Data Supplement for Peterson et al., Molecular genetic analysis subdivided by adversity exposure suggests etiologic heterogeneity in major depression. Am J Psychiatry (doi: 10.1176/appi.ajp.2017.17060621)

Contents

Supplemental Methods

1.   Sample collection

Recurrent major depression (MD) cases were recruited from 58 provincial mental health centers and psychiatric departments of general medical hospitals in 45 cities and 23 provinces of China (1). Controls were recruited from multiple locations including general hospitals and local community centers. All subjects were Han Chinese women with four Han grandparents. Cases were between 30 and 60 and had two or more episodes of MD meeting DSM-IV criteria with the first episode between ages 14 and 50, had not abused drugs or alcohol before their first depressive episode, and reported no history of schizophrenia or mania. The study protocol was approved by the Ethical Review Board of Oxford University and the ethics committees of all participating hospitals. All participants provided their written informed consent.

2.   Sample phenotyping

All subjects were interviewed using a computerized assessment system, which lasted on average two hours for a case and one hour for a control. All interviewers were trained by the **C**hina, **O**xford, a**n**d **V**CU **E**xperimental **R**esearch on **G**enetic **E**pidemiology (CONVERGE) team for a minimum of one week in the use of the interview. The interview includes assessment of psychopathology, demographic and personal characteristics, and psychosocial functioning. Interviews were tape-recorded and a proportion of them were listened to by the trained editors who provided feedback on the quality of the interviews. Diagnosis of MD was established with the Composite International Diagnostic Interview (CIDI) (WHO lifetime version 2.1; Chinese version), which classifies diagnoses according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) criteria. The interview was originally translated into Mandarin by a team of Chinese psychiatrists with the translation reviewed and modified by members of the CONVERGE team.

3.   Adversity measures

A binary measure of adversity was derived from self-reported stressful life events (SLE) and childhood sexual abuse (CSA) in order to identify individuals exposed to *severe* environmental adversities while reducing the burden of multiple testing. The SLE questionnaire was adapted from the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (2) and assessed 16 traumatic lifetime events and the age of their occurrence (Table S1). Because of evidence that sensitive subjects like CSA are more accurately reported with more confidential methods of assessment (3), participants were asked to fill in a paper questionnaire about CSA. The questions asked whether, before the subject was 16, did any adult, or any other older person, involve the subject in any unwanted incidents like inviting or requesting them to do something sexual, kissing or hugging in a sexual way, touching or fondling private parts, showing their sex organs, making them touch the person in a sexual way, or attempting or having sexual intercourse. The possible responses were "never," "once," and "more than once." We used these responses to define three forms of CSA (12): i) *nongenital* CSA including sexual invitation, sexual kissing,

and exposing ii) *genital* CSA including fondling and sexual touching and iii) attempted or completed *intercourse*. While it is known that the patterns of association with MD differ between some forms of environmental stressors (4), this heterogeneity is expected to be comparatively small and so for these analyses we grouped them together to maximize statistical power and reduce multiple testing. Subjects were considered an "adversity case" if they i) had non-missing data on SLE and CSA scales and ii) endorsed any CSA and/or had high aggregate SLE scores (+3SD). Since SLE vary in severity, the SLE score was constructed by weighting each item by their effect size on MD and summed across each of the 16 items. SLE in MD cases were only included if they preceded MD onset. We assume this data-adaptive weighting does not significantly bias our downstream inferences: (1) we are focused on the interaction effect (even though various forms of adversity are known risk factors for developing MD, there is no *a priori* reason to assume that it would impact the direction of an interaction of such a variable with genetic risk), (2) the weights are not used directly in analyses - only used to identify if the individual has experienced extreme adversity or not, and (3) our weighted score is similar to the SLE count score ($r = 0.698$, *P*-value $< 2.2 \times 10^{-16}$) and previous weights from an independent US sample of European descent (e.g., assault/marital problems strongest, financial/job-loss weakest) (5). Using adversity exposure status we grouped samples into those that are "adversity exposed" and those that are "not exposed". In order to reduce effects of retrospective recall, both sections of the CSA and SLE questions were separate from the MD section, occurring many minutes later toward the end of the interview. We did this in part to reduce the chances of "correlated errors" in recall that might arise in "seeking after meaning."

4.   Sample sequencing, imputation, and genotype quality control

CONVERGE obtained sequence on 11,670 samples and imputed genotypes using methods described in Cai et al. (6,7). Twenty-nine samples were excluded because of low imputation quality (maximum genotype probability < 0.9) in more than 10% of imputed sites. 392 samples were removed for being likely duplicates or first-degree relatives and beyond. 117 samples were excluded due to an excess number of private variants in the genic regions of their nuclear genome. A further 90 samples were excluded on the basis of an excess number of heteroplasmic sites in their mitochondrial genome. Removing a further 431 samples with incomplete phenotype information would give the set of 10,640 samples on which we previously reported our primary GWAS on MD (5,303 cases of MD, 5,337 controls) (6). As our current analyses involved testing associations using the logistic regression framework which is less effective in controlling for relatedness between samples, we further removed 164 samples with genetic relatedness greater than 0.15 (leaving us with 5,279 cases, 5197 controls), of which 877 samples did not self-report information on adversity, giving us a final set of 9,599 samples (4,785 cases, 4,814 controls) with full phenotypic information for our final analysis.

5.   Obtaining odds ratio for major depression using logistic regression

While the linear mixed model provides increased power, effect sizes obtained on a linear scale from retrospective, ascertained case-control data are difficult to relate to the prospective effect sizes that are of primary interest. To obtain and access the difference between odds ratios (OR) of associations in different cohorts, we perform logistic regression on all 4,313,801 SNPs for all

the above analyses in PLINK (version 1.9) (8,9) with 10 principal components (PC) obtained from eigen-decomposition of a kinship matrix constructed with 413,669 linkage disequilibrium (LD)-pruned SNPs (that were used for linear mixed model association in BOLT-LMM (10)) computed using GCTA (version 1.26.0) (11). We consistently use these PCs for all logistic regression analyses (including gene-by-environment interaction test).

6.   Determining differences in odds ratios of major depression associated loci between the adversity unexposed group and the complete CONVERGE cohort

To determine whether difference in the OR we observe between the whole cohort and the adversity unexposed subset at rs7526682 and rs11577545 on chromosome 1, rs950893 on chromosome 8, and rs12415800 and rs35936514 on chromosome 10 were due to chance fluctuations, we obtained empirical distributions of the ORs in subsets of samples of the same size as the unexposed subset, and compared them to the observed values. We generate 10,000 subsets of the cohort by randomly excluding 2,702 samples (1,646 cases, 982 controls) from the full cohort (matching numbers of MD cases and controls with self-reported adversity), and run logistic regression in PLINK v1.9 (8,9) between the three SNPs as well as rs12415800 and rs35936514 on chromosome 10 for comparison, using 10 PCs as covariates. We generate an empirical distribution of ORs using the 10,000 random subsets, and obtain the percentile of the observed ORs in the adversity unexposed cohort in the empirical distribution. We consider the observed ORs from the unexposed cohort a significant change in ORs from the full cohort if it is in the top or bottom 0.5th percentile of the empirical distribution (corresponding to P-value < 0.05, after multiple testing correction for 5 SNPs). Figure S5 and Figure S6 show the distribution of 10,000 empirically derived ORs with the vertical red line on each histogram showing the observed OR for the same SNP in the adversity unexposed group.

7.   Gene-by-environment interactions on the additive scale

Interactions significant on one scale may not be so on another (12). Because its results are more interpretable, the additive scale is also commonly used (13), but there are statistical issues with testing additive regression models for binary data which may lead to loss of power or false positive tests of interaction (12). With these caveats in mind, we report in Table S2 and Table S5 interactions on the additive scale using one such test (BLM) (14). For all analyses we examined boundary constraints and found none were active at the default criterion of 10^{-6}.

8.   Aggregate gene-by-environment interactions

Aggregate GxE interactions were tested by two methods. First, GCTA was used to estimate the proportion of variance in major depression due to aggregate additive gene-by-environment interaction between adversity and all GRM SNPs (11,15). Here, the main effect of the environmental variable (adversity) is included in the model as a fixed effect and the GxE interaction effect is treated as a random effect. A significant variance component for GRMxE indicates support for 'omnibus' GxE interaction across the genome regardless of direction and

can be considered a 'qualitative test'. Second, we tested for significant polygenic risk score by environment interaction (PRSxE). In contrast to GRMxE interaction, PRSxE interactions can be considered a 'quantitative test' and is dependent on a consistent direction of SNPxE interactions. If the PRS includes too few SNPs that have true GxE effects in the same direction, too many null SNPs, or is not a sufficient indicator of genetic risk (low r^2) then a PRSxE interaction would not be detected which highlights some limitations of a PRS-based approach.

9.   Adjustment for ascertainment in SNP-based heritability estimation

To adjust for ascertainment, we obtain adjusted SNP-based heritability (h^2_{SNP}) at best-estimate population prevalences for each group. We previously reported our best estimate of a population prevalence (K) of MD of 8% (16). Assuming the sample adversity prevalences within cases and controls are similar to their population counterparts, we estimate that the population prevalence of MD in the exposed is 0.128 and of the unexposed is 0.066 (see below).

In detail, we start off with a prevalence estimate of $K = 0.08$ for MD in the whole cohort, then we make the following premises:

1. We have ascertained MD in our study since it is a case control study

2. BUT the adversity measure is NOT ascertained as we have not collected our cases and controls based on whether they have had any adversities in the past. This means proportions of adversity-exposed vs unexposed samples in our depression (MD) cases and controls mirror that in cases of MD and non-cases in the population. This is regardless of whether MD and adversity are independent.

 Therefore, in our data where:

nCases = nCaseAdversity + nCaseNoAdversity = 1646+3139

nAdversity = nAdversityCases + nAdversityControls = 1646+982

Assuming premise 2 is true, then prevalence of adversity in depression cases and controls can be obtained from our data (K denotes prevalence):

K_AdversityCases = nAdversityCases/nCases = 1646/(1646+3139) = 0.34

K_AdversityControls = nAdversityControls/nControls = 982/(982+3832) = 0.20

Then, knowing prevalence for depression is 0.08, we estimate population prevalence of adversity by:

K_AdversityAll = K_AdversityCases*K_MDAll + K_AdversityControls*(1-K_MDAll) = 0.34*0.08 + 0.20*0.92 = 0.215

Prevalence of depression in the adversity exposed group is essentially the conditional probability of having depression when one had been exposed to adversity, and that of depression in non-exposed group is the conditional probability of having depression when one had not been exposed to adversity. Hence, we are able to apply Bayes' Rule $P(A|B) = P(A \text{ and } B)/P(B)$ to arrive at the conditional probabilities as stated above. This is shown explicitly below:

K_MDAdversity = P(MD|Adversity)

= P(MD intersect Stress)/P(Adversity)

= P(Adversity|MD)*P(MD)/P(Adversity)

= K_AdversityCases*K_MD/K_Adversity

= 0.34*0.08/0.215

= 0.127

K_MDNoAdversity = P(MD|NoAdversity)

= P(MD intersect NoStress)/P(NoAdversity)

= P(NoAdversityMD)*P(MD)/P(NoAdversity)

= (1-K_AdversityCases)*K_MD/(1-K_Adversity)

= (1-0.34)*0.08/(1-0.215)

=0.067

10.   Estimation of heritability under different assumptions and frameworks

We note that differences in methods used for h^2_{SNP} estimations may have large impacts on the estimates and their interpretations (17-20). In addition to using GCTA for our estimates, we performed the following analysis to formally account for uneven LD in dense imputed data and potential biases in h^2_{SNP} estimations from restricted maximum likelihood (REML). First, to fully utilize all imputed SNPs in capturing genetic variation, while accounting for uneven LD, we also calculated h^2_{SNP} estimates using an LD-weighted GRM constructed with all 4,313,801 SNPs used in association testing and their LD-weights computed in LDAK (version 5.9) (17), with 10 PCs from the eigen-decomposition of this GRM as covariates. Second, we assessed potential underestimates from REML in ascertained case/control data using PCGC (19)), using the same GRM we use for GCTA, and the 10 PCs from the eigen-decomposition of this GRM as covariates. For both analyses, we report h^2_{SNP} estimates on the liability scale corrected for population prevalence as stated above. The results are displayed in Table S4 and are consistent with the trends shown in h^2_{SNP} from GCTA analyses.

11.   Assessment of G-E correlation

Since G-E correlation can bias GxE results, we tested for G-E correlation by three methods. First, using the bivariate option in GCTA (15) the genetic correlation (ρ) was estimated between MD and adversity and tested if different from 0 or 1. Second, G-E correlation was examined by testing for association of MD-PRS with adversity in the full sample, MD-cases only, and controls only. Finally, an exploratory test of G-E correlation was examined by genome-wide correlation of SNP odds ratios from comparing results from adversity GWAS of MD-cases to controls. In order to obtain odds ratios for case-control group comparisons logistic regression was run in PLINK as described above in Supplemental Methods (5. Obtaining odds ratio for major depression using logistic regression).

12.   Simulations of a single SNP effect under the presence and absence of etiologic heterogeneity

Overview

We use simulations to mirror genetic approaches presented above to discern features of etiologic heterogeneity and to demonstrate stratification of samples (such as adversity) are appropriate means for uncovering heterogeneous genetic effects. The three simulated scenarios were as follows:

   a) SNP effect and adversity exposure contribute additively to MD liability (no etiologic heterogeneity),
   b) SNP effect is only present in the adversity unexposed MD subtype (reflecting etiologic heterogeneity), and
   c) h^2_{SNP} estimates under the presence and absence of etiologic heterogeneity by replacing the single causal SNP with polygenic contributions in models a) and b) above.

