Genome-wide interaction study of stress-sensitivity and its prediction of major depressive disorder in UK Biobank and Generation Scotland.

Running title: **Stress-sensitivity and depression**

Aleix Arnau-Soler MSc[1], Mark J. Adams PhD[2], Generation Scotland[3], Caroline Hayward PhD[4], Pippa A. Thomson PhD[1,5]

[1] Medical Genetics Section, University of Edinburgh, Centre for Genomic and Experimental Medicine and MRC Institute of Genetics and Molecular Medicine, Edinburgh, UK

[2] Division of Psychiatry, Deanery of Clinical Sciences, University of Edinburgh, Royal Edinburgh Hospital, Morningside Park, Edinburgh EH10 5HF, UK

[3] A collaboration between the University Medical School and NHS in Aberdeen, Dundee, Edinburgh and Glasgow, Scotland, UK

[4] Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK

[5] Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK

Correspondence to:

Dr. Pippa A. Thomson

Centre for Genomic and Experimental Medicine

Institute of Genetics and Molecular Medicine

University of Edinburgh

Western General Hospital

Crewe Road South

Edinburgh, EH4 2XU, UK

Phone: +44(0)1316518756

Email: Pippa.Thomson@ed.ac.uk

ABSTRACT

Stressful life events (SLE) are important predictors of neuroticism and the onset of major depressive disorder (MDD), complex traits involving both genetic and environmental factors. Identification of the genetics underpinning the individual differences in response to SLE may improve our understanding of molecular pathways involved, and the effects of stress-sensitivity on stress-related illnesses. We sought the combined effects of SLE on MDD status and neuroticism, as a proxy for stress-sensitivity, by modeling the interaction between SNP allele and MDD status on neuroticism score. We performed genome-wide interaction studies (GWIS) in UK Biobank (N = 23 092; UKB) and Generation Scotland: Scottish Family Health Study (N = 7 155; GS:SFHS). No SNP in the GWIS achieved genome-wide significance, but gene-based tests identified a genome-wide significant association between stress-sensitivity and *ZNF366*, a negative regulator of glucocorticoid receptor function that has been implicated in alcohol dependence ($p = 1.48 \times 10^{-7}$; Bonferroni-corrected significance threshold $p < 2.79 \times 10^{-6}$). From the GWIS results, SNP heritability for stress-sensitivity was estimated at 5%. We next tested whether polygenic risk scores (PRS) for stress-sensitivity improved the prediction of MDD. In models fitting both PRS of the main effects of MDD and neuroticism, stress-sensitivity PRS significantly improved the prediction of MDD beyond those based on main additive effects. This study may improve understanding of the etiological mechanisms underpinning stress-sensitivity, and understanding of the results of larger genome-wide association studies of MDD and other stress-related illnesses.

INTRODUCTION

Stressful life events (SLE) are known to effect mental and disease-related traits including neuroticism, MDD, autoimmune diseases and some cancers[1], A greater understanding of the causal mechanism by which SLE affect disease risk or outcome may be beneficial in identifying individuals for targeted support. In particular, it is increasingly clear that the effect on an individual results more from the perceived stress than the event itself. Few studies of stress-related traits have prospective data on SLE and where data is retrospectively ascertained this is primarily through questionnaires that assess a limited number of events and may be affected by the current emotional state of a respondent[2]. Previous studies have shown that the effects of environmental factors on traits may be mediated by genetics[3] or gene-environment interactions such as those studied through emerging genome-wide by environment interaction studies (GWEIS). This may provide an approach to identify the genetic variation that affects an individual's sensitivity to stressful life events, even in the absence of data on SLE, and to improve genetic prediction of an individual's liability to negative outcomes. We sought to model genetic sensitivity to stress through the combined effects of SLE on neuroticism, as a proxy for sensitivity to negative experiences, and on lifetime diagnosis of depression.

Major depressive disorder (MDD) is a common psychiatric illness with a prevalence of 12% in the Scottish population[4]. It is a complex disorder influenced by both genetic contributions, with heritability estimates from twin and family studies of between 31-42%[5], and environmental risk factors. Genome-wide association studies (GWAS) of depression have identified common variants that account for approximately 0.6% of the genetic heritability and shown that individuals with a high loading of common risk variants have an increased risk of developing MDD[6-8]. Polygenic risk scores (PRS), derived from the summary data of MDD-GWAS, significantly predict variability in MDD liability, with greater accuracy as GWAS sample sizes and the numbers of variants used for prediction increase. Confirmed environmental risk factors for MDD include maternal infections, childhood maltreatment and adverse life events[8-12], however, few genetic studies have such information and even fewer prospective studies exist.

