	1.028	0.433	1.007	0.010	4.884e-01	114,148
7.	rs12886138	14	64871010	T	0.356	8.721e-09	0.968	0.975	0.339	0.997	0.010	7.533e-01	111,943
8.	rs7198928	16	7666402	C	0.608	1.035e-08	1.033	1.023	0.253	0.985	0.010	1.238e-01	111,672
9.	rs61867293	10	106563924	T	0.150	1.646e-08	0.961	0.985	0.033	0.977	0.012	5.483e-02	113,996
10.	rs10214154	5	87545319	G	0.281	1.689e-08	1.037	1.010	0.023	0.997	0.011	7.852e-01	113,303
11.	rs1806153	11	31850105	T	0.312	1.707e-08	1.039	1.011	0.029	1.014	0.011	2.208e-01	114,910
12.	rs12658032	5	103904226	A	0.334	2.346e-08	1.033	1.011	0.032	1.033	0.010	1.232e-03*	114,912
13.	rs11135349	5	164523472	A	0.418	2.480e-08	0.969	0.998	0.004	0.967	0.010	4.805e-04*	115,120
14.	rs3095337	6	30737591	C	0.204	3.256e-08	0.960	0.998	0.004	0.953	0.018	6.535e-03	115,122
15.	rs12958048	18	53101598	A	0.385	3.376e-08	1.033	1.001	0.003	1.027	0.010	1.109e-02	114,873
16.	rs10514301	5	87939654	T	0.165	3.627e-08	1.047	1.002	0.003	1.024	0.015	1.124e-01	114,704

The top 16 independent genome-wide significant loci for PGC lifetime MDD diagnosis tested for association with estimated lifetime diagnosis in the independent UKBB sample. Estimated effect in the PGC were corrected for bias due to the winner's curse and used to estimate power to detect associations in the UKBB. Three loci were significant in the UKBB at $p < .05/16$. The 95% CI for the number of replications to be expected given power in the UKBB was 2 - 7 (S4.6). See S4.5 for further details.

†Unbiased estimator correcting for winner's curse ([38]); ‡Power to detect corrected locus effects in UKBB at $\alpha_{\text{CG}} = .05/16$; *Significant at $p < .05/16$.

References

[1] Cock, P. J. A. *et al.* Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422–1423 (2009). URL https://academic.oup.com/bioinformatics/article/25/11/1422/330687.

[2] National Center for Biotechnology Information, US National Library of Medicine & National Institutes of Health. PubMed. URL https://www.ncbi.nlm.nih.gov/pubmed/.

[3] HUGO Gene Nomenclature Committee (HGNC) & European Molecular Biology Laboratory. HGNC Database. URL www.genenames.org.

[4] Yates, B. *et al.* Genenames.org: the HGNC and VGNC resources in 2017. *Nucleic Acids Research* **45**, D619–D625 (2017).

[5] Wang, W. Exact Optimal Confidence Intervals for Hypergeometric Parameters. *Journal of the American Statistical Association* **110**, 1491–1499 (2015). URL https://doi.org/10.1080/01621459.2014.966191.

[6] Ferreira, M. A. R. *et al.* Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1c in bipolar disorder. *Nature genetics* **40**, 1056–1058 (2008). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2703780/.

[7] Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics (Oxford, England)* **26**, 2190–2191 (2010).

[8] Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics* **50**, 668–681 (2018). URL https://www.nature.com/articles/s41588-018-0090-3.

[9] Bycroft, C. *et al.* Genome-wide genetic data on ˜500,000 UK Biobank participants. *bioRxiv* 166298 (2017). URL https://www.biorxiv.org/content/early/2017/07/20/166298.

[10] Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, 7 (2015).

[11] Purcell, S. M. & Chang, C. PLINK 1.9. URL www.cog-genomics.org/plink/1.9/.

[12] Border, R. *et al.* Imputation of Behavioral Candidate Gene Repeat Polymorphisms in 486,551 Publicly-Available UK Biobank Individuals. *bioRxiv* 358267 (2018). URL https://www.biorxiv.org/content/early/2018/06/29/358267.

[13] Heils, A. *et al.* Allelic variation of human serotonin transporter gene expression. *Journal of Neurochemistry* **66**, 2621–2624 (1996).

[14] Smith, D. J. *et al.* Prevalence and Characteristics of Probable Major Depression and Bipolar Disorder within UK Biobank: Cross-Sectional Study of 172,751 Participants. *PLOS ONE* **8**, e75362 (2013). URL http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0075362.