For each simulation, SNP effects were then tested under four logistic regression models: I) ignoring the effect of adversity on MD, II) controlling for the effect of adversity on MD by incorporating it as a covariate, III) including an interaction between SNP and adversity, and IV) analyzing adversity exposed and unexposed cohorts separately.

Liability threshold model

We simulated SNP, adversity, and SNP-by-adversity effects using a liability threshold model (equation 1): a single causal SNP (g), an additive effect of adversity (s), and a SNP-by-adversity interaction ($g * s$, using * for element-wise multiplication):

(1) $$y_i = g_i \beta_g + s_i \beta_s + (g_i * s_i) \beta_{gs} + \epsilon_i$$

Above, y_i is the latent MD liability for individual i; g_i is their genotype at the causal SNP drawn independent and identically distributed (i.i.d.) from a Binomial distribution (Bin(2,.3)) with minor allele frequency (MAF) of 0.30; s_i is a 0-1 variable indicating whether individual i was adversity exposed, drawn i.i.d. from a Bernoulli distribution (Ber(.23)) with mean similar to the proportion of the CONVERGE samples endorsing extreme adversity; and e_i is i.i.d. Gaussian noise with variance 1.

The regression coefficients (β) define the contribution of each model term and are used to modify equation 1 to reflect simulation scenarios a) and b) such that when $\beta_{gs} = 0$ there is no SNP-by-adversity interaction modeled; when $\beta_g = 0$ but $\beta_{gs} \neq 0$ the SNP acts only on the adversity exposed individuals; and when $\beta_g = -\beta_{gs}$ the SNP acts only on the unexposed individuals.

Ascertainment of MD cases and adversity

To simulate ascertainment of MD cases and adversity exposure, we generate genotypes, adversity variables, and MD liabilities for one million samples (the population), define individuals with the largest 5% of liabilities to be MD cases (similar to 8% prevalence we assume for severe recurrent MD in Han Chinese women), then draw 5100 cases and 5100 controls from their respective populations. Specifically, the adversity effect size and prevalence were chosen so that after ascertainment the average fraction of adversity among cases (35.3%) and controls (22.3%) approximately matched the observed proportions in CONVERGE (34.4% and 20.4%, respectively). For scenarios a) and b), we simulated 1,000 independent replicate datasets with a single causal SNP effect. For scenario c), the single SNP was replaced with a polygenic contribution from 10,000 simulated, independent SNPs (below).

Simulations of a single SNP effect under the absence of genetic heterogeneity

As detailed above, 1,000 independent replicates of a SNP effect were used to generate average test statistics from Equation 1 under four models (I - only SNP effect, II - SNP effect plus adversity covariate, III - model II plus SNP-by-adversity interaction term, IV - analysing effect of SNP in adversity exposed and unexposed cohorts separately). The genetic effect size was chosen so the logistic regression results from the simulated data (Table 3) were comparable to the observed CONVERGE results in Tables 1 and 2 (i.e., chromosome 10 loci). Specifically, to reflect the absence of heterogeneity, the SNP, adversity, and GxE regression coefficients in equation 1 were set to 0.078, 0.30, and 0, respectively. The average test statistics from the logistic regression models (I-IV) are displayed on the left panel of Table 3.

Simulations of a single SNP effect under the presence of genetic heterogeneity

To model etiologic heterogeneity, the baseline simulation was modified by setting $\beta_g = 0.0975 = -\beta_{gs}$, so that the SNP acts only on the unexposed individuals. We multiply β_g by 1.25 in all simulations where the genotype acts only in the unexposed cohort to compensate for

the reduced sample size and to obtain qualitatively similar *z*-scores. The results from this simulation are displayed in the right panel of Table 3.

To complement the etiologic heterogeneity simulation above, where the SNP was causal only in the unexposed cohort, we also reversed the simulation, where the SNP was only causal in the exposed cohort (Table S5). To do this, we set $\beta_g = 0.156$ and $\beta_{gs} = 0$.Results did not qualitatively change, though all tests had lower power, which is expected since the adversity exposed sample size is smaller than the unexposed group.

We also considered a variety of modifications to this baseline simulation, none of which affected the qualitative conclusions in the main text: we modified the population MD prevalence from 5% to 1% and 20%; the MAF from 30% to 5% and 50%; the Gaussian noise in the liability simulation to logistic noise; and finally scenarios where case/control labels are simulated directly from a generalized linear model using either the probit or logit link functions. Full simulation details for these extensions are available upon request.

13.   Simulations of polygenic effects under the presence and absence of etiologic heterogeneity

Overview

We use simulations similar to those above, replacing the single causal SNP with a polygenic contribution from 10,000 simulated, independent SNPs. We then estimate h^2_{SNP} to determine the expected features when the full cohort and subgroups are simulated to have: 1) equal h^2_{SNP}, 2) varying degree of genetic correlation (ρ = -1 to 1), and 3) unequal h^2_{SNP}. We also compare these h^2_{SNP} estimates to models including the environmental stressor (adversity) as a covariate and estimate the proportion of MD variance due to aggregate GxE interaction. Since estimates can vary under alternate assumptions all estimation was performed using both maximum likelihood (ML) and the Haseman-Elston regression (HE) (19,20). The resulting h^2_{SNP} estimates from the overall cohort with and without accounting for adversity exposure are shown in Figure S8. The two within-cohort measures (exposed versus unexposed), along with genetic correlation and GxE estimates, are shown in Figure S9. The results confirm our prior intuition: without heterogeneity (where all causal SNPs are shared), within-group average heritabilities coincide with the overall average heritability, though the attenuated sample sizes induce larger variance in the within-group estimators; however, as heterogeneity increases (or causal variant sharing decreases), the overall heritability decreases while the within-cohort heritabilities remain constant. On a methodological note, ML gave downward biased estimates compared to HE, with the exception of genetic correlation estimates (Figure S8, Figure S9), as expected (19).

Generation of major depression liability from a standard polygenic linear mixed model

We modify our above single-SNP simulations to be polygenic by using 10,000 causal SNPs. Unlike in the above simulation, however, the causal effect sizes will be determined randomly, and independently, for each of 200 simulated datasets. We generate liabilities within each adversity exposed and unexposed cohort from a standard linear model:

$$(2) \qquad y_i' = \frac{1}{\sqrt{L}} G_i \beta^0 + \epsilon_i \text{ if } s_i = 0$$

$$(3) \qquad y_i' = \frac{1}{\sqrt{L}} G_i \beta^1 + .32 + \epsilon_i \text{ if } s_i = 1$$

We again use i.i.d. Bernoulli draws for entries of adversity status (s), now with mean *.24*; the noise term is distributed normally ($e \sim N(0,.57I)$). Each column of G, the vector of genotypes at a specific SNP, is drawn i.i.d. Binomial (Bin(2,q)), where q is drawn i.i.d. for each locus from a uniform distribution (Unif[0.05,0.5]).

We marginally draw the causal SNP effects in each β^s from a spherical Gaussian distribution. Without etiologic heterogeneity, $\beta^0 = \beta^1$. Complete heterogeneity, on the other hand, means β^0 and β^1 are entirely independent. To bridge these extremes, we define a set S of SNPs where the β agree, i.e. $\beta_S^0 = \beta_S^1$. Outside of this set S, each SNP is causal in only one adversity cohort. $|S| = 0$ gives complete heterogeneity; $|S| = L$ means all SNPs (and thus all genetic liability) are shared between the two adversity cohorts. Formally, let S be the set of common SNPs and S_0 and S_1 the SNPs unique to the adversity unexposed and exposed cohorts respectively. We take $|S_0| = |S_1| = ((L-|S|)/2)$. We then define β by:

$$\beta_S^0 = \beta_S^1 \sim \mathcal{N}\left(0, \frac{|S|}{(L+|S|)/2} \sigma_g^2 I_{|S|}\right)$$

$$\beta_{S_0}^0 \sim \mathcal{N}\left(0, \frac{|S|}{(L+|S|)/2} \sigma_g^2 I_{|S_0|}\right)$$

$$\beta_{S_0}^1 = 0$$

$$\beta_{S_1}^1 \sim \mathcal{N}\left(0, \frac{|S|}{(L+|S|)/2} \sigma_g^2 I_{|S_1|}\right)$$

$$\beta_{S_1}^0 = 1$$

These definitions ensure the overall genetic variance (σ_g^2) is 0.35.

After defining the liabilities y' from this linear mixed model, we define samples with the largest 15% of liabilities to be cases, and then draw from this latent population of size 1,000,000 the same number of cases and controls as in the observed CONVERGE data (4,745 and 4,830, respectively). Finally, we drop rows of the simulated genotype matrix G corresponding to samples that were not ascertained, scale its columns (SNPs) to zero mean and unit variance, and then create kinship matrices by:

$$(4) \qquad K(G) := \frac{1}{L} GG^T$$

Alternative parameterization of genetic correlation

We now consider an alternative way to define the effect sizes in β. Above, some fraction of loci had identical effects in both adversity cohorts, while others were active in only one ('Different Causal SNPs'). Instead, we can set all loci to affect both cohorts, but with different, yet

correlated, effect sizes. Independently at each locus l, we draw the two effect sizes from a Normal distribution:

$$\begin{pmatrix} \beta_l^0 \\ \beta_l^1 \end{pmatrix} \sim \mathcal{N}\left(0, \begin{pmatrix} \sigma_{g0}^2 & \rho\sigma_{g0}\sigma_{g1} \\ \rho\sigma_{g0}\sigma_{g1} & \sigma_{g1}^2 \end{pmatrix}\right) \quad (5)$$

When $\sigma_{g0}^2 \neq \sigma_{g1}^2$, the two cohorts have different heritabilities: this may well happen in the real data if, for example, predisposition to develop depression is highly heritable, but after exposure to stress this predisposition becomes essentially irrelevant. The parameter rho (ρ) plays a role similar to the fraction of shared causal SNPs in the original simulation above because in both cases, a value of zero results in completely independent genetic architectures between the adversity exposed and unexposed cohorts, and a value of one indicates the genetic architectures are completely overlapping. One difference is that ρ can be negative.

Results from this simulation are displayed in Figures S8 and S9, when $\sigma_{g0}^2 = \sigma_{g1}^2$ ('Equal Heritability'), $\sigma_{g0}^2 = 4\delta\sigma_{g1}^2$ ('Unequal Heritability'), along with results using the above simulation ('Different Causal SNPs') where SNP effects are equal between-group but only a portion of effects are shared. Because these 'Equal Heritability' and 'Different Causal SNPs' simulations parameterize identical models, results are identical – the only difference is that only in the former scenario is $\rho < 0$ feasible.

SNP-based heritability estimation

After simulating case/control phenotypes y, genotypes G, and adversity indicators s, we fit a variety of heritability estimates. Each is defined by a kinship matrix and phenotype vector:
- h^2: overall heritability using all phenotypes and all genotypes to create the kinship matrix
- h^2_s: overall heritability accounting for stress, modifying the above phenotypes and kinship by projecting out the adversity covariate s; e.g., for the phenotype, if s is centered and scaled, this amounts to replacing y with $y - s^T y$
- $h^2 s = 0$: heritability using only adversity unexposed samples
- $h^2 s = 1$: heritability using only adversity exposed samples

Each of these parameters is estimated both with ML and HE regression. Assume, for notational simplicity, that the case/control labels y have been standardized to mean zero and unit variance. ML fits heritability by maximizing a (misspecified) Gaussian likelihood for y:

$$(6) \qquad \hat{h}_{ML}^2 = \arg\min_{h^2} \ \log|h^2 K + (1 - h^2)I| + y^T \left(h^2 K + (1 - h^2)I\right)^{-1} y$$

The optimization is performed with the R package phenix (21), which uses well-known eigendecomposition identities to expedite inference (22).

This likelihood function is motivated by the Gaussian linear model used to create the above liabilities y'. However, the likelihood is clearly misspecified for two reasons: first, the y variables are binary, not continuous (because of liability thresholding); second, ascertainment is

not accounted for. One proposal to correct these two problems is to simply rescale the heritability estimates (23):

$$(7) \qquad \tilde{h}^2 := \frac{K^2(1-K)^2}{P(1-P)\varphi(\Phi^{-1}(K))^2}\hat{h}^2$$

where P and K are the sample and population disease prevalences, respectively, and ϕ and φ are the standard normal density and distribution functions, respectively. We only ever report these rescaled estimates, $\tilde{h}^2$.

Even after this correction, however, MLEs remain downwardly biased (19). So we also implement HE regression to estimate heritability, which simply regresses phenotypic similarity on genetic similarity: if $\tilde{x}$ and $\tilde{y}$ contain, respectively the upper triangular entries of K and yy^T:

$$(8) \qquad \hat{h}^2_{HE} = \frac{1}{\|\tilde{x}\|^2}\tilde{x}^T\tilde{y}$$

For each of the four heritabilities defined above, both ML and HE estimates–adjusted for non-normality and ascertainment–were obtained from between 50 and 300 independently simulated datasets for each (simulation type, estimation type, ρ) combination. Figure S8 plots the resulting averages ($\pm$ 2 empirical standard deviations) for the two estimates that aggregate groups (h^2 and h^2_s), while Figure S9 plots results from analyses that distinguish the two groups ($h^2_{s=0}$ and $h^2_{s=1}$). The observed CONVERGE data estimates are also provided for reference, though their error bars represent standard errors, rather than standard deviations.

Broadly, the overall heritability increases linearly as the causal SNP overlap increases (this linearity is provable for HE regression); however, only when the causal SNPs perfectly overlap is the whole-cohort heritability as large as the within adversity exposure groups heritabilities. Further, the downward bias of ML is evident for all causal overlaps and types of heritabilities.

We also estimate genetic correlation ρ (15) and GxE variance (σ^2_{GxE}) (11). These estimates can also be used to discern between-group heterogeneity and are included as additional rows of Figure S9. The ρ is unbiasedly estimated by both ML and HE and the methods give similar standard deviations (with the possible exception that HE behaves strangely for extreme values of ρ and unequal variance, which is at least partially because we did not truncate marginal HE heritabilities to [0, 1]). ML estimates of σ^2_{GxE}, however, appear to be downwardly biased.