Incorporation of SLE data has been shown to improve the ability to predict MDD[13, 14]. A genetic contribution has been suggested for the effects of adverse life events on risk of depression[15, 16].

The personality trait neuroticism is highly heritable[17-19], is higher in individuals with depression compared to controls[20, 21] and is known to have shared genetic aetiology with depression[22, 23]. Neuroticism is strongly correlated with measures of sensitivity to punishment but not reward[24] and has been used previously as a measure of stress-sensitivity[25]. Neuroticism is thought to mediate or interact with the effects of adverse life events on risk of depression[9, 26]. It has a substantial stable component[27], however, there is evidence for change, as well as stability, across the life span[28-31]. Whereas the stable component of neuroticism is strongly determined by genetics, change in neuroticism score is attributed to the effects of unshared environment[28]. Persistent change in neuroticism score has been shown in response to life events[29-31].

The relationship between neuroticism scores and adverse life events is unclear, with studies suggesting both that neuroticism may predict future exposure to adverse life events and that exposure to such events may result in increased measures of neuroticism, particularly when adversity is perceived as highly stressful[32-35]. Multiple regression analyses suggest that, when combined, both recent adverse life events and high neuroticism scores predict development of a mental disorder and poorer outcome in those with a disorder at baseline[36]. This suggests that the effects of adverse life events and neuroticism on depression risk are to some extent distinct. Although, Colodro-Conde *et al.* showed a positive interaction between PRS for MDD and personal SLE on liability to depressive symptoms[14]. However, multivariate structural equation models confirm a genetic effect of perceived stress, distinct from that of MDD or neuroticism[37]. These studies suggest that a genetic variable derived from the difference in neuroticism levels seen in individuals with MDD compared to controls may allow us to identify genetic loci important on stress-sensitivity. Further, PRS derived from this stress-sensitivity variable may improve prediction of MDD over that based on MDD or neuroticism PRS alone.

Using unrelated individuals from two large population-based samples, UK Biobank (UKB; N = 23 092) and Generation Scotland: Scottish Family Health Study (GS:SFHS; N = 7 155), we sought to identify genes involved in stress-sensitivity (SS) by performing GWIS for the interaction between MDD status and SNP allele on neuroticism score. We identify a gene significantly associated with stress-sensitivity and show that a PRS derived from the GWIS, significantly predicts liability to depression independently of the PRS for MDD and/or neuroticism.

MATERIALS AND METHODS

Phenotype assessment

Study ethics and cohort descriptions are detailed in Supplemental Information: Ethics statement & Cohorts' description.

MDD status and the Eysenck Personality Questionnaire-Revised Short Form's Neuroticism Scale (EPQN) were available in both UKB and GS:SFHS cohorts[4, 38]. Further details on phenotypes assessment are given in Supplemental Information: Phenotype assessment.

GWIS and derivation of a genetic stress-sensitivity (SS) effect

The effect size of an stress-sensitivity effect (β_{SS}) was derived by performing a GWIS for the effect of the MDD-status and SNP allele interaction on EPQN (dependent variable) in both UKB and GS:SFHS cohorts using PLINK 1.90 (PLINK-command --gxe; fitting MDD diagnosis as a binary "group" effect)[39]. PLINK-command --gxe estimates the difference in allelic association with a quantitative trait (EPQN) between two groups (MDD cases vs. controls) producing effect estimates on each group and a significance estimate for the interaction between SNP allele and MDD status. The interaction *p* value reflects the difference between the regression coefficient of the allelic effect in a linear model for EPQN in MDD cases (β_A) and the same regression coefficient in a linear model for EPQN in controls (β_B).

Considering one SNP, the effect it confers to EPQN can be modelled by MDD status (control = 0, MDD case = 1) as follows:

$$\begin{cases} MDD = 0; \;\; EPQN = \beta_0 + \beta_B SNP + COV + \varepsilon \\ MDD = 1; \; EPQN = \beta_1 + \beta_A SNP + COV + \varepsilon \end{cases}$$

This is equivalent to modelling the effect on MDD cases as follows:

$$\begin{cases} MDD = 0; \;\; EPQN = \beta_0 + \beta_B SNP + COV + \varepsilon \\ MDD = 1; \; EPQN = \beta_1 + \beta_B SNP + (\beta_A - \beta_B) SNP + COV + \varepsilon \end{cases}$$

Or, it can be modelled as a whole as:

$$EPQN = (\beta_1 - \beta_0)MDD + \beta_B SNP + (\beta_A - \beta_B)SNP * MDD + COV + \varepsilon$$

Thus, the interaction effect (β_{SS}) can be estimated as the difference in allelic effect on EPQN between MDD cases (β_A) and controls (β_B) as follows,

$$\hat{\beta}_{SS} = \hat{\beta}_A - \hat{\beta}_B$$

$\hat{\beta}_{SS}$ is therefore defined as the effect size reflecting the genetic stress-sensitivity effect on MDD cases compared to controls (see Supplemental Figure S1).