[15] Spitzer, R. L., Kroenke, K., Williams, J. B. & Group, P. H. Q. P. C. S. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Jama* **282**, 1737–1744 (1999).

[16] First, M. B., Gibbon, M., Spitzer, R. L. & Williams, J. B. W. User's guide for the structured clinical interview for DSM-IV axis I Disorders—Research version. *New York: Biometrics Research Department, New York State Psychiatric Institute* (1996).

[17] Townsend, P. Deprivation. *Journal of social policy* **16**, 125–146 (1987). Citation Key: townsend1987deprivation bibtex[publisher=Cambridge University Press].

[18] Yzerbyt, V. Y., Muller, D. & Judd, C. M. Adjusting researchers' approach to adjustment: On the use of covariates when testing interactions. *Journal of Experimental Social Psychology* **40**, 424–431 (2004). URL http://www.sciencedirect.com/science/article/pii/S0022103103001598.

[19] Hull, J. G., Tedlie, J. C. & Lehn, D. A. Moderator Variables in Personality Research: The Problem of Controlling for Plausible Alternatives , Moderator Variables in Personality Research: The Problem of Controlling for Plausible Alternatives. *Personality and Social Psychology Bulletin* **18**, 115–117 (1992). URL https://doi.org/10.1177/0146167292182001.

[20] Keller, M. C. Gene Environment Interaction Studies Have Not Properly Controlled for Potential Confounders: The Problem and the (Simple) Solution. *Biological Psychiatry* **75**, 18–24 (2014). URL https://www.biologicalpsychiatryjournal.com/article/S0006-3223(13)00825-1/abstract.

[21] Björck, Å. *Numerical methods in matrix computations*, vol. 59 (Springer, 2015).

[22] Leeuw, C. A. d., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Computational Biology* **11**, e1004219 (2015). URL http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004219.

[23] MacArthur, J. *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Research* **45**, D896–D901 (2017). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5210590/.

[24] Lips, E. S. *et al.* Functional gene group analysis identifies synaptic gene groups as risk factor for schizophrenia. *Molecular Psychiatry* **17**, 996–1006 (2012).

[25] Ruano, D. *et al.* Functional gene group analysis reveals a role of synaptic heterotrimeric G proteins in cognitive ability. *American Journal of Human Genetics* **86**, 113–125 (2010).

[26] Liu, J. Z. *et al.* A versatile gene-based test for genome-wide association studies. *American Journal of Human Genetics* **87**, 139–145 (2010).

[27] Consortium, T. . G. P. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015). URL https://www.nature.com/articles/nature15393.

[28] Purcell, S., Cherny, S. S. & Sham, P. C. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* **19**, 149–150 (2003). URL https://academic.oup.com/bioinformatics/article/19/1/149/316873.

[29] Kessler, R. C. & Bromet, E. J. The epidemiology of depression across cultures. *Annual review of public health* **34**, 119–138 (2013). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4100461/.

[30] Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nature Genetics* **47**, 1236–1241 (2015). URL https://www.nature.com/articles/ng.3406.

[31] Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics* **47**, 291–295 (2015). URL https://www.nature.com/articles/ng.3211.

[32] Loh, P.-R., Kichaev, G., Gazal, S., Schoech, A. P. & Price, A. L. Mixed-model association for biobank-scale datasets. *Nature genetics* 1 (2018).

[33] Wood, A. R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics* **46**, 1173–1186 (2014).

[34] Morris, A. P. *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nature Genetics* **44**, 981–990 (2012).

[35] Holman, N., Young, B. & Gadsby, R. What is the current prevalence of diagnosed and yet to be diagnosed diabetes in the UK. *Diabetic Medicine* **31**, 510–511. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/dme.12397.

[36] Huang, J. *et al.* Improved imputation of low-frequency and rare variants using the UK10k haplotype reference panel. *Nature Communications* **6**, 8111 (2015). URL https://www.nature.com/articles/ncomms9111.

[37] Xiao, R. & Boehnke, M. Quantifying and correcting for the winner's curse in genetic association studies. *Genetic epidemiology* **33**, 453–462 (2009). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2706290/.

[38] Zhong, H. & Prentice, R. L. Bias-reduced estimators and confidence intervals for odds ratios in genome-wide association studies. *Biostatistics* **9**, 621–634 (2008). URL https://academic.oup.com/biostatistics/article/9/4/621/258822.

[39] Cross-validated mixed-datatype bandwidth selection for nonparametric cumulative distribution/survivor functions: Econometric Reviews: Vol 36, No 6-9. URL https://www.tandfonline.com/doi/full/10.1080/07474938.2017.1307900?src=recsys.