Supplemental Tables

TABLE S1. Prevalence of 16 stressful life events and odds ratio for major depression.

This table shows the number of samples with self-reported answers for each life event item (*n* Total), the number of which are cases of MD (*n* Case), the number of which are controls (*n* Cont.), and the percentage of samples that replied "Yes" to the life event question are shown below (%). The odds ratio (OR) with 95% confidence interval [95%CI] of each life event in MD from logistic regression are displayed with corresponding *P*-value.

Life Event Question	*n* Total	*n* Case (%)	*n* Con. (%)	OR [95%CI]	*P*-value
1. Have you ever had a spouse, child, or sibling die?	11848 18.2%	1178 19.8%	988 16.5%	1.62 [1.46,1.78]	1.07×10^{-21}
2. Have you ever been divorced or long-term marital separation?	11921 13.0%	1091 18.3%	462 7.7%	2.51 [2.23,2.83]	6.48×10^{-53}
3. Have you ever been unemployed or seeking work for more than a month?	11922 13.4%	940 15.8%	657 11.0%	1.28 [1.14,1.43]	1.36×10^{-5}
4. Have you ever been fired from a job?	11923 5.3%	440 7.4%	187 3.1%	2.03 [1.70,2.43]	1.12×10^{-14}
5. Have you ever had a <u>major</u> financial crisis?	11923 16.4%	1166 19.6%	784 13.1%	1.58 [1.43,1.75]	8.05×10^{-19}
6. Have you ever had problems with the police or go to court?	11923 2.4%	223 3.7%	64 1.1%	3.38 [2.55,4.53]	6.51×10^{-17}
7. Have you ever had a <u>serious</u> illness?	11293 9.8%	712 12.0%	460 7.7%	1.82 [1.60,2.07]	1.98×10^{-20}
8. Have you ever been involved in a life-threatening accident?	11304 7.6%	479 8.5%	379 6.7%	1.25 [1.08,1.45]	0.002
9. Have you ever been involved in a fire, flood, or natural disaster?	11305 11.0%	614 10.9%	633 11.1%	1.01 [0.90,1.14]	0.830
10. Have you ever witnessed someone being badly injured/killed?	11305 8.0%	476 8.5%	431 7.6%	1.10 [0.95,1.26]	0.189
11. Have you ever been raped?	11305 0.99%	108 1.9%	4 0.1%	21.35 [8.91,69.92]	2.16×10^{-9}
12. Have you ever been physically attacked or assaulted?	11305 5.1%	411 7.3%	160 2.8%	2.41 [1.99,2.92]	2.36×10^{-19}
13. Were you ever physically abused as a child?	11305 2.8%	259 4.6%	53 0.9%	4.63 [3.45,6.33]	4.19×10^{-23}
14. Were you ever seriously neglected as a child?	11305 6.3%	597 10.6%	118 2.1%	4.82 [3.94,5.94]	2.47×10^{-51}
15. Were you ever threatened with a weapon, held captive or kidnapped?	11305 1.0%	87 1.5%	23 0.4%	3.01 [1.92,4.92]	4.10×10^{-5}
16. Have you had any other terrible experience in your lifetime?	11929 6.2%	472 7.9%	264 4.4%	1.68 [1.43,1.97]	1.66×10^{-10}

TABLE S2. Association of clinical features with adversity exposure.

This table shows the association of clinical characteristics with adversity exposure separately for major depression (MD) cases and controls. The mean value or proportion (Prop.) of the sample endorsing each clinical feature is listed for adversity unexposed (unexp.) and exposed (exp.) subgroups. The corresponding odds ratio (OR), 95% confidence interval (95%CI), *P*-value, and proportion of variance (R2) due to each clinical characteristic from linear or logistic regression is displayed. BMI is body mass index, Age-of-onset is age of MD onset, and GAD is generalized anxiety disorder.

Clinical Feature	MD Cases					Controls				
	Mean/Prop. (unexp./exp.)	OR	95%CI	P-Value	R2	Mean/Prop. (unexp./exp.)	OR	95%CI	P-Value	R2
Age	45.33/42.21	0.961	0.955 0.968	8.87E-33	0.0373	47.69/47.54	0.995	0.983 1.007	0.4248	0.0002
BMI	22.87/22.42	0.96	0.944 0.977	3.89E-06	0.0056	23.25/23.48	1.026	0.998 1.055	0.0674	0.0014
Obesity	0.066/0.065	0.974	0.773 1.222	0.8231	1.29E-05	0.060/0.080	1.356	0.979 1.848	0.0596	0.0015
Education (ref Primary School)	0.747/0.253	-	-	-	0.0189	0.811/0.189	-	-	-	0.0007
Middle School	0.668/0.332	1.467	1.251 1.724	2.72E-06	-	0.831/0.169	0.868	0.692 1.095	0.2276	-
Technical/Junior College	0.597/0.403	1.993	1.667 2.386	4.75E-14	-	0.822/0.178	0.928	0.706 1.220	0.5909	-
Bachelor Degree	0.593/0.407	2.028	1.656 2.484	7.78E-12	-	0.817/0.183	0.962	0.684 1.342	0.8206	-
Marital Status (ref Married)	0.697/0.303	-	-	-	0.0542	0.820/0.180	-	-	-	0.0542
Separated	0.359/0.641	4.106	2.748 6.226	1.1E-11	-	0.469/0.531	5.171	2.568 10.519	3.96E-06	-
Divorced	0.392/0.608	3.576	2.919 4.391	1.97E-34	-	0.403/0.597	6.751	4.984 9.193	1.7E-34	-
Widowed	0.676/0.324	1.102	0.791 1.517	0.559166	-	0.635/0.365	2.62	1.873 3.634	1.15E-08	-
Never Married	0.500 0.500	2.302	1.703 3.110	5.5E-08	-	0.844/0.156	0.845	0.286 2.021	0.730154	-
Neuroticism	11.73/14.62	1.091	1.079 1.102	1.68E-60	0.0736	3.12/4.65	1.095	1.078 1.112	8.32E-30	0.0358
Age-of-onset	35.6/33.2	0.975	0.970 0.98	4.11E-17	0.0186	-	-	-	-	-
Episode Duration (weeks)	43.5/51.6	1.001	1.000 1.002	0.0011	0.0027	-	-	-	-	-
DSM Sx Count	8.6/8.7	1.259	1.143 1.390	4E-06	0.0067	-	-	-	-	-
Number Episodes	3.7/4.2	1.036	1.021 1.052	3.83E-06	0.0054	-	-	-	-	-
Suicidal Thoughts	0.594/0.635	1.188	1.057 1.335	0.0039	0.0022	-	-	-	-	-
Suicidal Plan	0.409/0.521	1.571	1.401 1.761	9.55E-15	0.0156	-	-	-	-	-
Suicidal Attempt	0.472/0.520	1.213	1.029 1.429	0.0213	0.003	-	-	-	-	-
Panic	0.054/0.086	1.666	1.337 2.073	4.92E-06	0.0053	-	-	-	-	-
GAD	0.217/0.307	1.593	1.402 1.810	8.11E-13	0.013	-	-	-	-	-
Dysthymia	0.067/0.152	2.493	2.075 2.996	1.74E-22	0.0242	-	-	-	-	-
Phobia	0.319/0.505	2.178	1.940 2.445	1.15E-39	0.0443	-	-	-	-	-

TABLE S3. Tests for gene-by-environment interaction between adversity and genetic variants on the additive scale.

This table shows the regression coefficient (Est.), standard error (SE), and *P*-value of SNP association with MD in the full cohort in linear regression model with a binomial link (Methods), with an interaction term (on the additive scale) between SNP and self-reported adversity (adversity:SNP) term included in Model 1, and without it in Model 2. All analyses performed included 10 principal components as covariates, bold font indicates significant genetic effect ($P < 5.0 \times 10^{-8}$) or gene-by-environment interaction ($P < 0.005$).

Test	Model 1: Interaction			Model 2: Covariate		
	Est.	**SE**	***P*-value**	**Est.**	**SE**	***P*-value**
chr1: rs7526682_G	0.065	0.014	7.93×10^{-6}	0.043	0.012	0.0004
adversity	0.195	0.015	2.81×10^{-39}	0.174	0.013	1.89×10^{-40}
adversity:rs7526682	-0.079	0.028	**0.0045**	-	-	-
chr1: rs11577545_T	0.060	0.012	1.28×10^{-6}	0.030	0.010	0.0032
adversity	0.215	0.016	3.39×10^{-40}	0.172	0.013	1.72×10^{-39}
adversity:rs11577545	-0.097	0.023	**2.01×10^{-5}**	-	-	-
chr8: rs950893_G	-0.055	0.011	3.90×10^{-7}	-0.037	0.009	6.72×10^{-5}
adversity	0.136	0.017	3.66×10^{-15}	0.172	0.013	1.33×10^{-39}
adversity:rs950893	0.067	0.020	**0.0010**	-	-	-
chr10: rs12415800_A	0.040	0.010	6.45×10^{-5}	0.036	0.008	1.21×10^{-5}
adversity	0.183	0.021	7.45×10^{-18}	0.173	0.013	2.18×10^{-40}
adversity:rs12415800	-0.011	0.018	0.5476	-	-	-
chr10: rs35936514_T	-0.040	0.011	0.0003	-0.044	0.009	5.48×10^{-6}
adversity	0.179	0.017	1.06×10^{-26}	0.175	0.013	$1.31 \times^{-40}$
adversity:rs35936514	-0.010	0.021	0.6118	-	-	-

TABLE S4. Meta-analysis association results of major depression in adversity exposed and unexposed groups.

This table shows the results from fixed-effect (FE) and random effect (RE2) meta-analysis in METASOFT for SNPs associated with MD in the subgroup unexposed to adversity, with effect sizes and standard errors from the BOLT-LMM approximation to infinitesimal leave-one-chromosome-out linear mixed model association analysis. The meta-analysis is performed for the SNP effects in adversity exposed and unexposed subgroups: 1) P-values and effect sizes (beta) are shown for fixed effect (FE) analysis; 2) new random effect (RE2) model proposed in Han and Eskin, 2011 (24) mean effect statistic (Mean Effect Stat) equivalent to the fixed effect statistic, heterogeneity effect statistic (Het. Stat) for non-zero between-study variance, and P-values; and 3) results from tests of heterogeneity including Cochran's Q test (Q), Het. Stat tests the same hypothesis as Cochran's Q test, and approaches Q statistic (stat) asymptotically (with increasing number of studies), and I^2 which describes the percentage of total variation across studies that is due to heterogeneity rather than chance (25).

Chr	RSID	1: FE		2: RE2			3: Heterogeneity		
		P-value	beta	*P*-value	Mean Effect Stat	Het. Stat	*Q* *P*-value	*Q* Het. Stat	I²
1	rs7526682	1.70×10^{-5}	0.045	5.37×10^{-6}	18.50	7.00	3.12×10^{-4}	12.99	92.30
1	rs11577545	2.02×10^{-4}	0.032	2.89×10^{-7}	13.82	12.88	9.42×10^{-6}	19.63	94.90
8	rs950893	6.84×10^{-6}	-0.036	1.35×10^{-7}	20.24	7.89	1.82×10^{-4}	14.01	92.86
10	rs12415800	7.33×10^{-8}	0.035	8.91×10^{-7}	24.52	0	0.388	0.74	0
10	rs35936514	8.04×10^{-7}	-0.039	9.82×10^{-7}	24.35	0	0.643	0.21	0

TABLE S5. SNP-based heritability of major depression estimated using GCTA, LDAK, and PCGC.

This table shows the SNP-based heritability (h^2_{SNP}) of MD in the full cohort and the respective subgroups, estimated with REML in LD-pruned SNPs in GCTA (10) and all association analysis SNPs in LDAK (17), and an extension of Haseman-Elston regression model using LD-pruned SNPs in PCGC, each with 10 PCs derived from eigendecomposition of the GRMs used as covariates. We show the number of cases and controls in each cohort (NCase/NControls), the case to control ratio (NCases:NControls) and the heritability estimates (h^2_{SNP}) with standard errors (SE), and P-values (P) obtained at best-estimate population prevalence (K).

	Full Cohort			Adversity Exposed			Unexposed		
NCase/NControls	4785/4814			1646/982			3139/3832		
NCases:NControls	0.994			1.676			0.819		
Prevalence	$K = 0.08$			$K = 0.128$			$K = 0.066$		
Estimate	h^2_{SNP}	SE	P	h^2_{SNP}	SE	P	h^2_{SNP}	SE	P
GCTA	0.305	0.037	$<10 \times 10^{-16}$	0.342	0.159	0.013	0.380	0.048	1.1×10^{-16}
LDAK	0.294	0.043	6.9×10^{-13}	0.237	0.190	0.105	0.365	0.056	1.4×10^{-11}
PCGC	0.258	0.046	2.2×10^{-8}	0.209	0.195	0.283	0.311	0.060	2.4×10^{-7}

TABLE S6. Impact on major depression of polygenic risk scores and their interaction with adversity.

This table shows the odds ratio (OR) or regression coefficient (Est.), 95% confidence interval (95%CI), *P*-value (P), and the predictive value of each model term reported in terms of Nagelkerke's pseudo-R^2 (fmsb package in R) (R2) for multiplicative and additive models predicting major depression (MD) including polygenic risk scores (PRS), adversity, and their interaction (PRS*Adversity). PRS were constructed by two methods (see Methods): CONVERGE trained scores using BLUP SNP-weights (Conv_Blup_Test, Conv_Blup_Train) (16), and PGC-MDD trained scores (PGC-Rec_pT0.2) using SNP-weights from PGC GWAS of recurrent MD at *P*-value threshold < 0.2 (26). All analyses performed included 10 principal components as covariates, bold font indicates significant effect corrected for 6 tests (*P*-value < 0.008). Results indicate no robust PRS by adversity interactions (*P*-value > 0.044).