Stress-sensitivity GWIS, main additive effect GWASs, meta-analysis and gene-set analysis. Details of genotyping, imputation and quality control are provided in Supplemental Information: Genotyping, imputation and quality control.

For GWIS and subsequent analyses, sample specific covariates were applied as follows: *UKB.* All phenotypes were adjusted for centre, array and batch as random effects before all analyses. Analyses were adjusted for age, sex and 15 informative principal components (PCs; UKB Data Dictionary items #22009.01 to #22009.15) as fixed effects to take account of possible population stratification. *GS:SFHS.* All the analyses were adjusted for age, sex and 20 PCs.

GWAS for MDD (MDD-GWAS) and neuroticism (EPQN-GWAS), using logistic and linear models of additive allelic effects respectively, were conducted on the same sample sets for comparison and generation of matched PRS.

Results from the GWIS of UKB and GS:SFHS were combined in a sample size weighted meta-analysis performed using METAL[40]. While the use of standard error weighting is more common, the different diagnostic schema and MDD prevalence between the two cohorts (GS:SFHS; 12.2%, UKB: 25.8%)[4, 38] may indicate systematic differences in the measurement of MDD. Generalized gene-based analysis of the meta-analysis was performed using MAGMA[41] implemented through FUMA[42] (http://fuma.ctglab.nl). Briefly, SNP summary statistics were mapped to 17 931 protein-coding genes. Individual SNP *p* values from a gene were combined into a gene test-statistic using a SNP-wise model and a known approximation of the sampling distribution used

to obtain a gene-based p value. Genome-wide significance was defined at p = 0.05/17 931 = 2.79×10^{-6}.

LD Score regression

To examine the genetic overlap between the polygenic architecture of SS, MDD and neuroticism, we used LD score regression to derived genetic correlations (r_G) between these traits assessed[43, 44]. SNP-based heritability was also estimated using LD score regression, which only requires summary statistics from single-SNP analysis.

Pathway, functional and gene expression analyses

Lead SNPs, independently associated with the phenotype, were identified using PLINK by clumping (p threshold < 2×10^{-5}; LD r2 > 0.1; physical kb threshold = 500kb; 1000 Genomes Project Phase 1 CEU, GBR, TSI genotype data), and analysed using DEPICT[45]. Further detail is given in the Supplemental Information: DEPICT analyses.

Lead SNPs were investigated for evidence of phenotypic association in the NCBI dbGaP database of genotypes and phenotypes[46] (https://www.ncbi.nlm.nih.gov/gap/phegeni) and regulatory DNA elements in normal cell lines and association with expression quantitative trait loci (eQTLs) using the RegulomeDB database[47] (http://www.regulomedb.org) and the Genotype-Tissue Expression (GTEx) Portal[48] (http://www.gtexportal.org).

Polygenic profiling

PRS were produced using PRSice[49] and standardized to a mean of 0 and a standard deviation of 1. PRS for stress-sensitivity (PRS_{SS}) were created by weighting the sum of the reference alleles in an individual by the SS-effect, β_{SS}. Additional PRS were generated weighting by the matched MDD main additive effects (PRS_D) and neuroticism main additive effects (PRS_N). PRS were also generated using summary statistics from the most recent Psychiatric Genetic Consortium (PGC) MDD meta-analysis[50] (excluding GS:SFHS, and UKB individuals when required; N = 155 866 &

138 884) and the Genetics of Personality Consortium (GPC) neuroticism meta-analysis[18, 51] (N = 63 661). Further detail is given in the Supplemental Information: Polygenic Profiling.

Generalized linear models were implemented in R 3.1.3[52]. The direct effect of PRS_{SS} (model 1), PRS_D (model 2) and PRS_N (model 3) on MDD risk were assessed in independent logistic regression models on GS:SFHS using GWAS statistics from UKB (the largest cohort) as the discovery sample to weight PRS. Multiple regression models fitting both PRS_D and PRS_N (model 4) and fitting each of them separately with PRS_{SS} (models 5 and 6) were also calculated. Finally, full additive multiple regression models fitting PRS weighted by all three effects was assessed using both match PRS_D and PRS_N or using PRS_D and PRS_N weighted by PGC2 MDD[50] and GPC neuroticism[51] summary statistics.