PRS	Scale	PRS				Adversity				PRS*Adversity			
		OR/Est.	95%CI	P	R2	OR/Est.	95%CI	P	R2	OR/Est.	95%CI	P	R2
Conv_Blup_Test	Multiplicative	1.233	1.151 1.323	**3.43E-09**	0.0156	2.117	1.858 2.415	**3.56E-29**	0.0353	0.922	0.808 1.052	0.225	0.0004
Conv_Blup_Train	Multiplicative	1.254	1.172 1.342	**6.34E-11**	0.0078	1.95	1.709 2.227	**4.39E-23**	0.0273	0.896	0.784 1.024	0.107	0.0007
PGC-Rec_pT0.2	Multiplicative	1.004	1.000 1.009	0.076	0.0015	6.811	2.091 22.293	**0.001**	0.0313	1.009	1.000 1.019	0.044	0.0006
Conv_Blup_Test	Additive	0.051	0.032 0.071	**1.89E-07**	0.0086	0.183	0.147 0.218	**1.38E-23**	0.0194	-0.023	-0.058 0.013	0.207	0.0003
Conv_Blup_Train	Additive	0.056	0.037 0.074	**4.67E-09**	0.0043	0.164	0.127 0.200	**1.77E-18**	0.0151	-0.028	-0.064 0.008	0.129	0.0005
PGC-Rec_pT0.2	Additive	0.001	0.000 0.002	0.129	0.0008	0.442	0.121 0.763	**0.007**	0.0172	0.002	0.000 0.005	0.100	0.0003

TABLE S7. Association of major depression polygenic risk scores with adversity as a test of G-E correlation.

This table shows the odds ratio (OR), 95% confidence interval (95%CI), *P*-value (P), and the predictive value of each model term reported in terms of Nagelkerke's pseudo-R^2 (fmsb package in R) (R2) for multiplicative models using polygenic risk scores (PRS) to predict adversity status. PRS were constructed by two methods: CONVERGE trained scores using BLUP SNP-weights (Conv_Blup_Test, Conv_Blup_Train) (16), and PGC-MDD trained scores (PGC-Rec_pT0.2) using SNP-weights from PGC GWAS of recurrent major depression (MD) (26) (see Methods). Four models were tested for each of the three PRS: M1 tested the association of the PRS with adversity including MD as a covariate in the entire sample (*n* = 9599), M2 tested the association of the PRS with adversity in the entire sample, M3 tested the association of the PRS with adversity in MD cases only (*n* = 4785), and M4 tested the association of the PRS with adversity in controls only (*n* = 4814). All analyses performed included 10 principal components (PCs) as covariates, bold font indicates significant effect Bonferroni corrected for 12 tests (*P*-value < 0.004). The PRS were not significantly associated with adversity in cases or controls (*P*-value > 0.105) indicating no significant G-E correlation.

Multiplicative Model	PRS				MD			
PRS	OR	95%CI	P	R2	OR	95%CI	P	R2
M1: Conv_Blup_Test (all)	1.000	0.936 1.069	0.991	0.0038	2.118	1.858 2.416	**3.91E-29**	0.0381
M2: (all)	1.036	0.971 1.105	0.288	0.0004	-	-	-	-
M3: (case)	0.967	0.887 1.055	0.453	0.0024	-	-	-	-
M4: (control)	1.049	0.947 1.163	0.357	0.0015	-	-	-	-
M1: Conv_Blup_Train (all)	0.975	0.913 1.041	0.451	0.0032	1.955	1.713 2.233	**3.81E-23**	0.0299
M2: (all)	1.007	0.944 1.075	0.833	0.0008	-	-	-	-
M3: (case)	0.930	0.852 1.015	0.105	0.0063	-	-	-	-
M4: (control)	1.038	0.938 1.148	0.472	0.0022	-	-	-	-
M1 : PGC-Rec_pT0.2 (all)	1.001	0.996 1.005	0.701	0.0001	2.026	1.847 2.223	**1.57E-50**	0.0338
M2: (all)	1.002	0.998 1.007	0.366	0.0007	-	-	-	-
M3: (case)	1.005	0.999 1.011	0.108	0.0007	-	-	-	-
M4: (control)	0.996	0.989 1.002	0.206	0.0005	-	-	-	-

Note:
M1: adversity ~ PCs + PRS + MD (entire sample)
M2: adversity ~ PCs + PRS (entire sample)
M3: adversity ~ PCs + PRS (MD-case only)
M4: adversity ~ PCs + PRS (control only)

TABLE S8. Association of top five major depression variants with adversity in the entire sample, major depression cases only, and controls only.

This table shows the odds ratio (OR), 95% confidence interval (95%CI), *P*-value (P), and the predictive value of each model term reported in terms of Nagelkerke's pseudo-R^2 (fmsb package in R) (R2) for multiplicative models predicting adversity from MD-associated single-nucleotide polymorphisms (SNP). Four models were tested for each of the five SNPs: M1 tested the association of the SNP with adversity including MD as a covariate in the entire sample ($n = 9599$), M2 tested the association of the SNP with adversity in the entire sample, M3 tested the association of the SNP with adversity in MD cases only ($n = 4785$), and M4 tested the association of the SNP with adversity in controls only ($n = 4814$). All analyses performed included 10 principal components (PCs) as covariates, bold font indicates significant effect corrected for 20 tests (*P*-value < 0.0025). As expected, this proof of principle exercise shows that SNP ORs are in opposite directions for MD case-only and control-only analyses for the three SNPs we identified with heterogeneous effects on MD, while we see no differential effects for the two SNPs on chromosome 10.

| Multiplicative Model | SNP | | | | MD | | | |
SNP	OR	95%CI	P	R2	OR	95%CI	P	R2
M1: chr1: rs7526682_G (all)	0.907	0.823 0.998	0.0456	0.0006	2.035	1.855 2.233	4.29E-51	0.0342
M2: (all)	0.934	0.849 1.027	0.1615	0.0003	-	-	-	-
M3: (case)	0.805	0.710 0.911	**0.0006**	0.0034	-	-	-	-
M4: (con)	1.086	0.933 1.260	0.2817	0.0004	-	-	-	-
M1: chr1: rs11577545_T (all)	1.050	0.969 1.137	0.2321	0.0002	2.023	1.844 2.219	2.61E-50	0.0336
M2: (all)	1.074	0.993 1.162	0.0748	0.0005	-	-	-	-
M3: (case)	0.891	0.803 0.989	0.0312	0.0013	-	-	-	-
M4: (con)	1.328	1.174 1.501	**5.60E-06**	0.0065	-	-	-	-
chr8: rs950893_G (all)	0.975	0.906 1.048	0.4898	6.96E-05	2.024	1.845 2.221	2.43E-50	0.0336
M2: (all)	0.949	0.883 1.019	0.1519	0.0003	-	-	-	-
M3: (case)	1.100	1.000 1.210	0.0496	0.0011	-	-	-	-
M4: (con)	0.826	0.737 0.924	**0.0009**	0.0036	-	-	-	-
chr10: rs12415800_A (all)	0.960	0.900 1.025	0.2226	0.0002	2.033	1.854 2.231	6.15E-51	0.0341
M2: (all)	0.986	0.925 1.052	0.6777	2.58E-05	-	-	-	-
M3: (case)	0.942	0.866 1.026	0.1696	0.0005	-	-	-	-
M4: (con)	0.989	0.893 1.095	0.8322	1.45E-05	-	-	-	-
chr10: rs35936514_T (all)	1.048	0.974 1.128	0.2096	0.0002	2.033	1.854 2.231	6.00E-51	0.0341
M2: (all)	1.017	0.945 1.093	0.6570	2.94E-05	-	-	-	-
M3: (case)	1.027	0.930 1.132	0.6018	7.77E-05	-	-	-	-
M4: (con)	1.080	0.965 1.206	0.1783	0.0006	-	-	-	-

Note:
M1: adversity ~ PCs + SNP + MD (entire sample)
M2: adversity ~ PCs + SNP (entire sample)
M3: adversity ~ PCs + SNP (MD-case only)
M4: adversity ~ PCs + SNP (control only)

TABLE S9. Examining G-E independence assumption through exploratory case-only and control-only genome-wide association of adversity.

In order to evaluate G-E independence, we performed exploratory major depression (MD) Case-Only and Control-Only genome-wide association studies (GWAS) of adversity (Figure S8). In table (a) the *P*-value threshold is listed with the corresponding number of SNPs below that threshold for Case-Only GWAS results for adversity (Case-Only # SNPs) and the number of those same SNPs in the Control-Only GWAS for adversity that had a *P*-value less than 0.05. Similarly, for table (b) the *P*-value threshold is listed with the corresponding number of SNPs below that threshold for Control-Only GWAS results for adversity (Control-Only # SNPs) and the number of those same SNPs in the MD-Case-Only GWAS for adversity that had a *P*-value less than 0.05. Given the number of tests performed, departure from G-E independence would be marked by the proportion of Case-Only to Control-Only SNPs (Percent) significantly exceeding 5%. The genome-wide correlation of odds ratios between the MD-Case-Only and Control-Only for adversity was only r = 0.008. These analyses taken together do not support significant systematic G-E correlation.

(a) Top adversity-associated SNPs for MD-Case only GWAS

P-Value Threshold	Case-Only # SNPs	Control-Only # SNPs P <0.05	Percent
0.00001	81	1	1.23
0.0001	492	9	1.83
0.001	6379	353	5.53
0.01	63456	3728	5.87
0.05	311775	15767	5.06

(b) Top adversity-associated SNPs for Control only GWAS

P-Value Threshold	Control-Only # SNPs	Case-Only # SNPs P <0.05	Percent
0.00001	97	1	1.03
0.0001	896	58	6.47
0.001	5569	260	4.67
0.01	62951	3421	5.43
0.05	310758	15767	5.07

TABLE S10. Post-hoc power calculations.

Post-hoc power calculations for single SNP effects were performed using Genetic Association Power Calculator (27) and 'powerGWASinteraction' package in R for GxE interactions (28) using an 8% major depression disease prevalence and 20% adversity prevalence.

Table S10a shows the power (at either P-value $< 5.0 \times 10^{-8}$ or 0.005) to detect each single-nucleotide polymorphism (SNP) with its corresponding chromosome (Chr), minor allele frequency (MAF), and odds ratio (OR) in (1) the combined sample (Combined, $n = 9,599$) which had non-missing information on adversity, (2) the adversity unexposed subset (Unexposed, $n = 6,971$), and (3) the adversity exposed subset (Exposed, $n = 2,628$). For each subset of the CONVERGE data the corresponding number of cases and controls are listed (case/cont) and the final column indicates the expected sample sizes needed for 80% to detect each SNP. These results indicate we had modest to high power to detect the newly identified variants on Chr 1 and 8 in the combined and unexposed samples (power > 0.549) but not in the adversity exposed subsample (power < 0.040) and limited power to detect the previously identified loci on Chr 10 (power < 0.271).

			Sample: case/cont:	Combined 4785/4814 power (5x10-8)	Unexposed 3139/3832 power (5x10-8)	Exposed 1646/982 power (5x10-8)	Exposed 1646/982 power (0.005)	n needed 80% power case/cont
Chr	SNP	MAF	OR					
1	rs7526682	0.13	1.31	0.961	0.736	0.040	0.813	3670/3670
1	rs11577545	0.21	1.25	0.948	0.735	0.033	0.790	3800/3800
8	rs950893	0.28	0.79	0.872	0.549	0.018	0.707	4400/4400
10	rs12415800	0.45	1.16	0.603	0.271	0.005	0.534	5850/5850
10	rs35936514	0.26	0.84	0.335	0.116	0.002	0.400	7550/7550

Table S10b: The CONVERGE sample is under-powered to perform GWASxE interaction scans (power < 0.38), which is why we have not included this type of analysis in the manuscript, we limited GxE testing to 5 SNPs (Bonferroni P-value 0.005). The post-hoc power calculations indicated that we had high power to detect such interactions (power $= 0.956$).

Chr	SNP	MAF	SNP OR	Adversity OR	SNP*Adversity OR	Power $P < 0.005$
1	rs7526682	0.13	1.29	2.21	0.73	0.956
1	rs11577545	0.22	1.28	2.41	0.67	0.991
8	rs950893	0.28	0.8	1.74	1.31	0.999

TABLE S11. Simulation results for single-SNP causal effect only in adversity exposed major depression cases.

Average test output from four types of logistic regression on 1,000 simulated datasets: one ignoring adversity (model I); one controlling for the additive effect of adversity (model II); one additionally incorporating an interaction between genotype and adversity (model III); and finally a model which analyzes adversity exposed and unexposed cohorts separately (model IV). For each row, Z statistic (Z Stat), odds ratio (OR) and *P*-value are shown for SNP effect (g), adversity effect (s), or an interaction effect (g:s). The three columns show results for simulations with heterogeneity between samples with and without adversity; opposite main Table 3, the genotype has a causal effect only in adversity exposed samples (Methods). Data was simulated using a liability threshold model with realistic ascertainment, effect sizes and allele frequencies.

Regression Model	With Genetic Heterogeneity		
	Z Stat	OR	*P*-value
Model I, g	3.86	1.13	1.14E-04
Model II, g	2.93	1.10	3.35E-03
Model II, s	18.44	2.32	<2E-16
Model III, g	-0.01	1.00	9.92E-01
Model III, s	10.20	1.91	<2E-16
Model III, g:s	4.42	1.36	9.99E-06
Model IV, g, no adversity	-0.01	1.00	9.92E-01
Model IV, g, adversity	5.27	1.36	1.39E-07

Supplemental Figures

FIGURE S1. Major depression cases endorse significantly more stressful life events than controls.

This figure shows the proportion of cases of MD and controls for number of stressful life events endorsed. A greater proportion of MD cases report higher levels of stressful life events than controls.
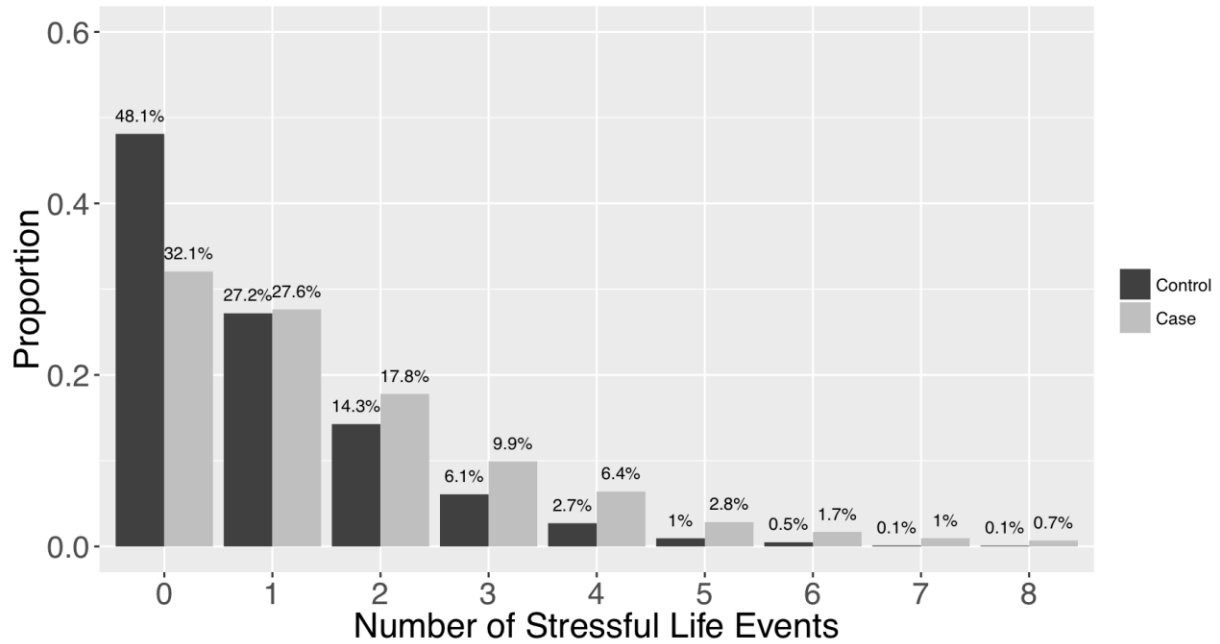
FIGURE S2. Major depression cases have a higher rate of childhood sexual abuse than controls.

This figure shows the proportion of cases of MD and controls reporting childhood sexual abuse (CSA). MD cases report higher levels of each of the types of CSA (NonGenital, Genital, and Intercourse), and for any of the three types of CSA (Any_CSA) than controls.
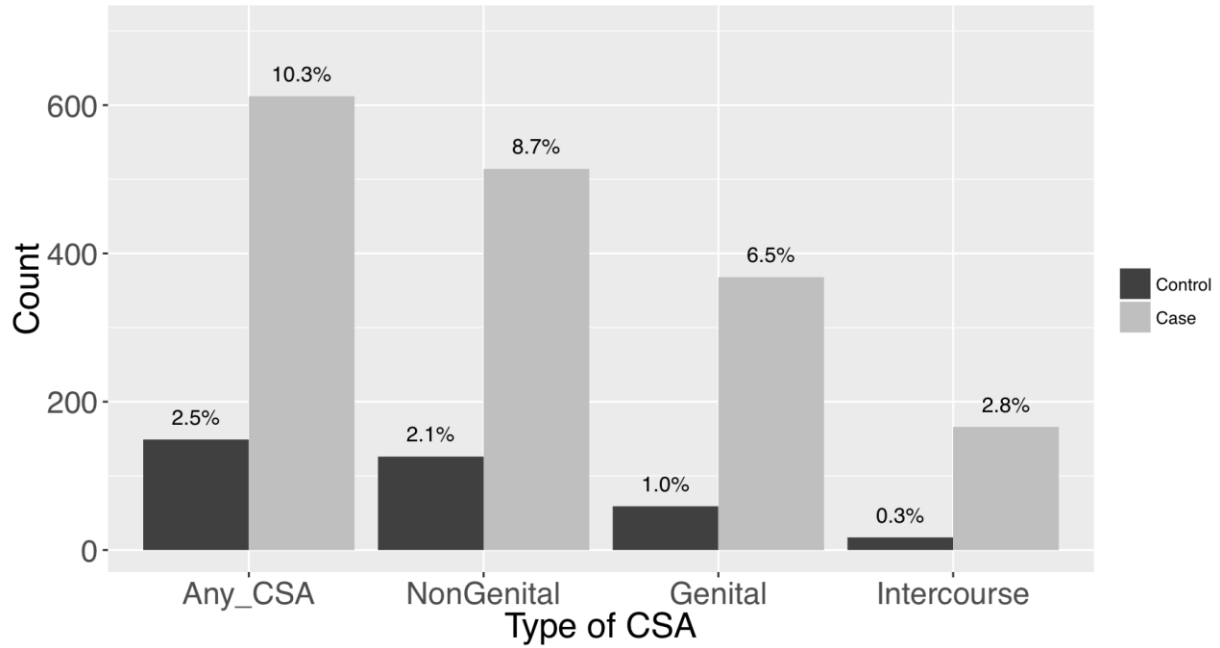
FIGURE S3. Quantile-Quantile plots for genome-wide association studies of major depression.

This figure shows Quantile-Quantile plots for GWAS of MD on a) the full cohort, b) the subgroups with self-reported adversity exposure, and c) without. The genomic control inflation factors were 1.047, 1, and 1.047; the adjusted measure for sample size to that of 1,000 cases and 1,000 controls (λ_{1000}) were 1.01, 1, and 1.014 respectively.
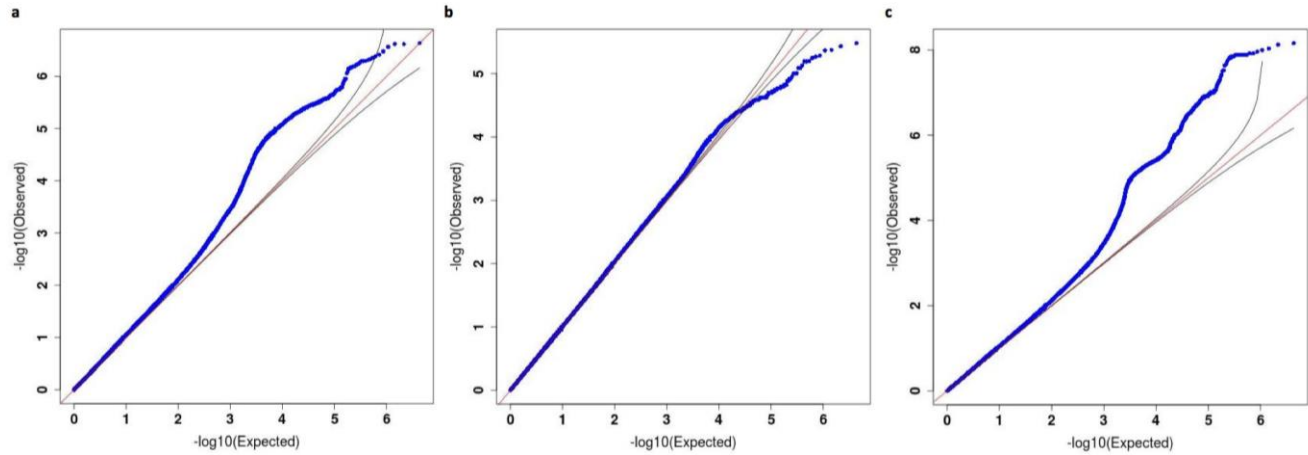
FIGURE S4. Comparison of SNP odds ratios on major depression in full cohort and adversity subgroups as well as results from the Psychiatric Genomics Consortium mega-analysis of European studies.

This figure shows forest plots of comparison of OR on MD in the full cohort and adversity subgroups in CONVERGE with results from the Psychiatric Genomics Consortium (PGC) mega-analysis of European studies (6) at a) rs7526682 on chromosome 1, b) rs11577545 on chromosome 1, and c) rs950893 on chromosome 8. The plots demonstrate all three SNPs have larger OR in the subgroup with no adversity than in the full cohort, OR of respective SNPs on MD are in opposite directions in adversity subgroups, significant association in PGC samples between rs950893 on chromosome 8 and MD ($P = 0.009$) and in the same direction as observed in CONVERGE, and the chromosome 1 loci (rs7526682 and rs11577545) findings were not significant in the PGC study ($P = 0.37$, $P = 0.81$), although were in the same direction as observed in CONVERGE.
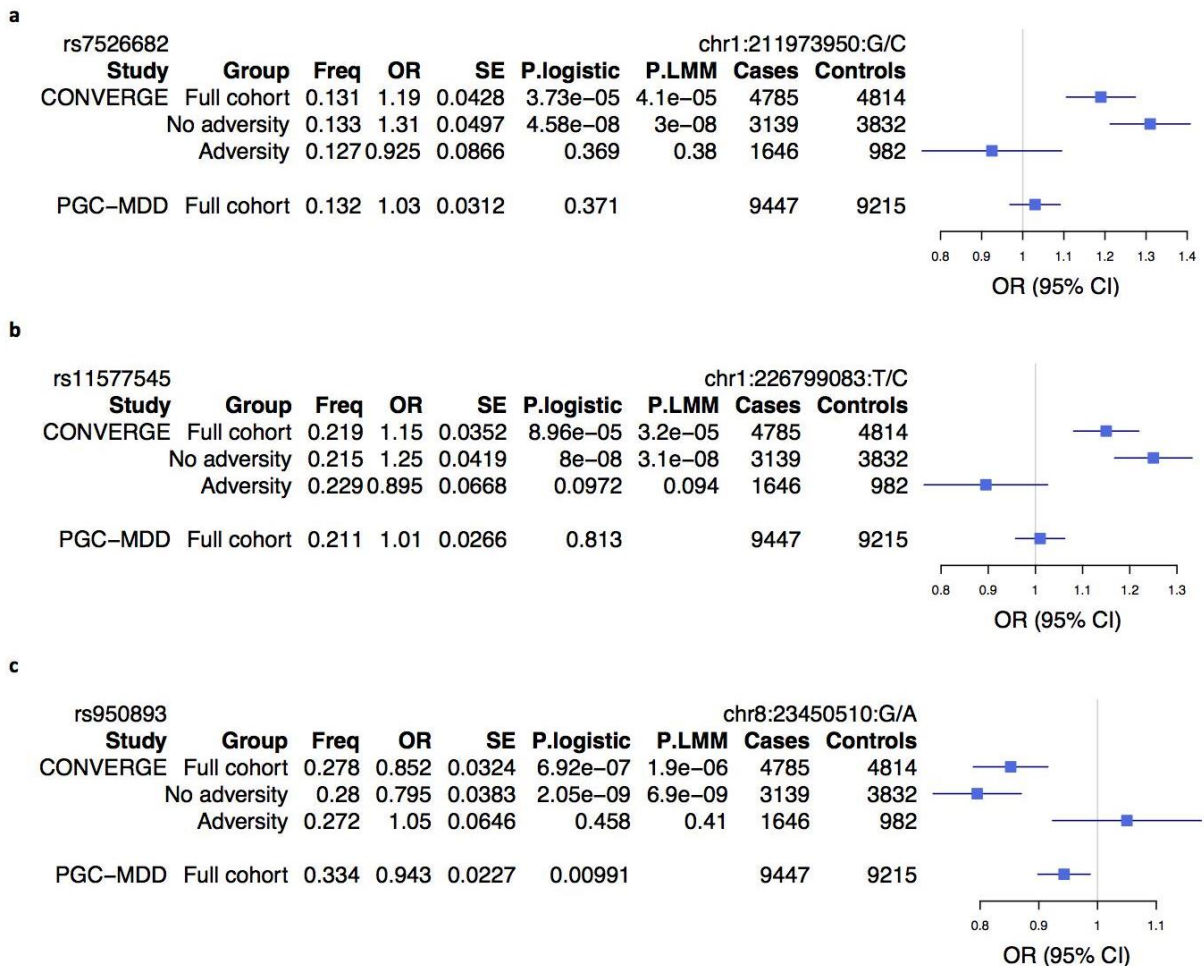
a

rs7526682   chr1:211973950:G/C

Study	Group	Freq	OR	SE	P.logistic	P.LMM	Cases	Controls
CONVERGE	Full cohort	0.131	1.19	0.0428	3.73e−05	4.1e−05	4785	4814
	No adversity	0.133	1.31	0.0497	4.58e−08	3e−08	3139	3832
	Adversity	0.127	0.925	0.0866	0.369	0.38	1646	982
PGC−MDD	Full cohort	0.132	1.03	0.0312	0.371		9447	9215

b

rs11577545   chr1:226799083:T/C

Study	Group	Freq	OR	SE	P.logistic	P.LMM	Cases	Controls
CONVERGE	Full cohort	0.219	1.15	0.0352	8.96e−05	3.2e−05	4785	4814
	No adversity	0.215	1.25	0.0419	8e−08	3.1e−08	3139	3832
	Adversity	0.229	0.895	0.0668	0.0972	0.094	1646	982
PGC−MDD	Full cohort	0.211	1.01	0.0266	0.813		9447	9215

c

rs950893   chr8:23450510:G/A

Study	Group	Freq	OR	SE	P.logistic	P.LMM	Cases	Controls
CONVERGE	Full cohort	0.278	0.852	0.0324	6.92e−07	1.9e−06	4785	4814
	No adversity	0.28	0.795	0.0383	2.05e−09	6.9e−09	3139	3832
	Adversity	0.272	1.05	0.0646	0.458	0.41	1646	982
PGC−MDD	Full cohort	0.334	0.943	0.0227	0.00991		9447	9215

FIGURE S5. Odds ratios of SNPs on chromosomes 1 and 8 significantly associated with major depression in adversity unexposed subgroup are unlikely to arise from stochastic sampling effects.