All models were adjusted by sex, age and 20 PCs. A null model was estimated from the direct effects of all covariates on MDD. Likelihood ratio test was used to assess significance of each PRS.

Cross-validation was performed using UKB. MDD status on UKB was adjusted by centre, array and genotyping batch as random effects and scaled (between 0 and 1) prior to analysis, giving a quasi-binomial distribution of MDD status on UKB. Models implemented on UKB (quasi-binomial regression) were adjusted by sex, age and 15 PCs.

Nagelkerke's R^2 coefficients were estimated to quantify the proportion of MDD liability explained at the observed scale by each model and converted into R^2 coefficients at the liability scale (prevalence: 12.2% in GS:SFHS[4] and 25.8% in UKB[38]) using Hong Lee's transformation[53] available from GEAR: GEnetic Analysis Repository[54].

Using stress-sensitivity to stratify depression

GS:SFHS MDD cases (n_{cases} = 2 016; n_{female} = 1 345, n_{male} = 671) have data available on MDD course (single or recurrent), age of onset (n = 1 964) and episode count (n = 2 016), as well as on neuroticism (n = 2 010). In addition, a subset were evaluated by Mood Disorder Questionnaire[55] (MDQ; n = 1 022) and Schizotypal Personality Questionnaire[56] (SPQ; n = 1 093). The reduced

sample number of MDQ and SPQ reflects the later addition of these questionnaires to the study, and does not reflect a particular subgroup of GS:SFHS.

Difference in PRS_{SS} and PRS_{D} between MDD cases and controls on GS:SFHS were tested using a Student's two sample t-test (two tailed). Cases of MDD on GS:SFHS with data available on each trait analyzed were stratified by quintiles based on PRS_{SS} and PRS_{D} (5x5 groups). Post hoc, the effects on each trait of quintiles based on PRS_{SS} and its interaction effect with quintiles based on PRS_{D} were assessed using linear regression models adjusting by sex and age. Same analysis was reproduced using PRSs as continuous variables.

RESULTS

In GS:SFHS, individuals with a diagnosis of MDD had significantly higher EPQN scores compared to controls ($p < 2.2x10^{-16}$; $mean_{controls} = 3.16$; $mean_{cases} = 6.42$) and UKB ($p < 2.2x10^{-16}$; $mean_{controls} = 2.79$; $mean_{cases} = 5.64$). This difference remained significant after matching cases with controls for PGC PRS_D, sex and age. (GS:SFHS: $p < 2.2x10^{-16}$, $mean_{controls} = 3.56$; UKB: $p < 2.2x10^{-16}$, $mean_{controls} = 2.97$; see Supplemental Table S1).

Meta-analysis of MDD-dependant stress-sensitivity GWIS in UKB and GS:SFHS

No SNPs were associated with stress-sensitive at the genome-wide significant threshold ($p < 5x10^{-8}$, Figure 1). However, 14 SNPs from 8 loci achieved suggestive *p* values ($p < 1x10^{-5}$; range $8.9x10^{-6}$-$5.1x10^{-7}$) (Meta-analysis: Table 1; UKB and GS:SFHS: Supplemental Tables S2 & S3; Meta-analysis QQ-plot with λ: Supplemental Figure S2; UKB and GS:SFHS QQ-plots: Supplemental Figure S3). Traits with prior evidence of association with the nearest genes to the 8 lead SNPs were identified using dbGap and are shown in Supplemental Table S4. Comparison between the SNP association profile along the genome between SS and MDD main additive effects meta-analysis is shown in Miami plots filtering for the most significant SS or MDD SNPs ($p < 0.001$; Meta-analysis: Figure 2; UKB and GS:SFHS: Supplemental Figure S4). Gene-based test identified *ZNF366* as the only gene achieving genome-wide significance ($p = 1.48x10^{-7}$; Bonferroni-corrected significance threshold $p < 2.79x10^{-6}$; Supplemental Table S5 & Supplemental Figure S5). SS SNP-based heritability was estimated at 5.0% ($h^2 = 0.0499$, s.e. = 0.017), compared to matched SNP-based heritabilities of 9.6% ($h^2 = 0.0962$, s.e. = 0.0179) and 10.1% ($h^2 = 0.1006$, s.e. = 0.0076) for MDD and neuroticism, respectively.

LD score regression was performed to obtain genetic correlations between SS, MDD and neuroticism. As previously shown, there was a significant genetic correlation between MDD and neuroticism ($r_G = 0.637$, s.e. = 0.0704, $p = 1.39x10^{-19}$). However, we found no evidence for a genetic correlation between SS and MDD (rG = -0.099, s.e. = 0.182, $p = 0.585$) or between SS and neuroticism ($r_G = 0.114$, s.e. = 0.107, $p = 0.285$).