We obtain empirical distributions of the odds ratios (OR) in subsets of samples of the same size as the adversity unexposed subset, and compared them to the observed values. We generate 10,000 subsets of the cohort by randomly excluding 2,702 samples (1,646 cases, 982 controls) from the full cohort (matching numbers of MD cases and controls with self-reported adversity), and run logistic regression in PLINK v1.9 (7,8) between the three SNPs. Vertical red line on each histogram shows the ORs for the same SNP significantly associated with MD in the unexposed group. a) rs7526682 (99.9th percentile), b) rs11577545 (100th percentile) and c) rs950893 (0.2th percentile) all show significant deviation in OR from the full cohort ($P < 0.01$ after multiple testing correction).
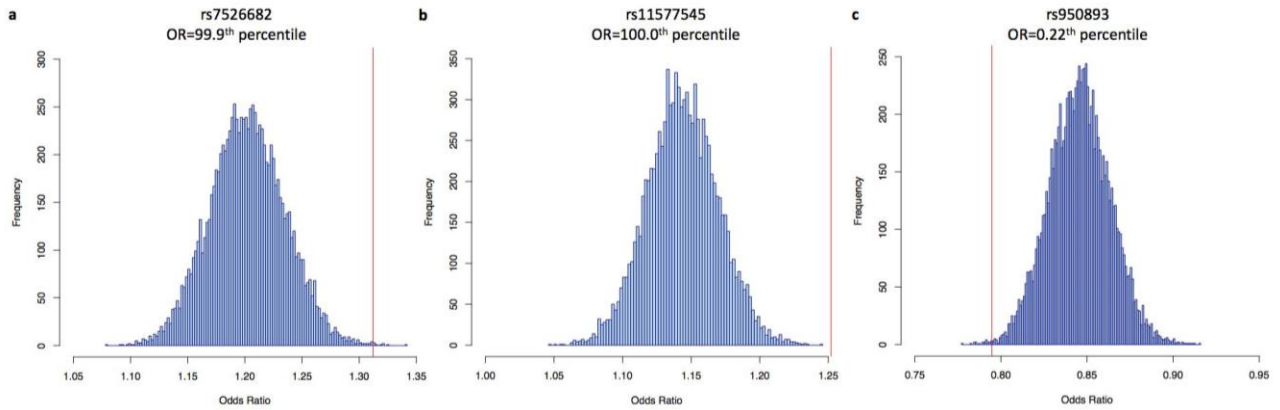
FIGURE S6. Odds ratios of major depression associated loci on chromosome 10 were not significantly different in the adversity unexposed group than from the whole CONVERGE cohort.

This figure shows the same analysis from Figure S5 performed on the two SNPs on chromosome 10 (rs12415800 and rs35936514) that were found to be associated with MD in the whole CONVERGE cohort, for comparison with the three SNPs that were only found to be significantly associated with MD in the subgroup unexposed to adversity. For both SNPs, their odds ratio (OR) for MD were not significantly different in the unexposed group than from the whole-cohort. On the left, OR at rs12415800 for MD is at the 79.9th percentile of the empirical distribution of 10,000 ORs generated from excluding 2,702 samples (1,646 cases, 982 controls) from the full cohort (matching numbers of MD cases and controls with self-reported adversity). On the right, OR at rs35936514 for MD is at the 69.2th percentile of the empirical distribution. Neither show significant deviation in OR from the full cohort.
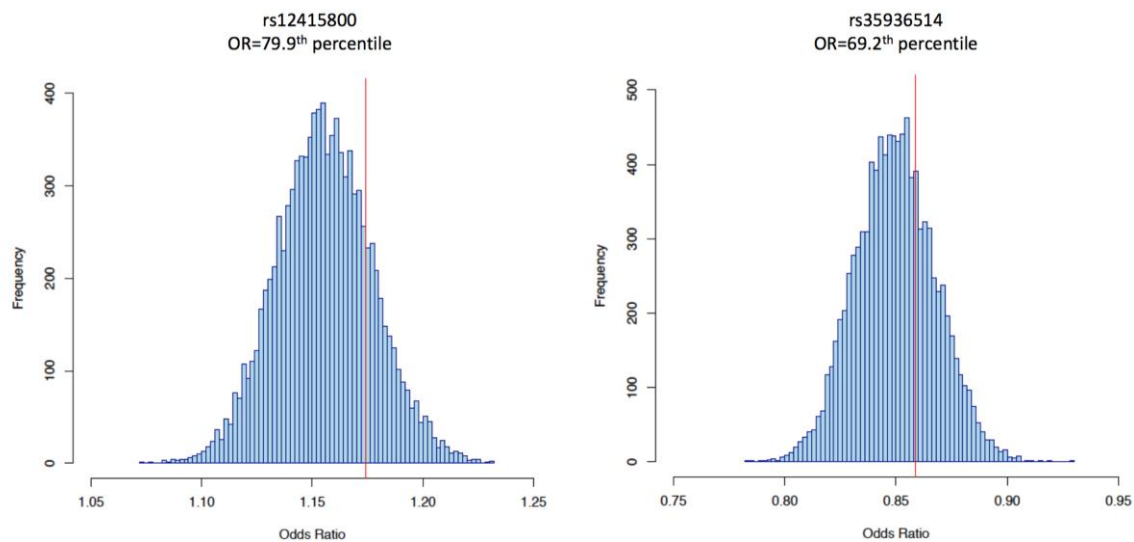
FIGURE S7. The proportion of major depression cases stratified by self-reported adversity exposure for each genetic variant identified depicting gene-by-environment interaction.

The significance for gene by environment interaction (GxE) on the multiplicative scale is denoted as P_{mult} and on the additive scale is P_{add}. Bars represent 95% confidence intervals.

(a) A significant (reversed fan-shaped) interaction detected between rs7526682 on chromosome 1 near *LPGAT1* and adversity (GxE $P_{mult} = 0.0016$, $P_{add} = 0.0045$).
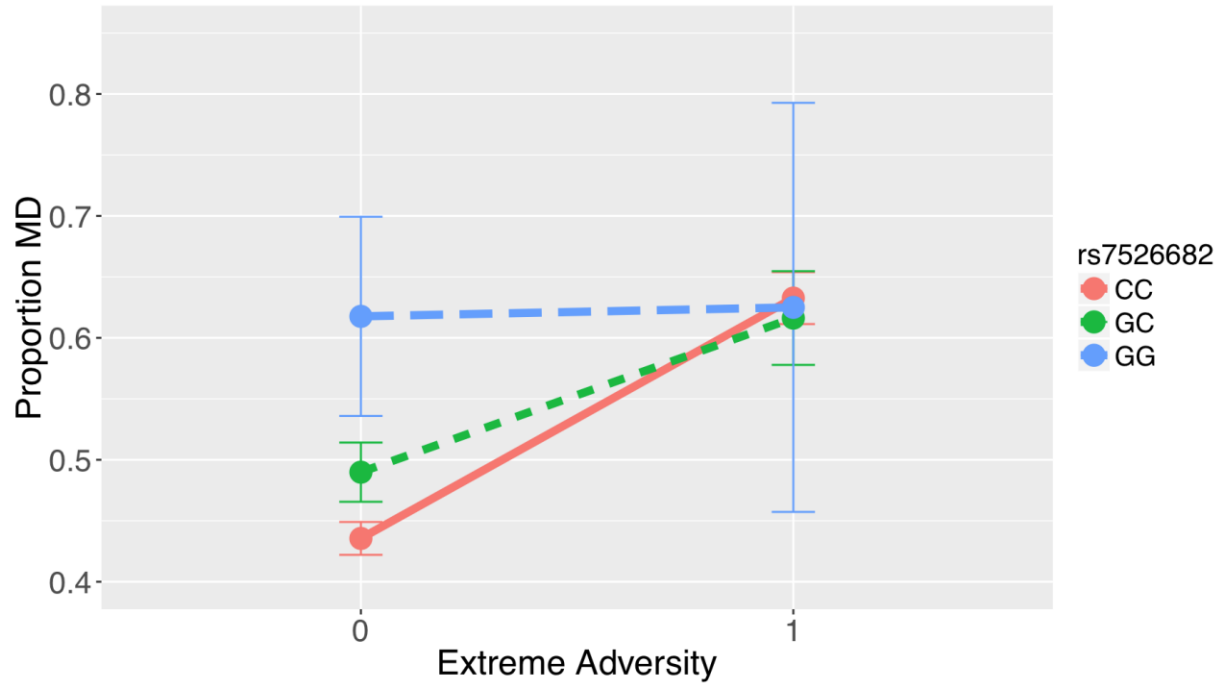
(b) A significant (reversed fan-shaped) interaction detected between rs11577545 on chromosome 1 in *C1ORF95* and adversity (GxE P_{mult} = 9.31x10^{-7}, P_{add} = 2.01x10^{-5}).
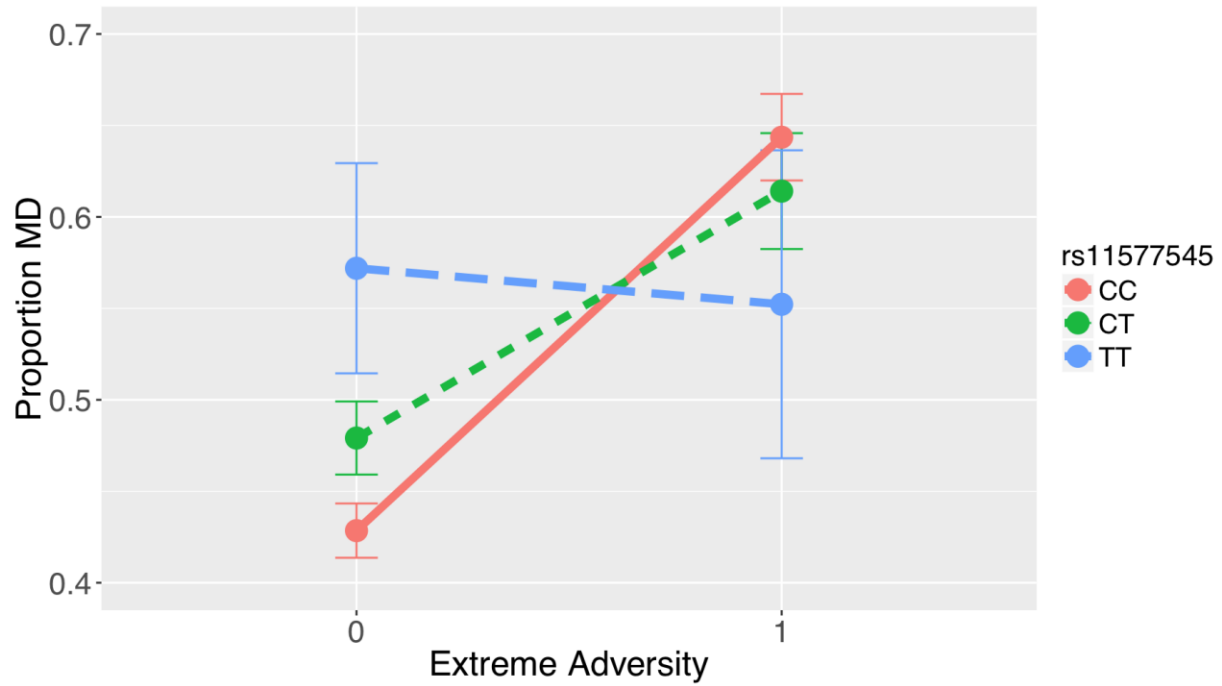
(c) A significant (reversed fan-shaped) interaction detected between rs950893 on chromosome 8 in *SLC25A37* and adversity (GxE P_{mult} = 0.0003, P_{add} = 0.0010).
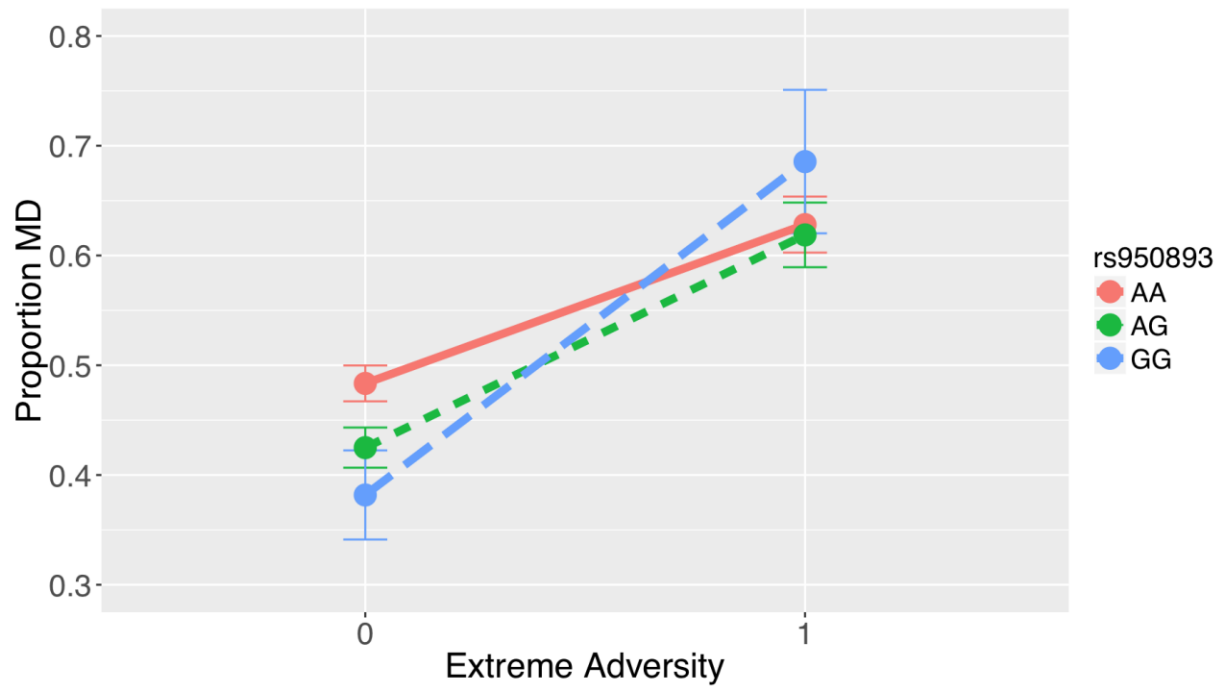
(d) No significant interaction detected between rs35936514 on chromosome 10 in *LHPP* and adversity (GxE P_{mult} = 0.5010, P_{add} = 0.6118).
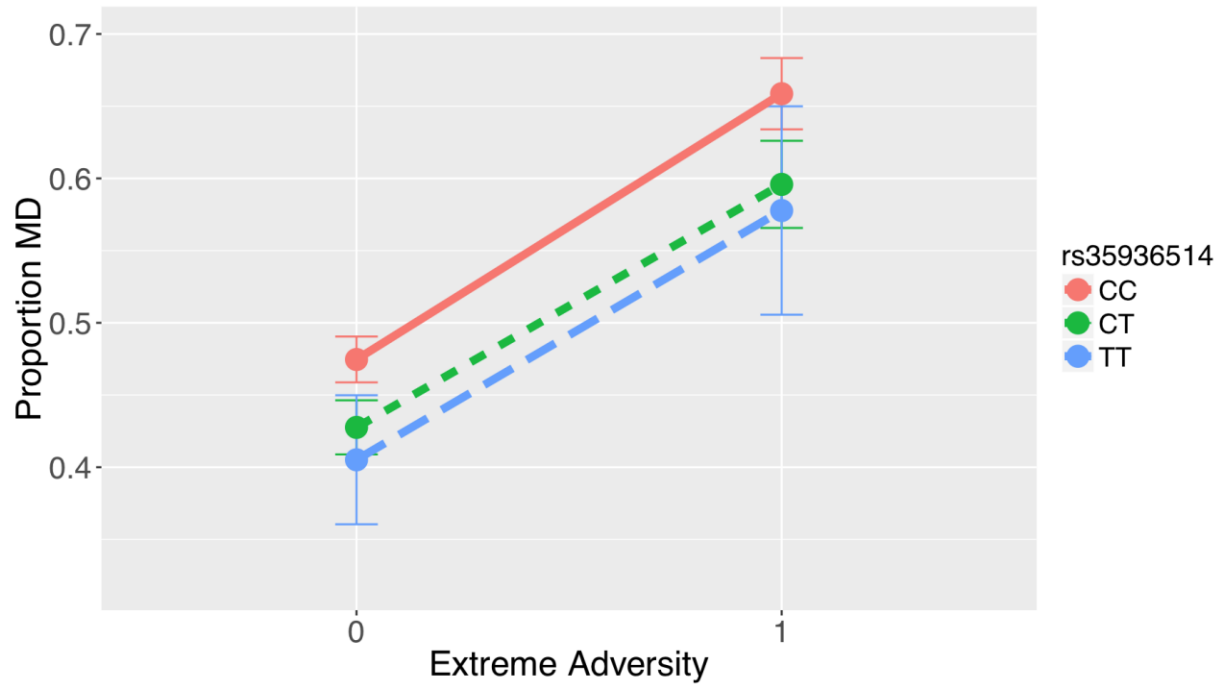
(e) No significant interaction detected between rs12415800 on chromosome 10 near *SIRT1* and adversity (GxE P_{mult} = 0.5630, P_{add} = 0.5476).
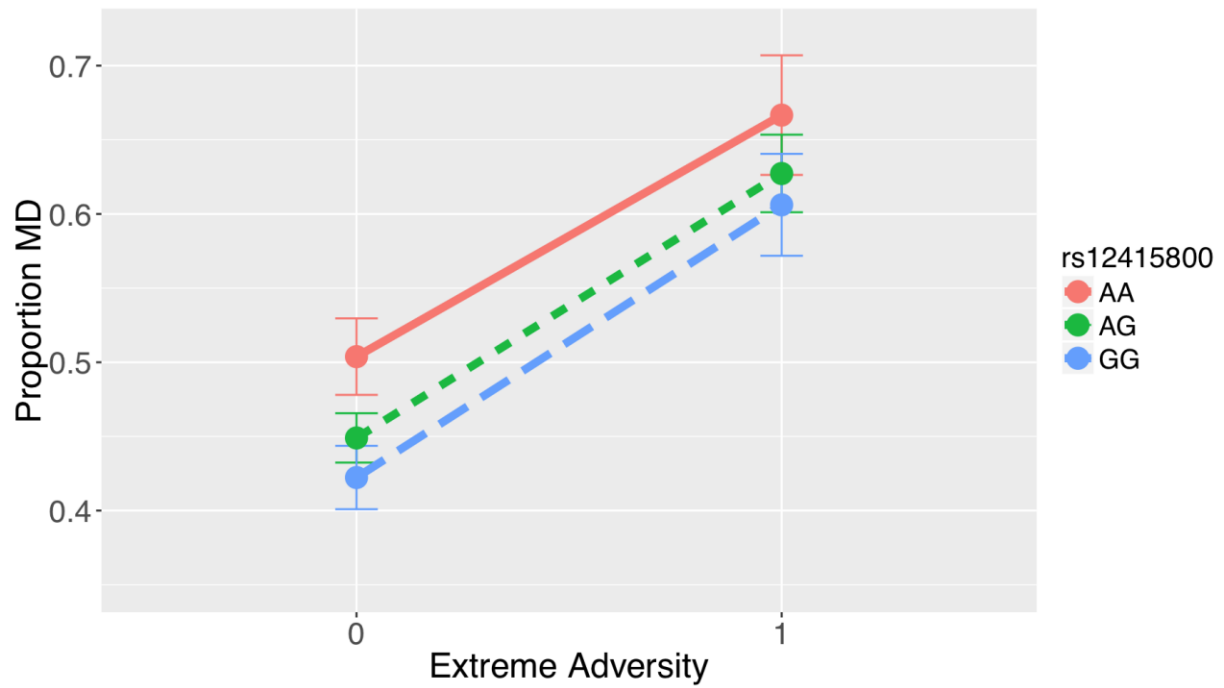
FIGURE S8. Manhattan and Quantile-Quantile plots for exploratory genome-wide association of adversity in major depression cases-only and controls-only.

(a) Manhattan plots of adversity for major depression (MD) cases only (top panel) and controls only (bottom panel). In each plot, the −log10 P-values of imputed SNPs associated with adversity by logistic regression (in order to obtain odds ratios for case-control group comparisons) are shown on the y-axes. The horizontal axis gives the position on each chromosome; chromosomes are numbered below the axis. All analyses included 10 principal components as covariates, red and blue lines indicate genome-wide (5.0×10^{-8}) and suggestive (1.0×10^{-6}) significance respectively.
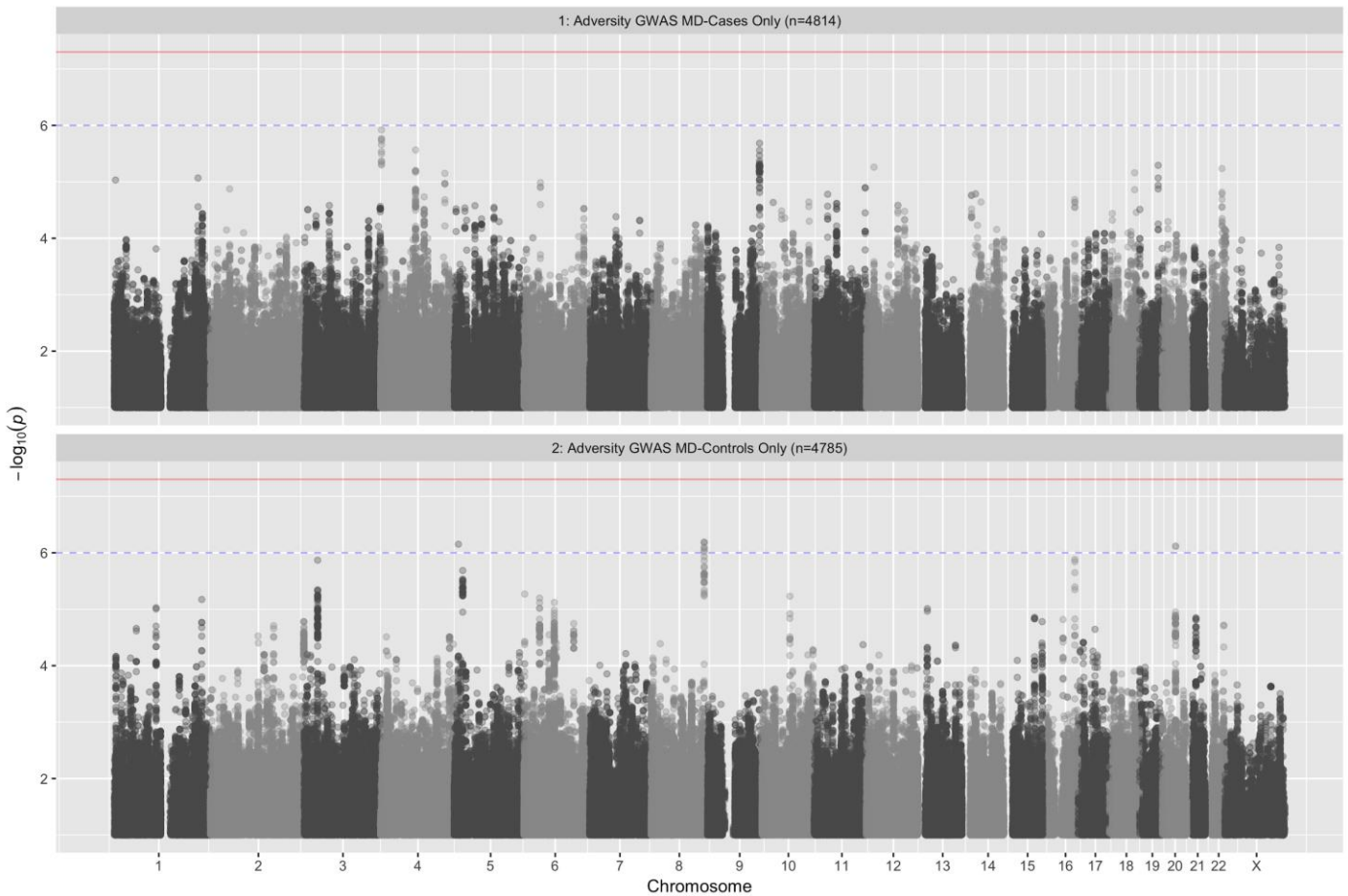
(b) This figure shows Quantile-Quantile plots for genome-wide association of adversity on 1: major depression (MD) cases only (black line) and 2: controls only (red line).
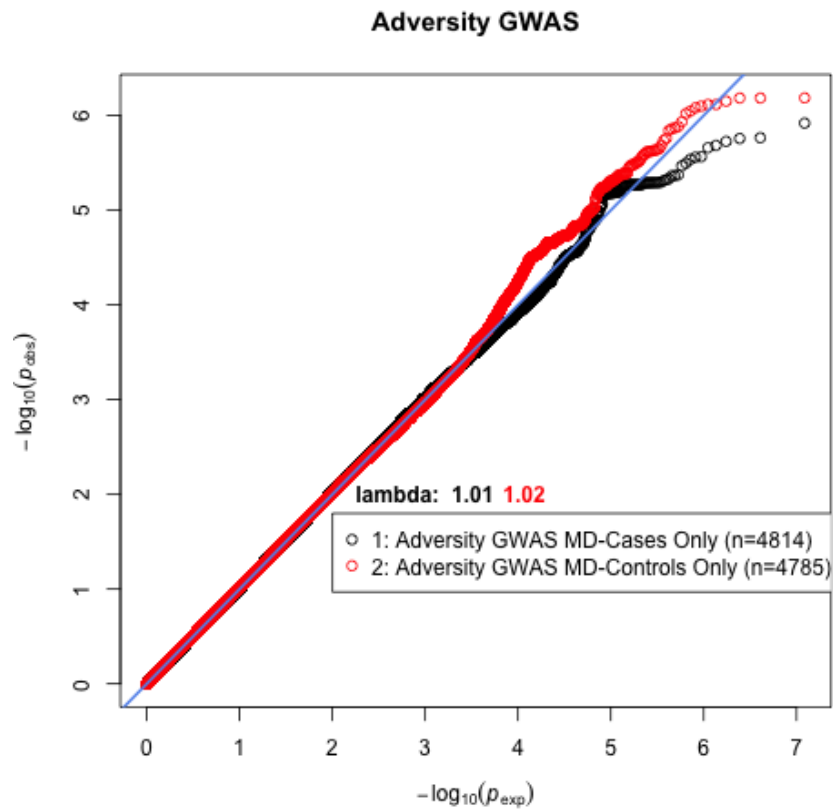

Adversity GWAS

FIGURE S9. Simulation results for overall heritability estimates before and after adjusting for a linear effect of adversity.

Simulation results for overall heritability estimates before (top row) and after (bottom row) adjusting for a linear effect of stress (adversity). Dotted error bars represent ± 2 standard deviations, while the dashed error bars represent ± 2 standard errors. The first two columns use distinct parameterizations of an equivalent model where both stress (adversity) groups are equally heritable and are partially genetically correlated. The third column uses the same genetic correlation parameterization as the first, but assumes the genetic variance in the adversity unexposed group is four times larger than in the exposed group.
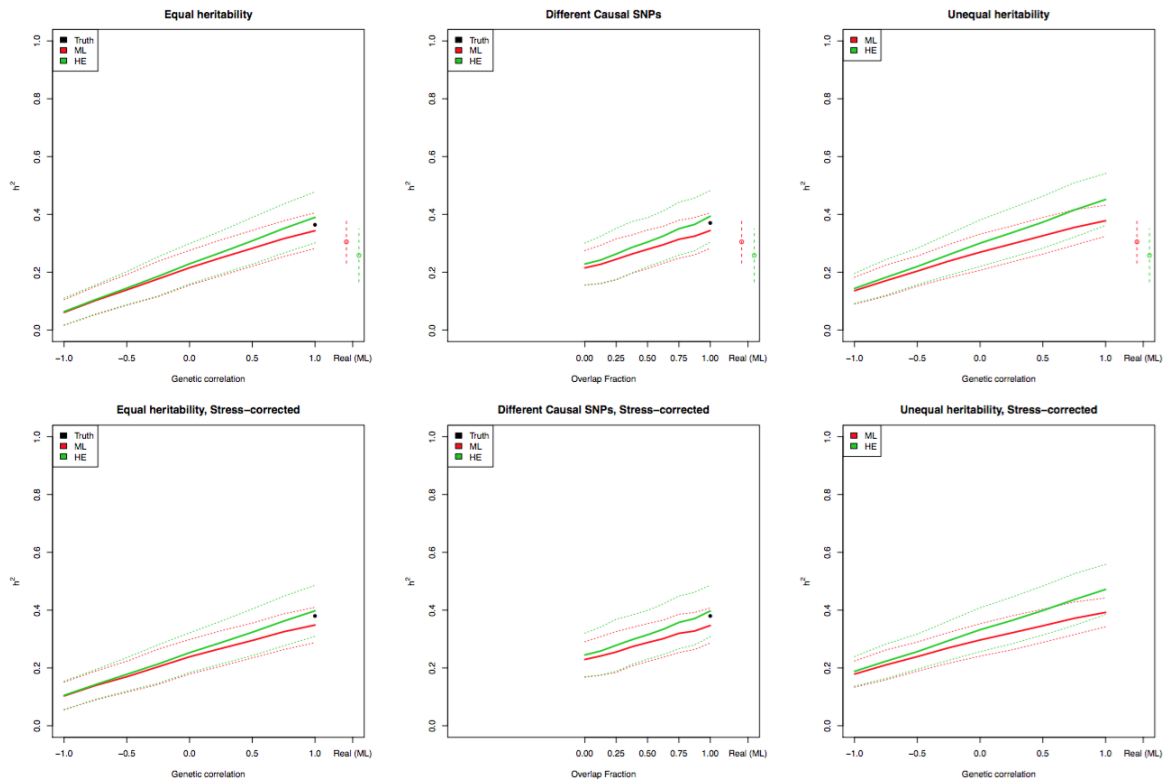
FIGURE S10. Simulation results for within-group heritability, genetic correlation, and GxE variance estimates.

Simulation results for within-adversity-group heritability (top row), genetic correlation (middle row), and GxE variance estimates (bottom row). Dotted error bars represent ± 2 standard deviations, while the dashed error bars represent ± 2 standard errors. Columns index simulation parameterizations as in Figure S8.
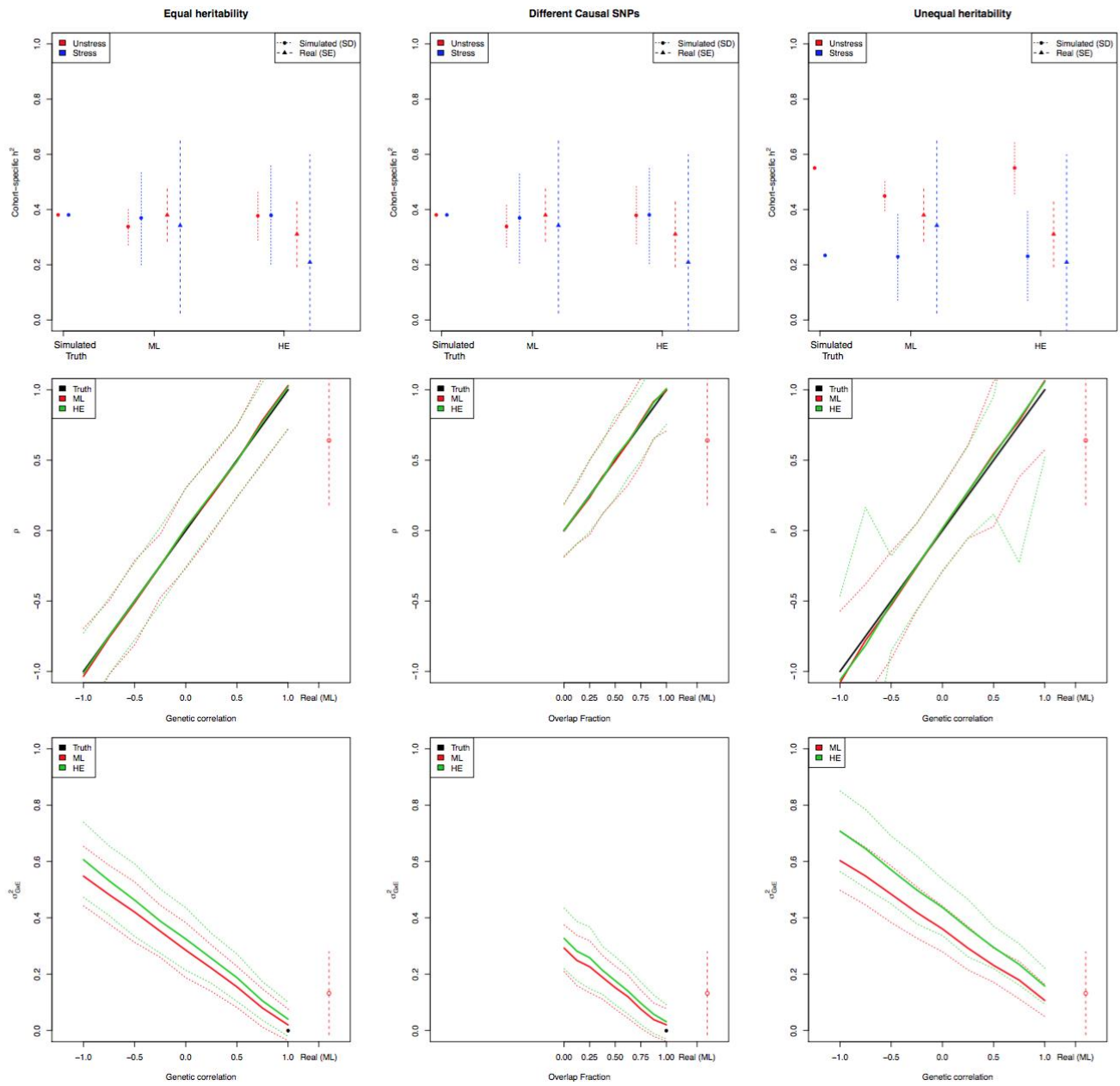
Author Contributions

JF, KSK, REP, NC designed the project; JF and KSK obtained funding for the project; JF and KSK collected the samples; NC, JF performed the genome sequencing and analysis of data quality; NC, REP, AWD, TBB, NZ, and JF performed the genetic analysis; REP, NC, AWD, KSK, and JF wrote the manuscript with assistance from TBB and NZ. All authors have read and approved the final manuscript.

CONVERGE Consortium

Na Cai, Tim B. Bigdeli, Warren Kretzschmar, Yihan Li, Jieqin Liang, Li Song, Jingchu Hu, Qibin Li, Wei Jin, Zhenfei Hu, Guangbiao Wang, Linmao Wang, Puyi Qian, Yuan Liu, Tao Jiang, Yao Lu, Xiuqing Zhang, Ye Yin, Yingrui Li, Xun Xu, Jingfang Gao, Mark Reimers, Todd Webb, Brien Riley, Silviu Bacanu, Roseann E. Peterson, Yiping Chen, Hui Zhong, Zhengrong Liu, Gang Wang, Jing Sun, Hong Sang, Guoqing Jiang, Xiaoyan Zhou, Yi Li, Yi Li, Wei Zhang, Xueyi Wang, Xiang Fang, Runde Pan, Guodong Miao, Qiwen Zhang, Jian Hu, Fengyu Yu, Bo Du, Wenhua Sang, Keqing Li, Guibing Chen, Min Cai, Lijun Yang, Donglin Yang, Baowei Ha, Xiaohong Hong, Hong Deng, Gongying Li, Kan Li, Yan Song, Shugui Gao, Jinbei Zhang, Zhaoyu Gan, Huaqing Meng, Jiyang Pan, Chengge Gao, Kerang Zhang, Ning Sun, Youhui Li, Qihui Niu, Yutang Zhang, Tieqiao Liu, Chunmei Hu, Zhen Zhang, Luxian Lv, Jicheng Dong, Xiaoping Wang, Ming Tao, Xumei Wang, Jing Xia, Han Rong, Qiang He, Tiebang Liu, Guoping Huang, Qiyi Mei, Zhenming Shen, Ying Liu, Jianhua Shen, Tian Tian, Xiaojuan Liu, Wenyuan Wu, Danhua Gu, Guangyi Fu, Jianguo Shi, Yunchun Chen, Xiangchao Gan, Lanfen Liu, Lina Wang, Fuzhong Yang, Enzhao Cong, Jonathan Marchini, Huanming Yang, Jian Wang, Shenxun Shi, Richard Mott, Qi Xu, Jun Wang, Kenneth S. Kendler & Jonathan Flint (6).

References

(1) Flint J, Chen Y, Shi S, Kendler KS. Epilogue: Lessons from the CONVERGE study of major depressive disorder in China. J Affect Disord 2012 Sep;140(1):1-5.

(2) Kendler KS, Prescott CA. Genes, Environment, and Psychopathology: Understanding the Causes of Psychiatric and Substance Use Disorders. : Guilford Press; 2007.

(3) Laumann EO. The social organization of sexuality: sexual practices in the United States. Chicago: University of Chicago Press; 1994.

(4) Kessler RC. The effects of stressful life events on depression. Annu Rev Psychol 1997;48:191-214.

(5) Kendler KS, Aggen SH. Clarifying the causal relationship in women between childhood sexual abuse and lifetime major depression. Psychol Med 2014 Apr;44(6):1213-1221.

(6) Converge Consortium. Sparse whole-genome sequencing identifies two loci for major depressive disorder. Nature 2015;523(7562):588-591.

(7) Cai N, Bigdeli TB, Kretzschmar WW, Li Y, Liang J, Hu J, et al. 11,670 whole-genome sequences representative of the Han Chinese population from the CONVERGE project. Sci Data 2017 Feb 14,;4:170011.

(8) Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 2015;4:7.

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

(10) Loh P, Tucker G, Bulik-Sullivan BK, Vilhjálmsson BJ, Finucane HK, Salem RM, et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. Nat Genet 2015 Mar;47(3):284-290.

(11) Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011 January 7,;88(1):76-82.

(12) Tyler J. VanderWeele, Mirjam J. Knol. A Tutorial on Interaction. Epidemiologic Methods 2014 Dec 1,;3(1):33-72.

(13) Kendler KS, Gardner CO. Interpretation of interactions: guide for the perplexed. Br J Psychiatry 2010 Sep;197(3):170-171.

(14) Kovalchik S, Varadhan R. Fitting Additive Binomial Regression Models with the R Package blm. Journal of Statistical Software 2013 September 3,;54.

(15) Lee SH, Yang J, Goddard ME, Visscher PM, Wray NR. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. Bioinformatics 2012;28(19):2540-2542.

(16) Peterson RE, Cai N, Bigdeli TB, Li Y, Reimers M, Nikulova A, et al. The Genetic Architecture of Major Depressive Disorder in Han Chinese Women. JAMA Psychiatry 2017 Feb 01,;74(2):162-168.

(17) Speed D, Hemani G, Johnson MR, Balding DJ. Improved heritability estimation from genome-wide SNPs. Am J Hum Genet 2012 December 7,;91(6):1011-1021.

(18) Lee SH. Implications of simplified linkage equilibrium SNP simulation. Proc Natl Acad Sci U S A 2015 Oct 06,;112(40):5449.

(19) Golan D, Lander ES, Rosset S. Measuring missing heritability: inferring the contribution of common variants. Proc Natl Acad Sci U S A 2014 December 9,;111(49):5272.

(20) Golan D, Rosset S, Lander ES. Reply to Lee: Downward bias in heritability estimation is

not due to simplified linkage equilibrium SNP simulation. Proc Natl Acad Sci U S A 2015 Oct 06,;112(40):5452.

(21) Dahl A, Iotchkova V, Baud A, Johansson Å, Gyllensten U, Soranzo N, et al. A multiple-phenotype imputation method for genetic studies. Nat Genet 2016 Apr;48(4):466-472.

(22) Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, et al. Efficient control of population structure in model organism association mapping. Genetics 2008 Mar;178(3):1709-1723.

(23) Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. American journal of human genetics 2011;88(3):294-305.

(24) Han B, Eskin E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. Am J Hum Genet 2011 May 13,;88(5):586-598.

(25) Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003 -9-6;327(7414):557-560.

(26) Bigdeli TB, Ripke S, Peterson RE, Trzaskowski M, Bacanu S-, Abdellaoui A, et al. Genetic effects influencing risk for major depressive disorder in China and Europe. Transl Psychiatry 2017 Mar 28,;7(3):e1074.

(27) Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet 2006 Feb;38(2):209-213.

(28) Kooperberg C, Leblanc M. Increasing the power of identifying gene x gene interactions in genome-wide association studies. Genet Epidemiol 2008 Apr;32(3):255-263.