Pathway enrichment, functional annotation and gene expression analyses

Lead SNPs from the GWIS meta-analysis were investigated using DEPICT. No gene showed statistically significant links to SS at a DEPICT false discovery rate (FDR) < 0.05. No significant result was found for either gene set analysis or tissue enrichment analysis at FDR < 0.05. Evidence of regulatory elements on normal cell lines/tissues was identified for 5 of the 12 lead SNPs (i.e. rs3762096, rs10987199, rs2221540, rs246565, rs319924). Two lead SNPs were associated with eQTLs: rs319924 (an intronic SNP in *EYS*) and rs9509508 (an intronic SNP in *LATS2*) and potentially regulate *LGSN/RP3-407E4.3* ($p = 6.31\text{x}10^{-12}$/ $p = 1.15\text{x}10^{-5}$) and *LATS2* ($p = 3.74\text{x}10^{-8}$), respectively.

Polygenic risk scores for stress-sensitivity predict MDD liability.

PRS were used to investigate whether common variants affecting stress-sensitivity predict MDD risk. We generated PRS (PRS_{SS}) based on the summary statistics from the meta-analysis of the GWIS for SS. PRS_{SS} significantly predicted MDD risk in GS:SFHS using weights from the larger UKB summary data ($p = 1.8\text{x}10^{-3}$; $\beta = 0.087$; s.e. = 0.028; p threshold = 0.004). This association was cross-validated in UKB using summary data from the smaller GS:SFHS GWIS ($p = 0.04$; $\beta = 0.027$; s.e. = 0.013; p threshold = $1.15\text{x}10^{-4}$; Table 2 and Supplemental Figure S6). On the liability scale, the MDD variance explained by PRS_{SS} was modest, 0.243% in GS:SFHS and 0.037% in UKB. This was less than predicted by the main effects of MDD or neuroticism (PRS_{D}: 0.386% in GS:SFHS and 0.211% in UKB; PRS_{EPQN}: 0.34% in GS:SFHS and 0.178% in UKB; Table 2 and Supplemental Figure S6).

Due to the known genetic correlations between MDD, neuroticism and SLE, models jointly fitting the effects of multiple PRS were analysed. Multiple regression analyses showed that, compared to MDD main additive effects alone, SS-effect significantly explains and additional 0.248% in GS:SFHS (a predictive improvement of 64.3%; $p = 1.6\text{x}10^{-3}$; PRS_{D}: $\beta = 0.113$ and s.e. = 0.028; PRS_{SS}: $\beta = 0.088$ and s.e. = 0.028) and an additional 0.036% in UKB (a predictive

improvement of 17.1%; $p = 0.044$; PRS_D: β = 0.064 and s.e. = 0.013; PRS_{SS}: β = 0.027 and s.e. = 0.013; Figure 3). When fitting both MDD and neuroticism main additive effects derived from the same sample sets, the inclusion of PRS_{SS} in the full model remained significant explaining an additional 0.245% (a predictive improvement of 35.8%: full model $p = 1.7x10^{-3}$; PRS_D: β = 0.107 and s.e. = 0.029; PRS_{EPQN}: β = 0.123 and s.e. = 0.036; PRS_{SS}: β = 0.088 and s.e. = 0.028) and an additional 0.036% (a predictive improvement of 11.2%; full model $p = 0.044$; PRS_D: β = 0.054 and s.e. = 0.014; PRS_{EPQN}: β = 0.048 and s.e. = 0.014; PRS_{SS}: β = 0.027 and s.e. = 0.013) in GS:SFHS and UKB, respectively.

In models fitting PRS_D and PRS_N, the variances explained were non-additive, demonstrating the partial overlap between MDD risk prediction from MDD main additive effects and prediction from neuroticism main additive effects. This overlap was not seen between variance explained by SS-effect and the variance explained by either MDD and/or neuroticism main additive effects. Multiple regression analyses fitting main additive effects of neuroticism and MDD derived from worldwide consortiums (Figure 3) showed that the increased sample size from GWAS used to derive MDD main additive effects, resulted in an increment of MDD variance explained by MDD main additive effects (GS:SFHS: from 0.386% to 1.233%; UKB: from 0.21% to 1.281%). However, there was no change in the proportion of the variance explained by the PRS_{SS} in the full model (PRS_{SS} $p = 1.29x10^{-3}$ and $p = 0.042$ in GS:SFHS and UKB, respectively). These results suggest that SS explains a proportion of MDD risk not accounted for by the main additive effects of MDD or neuroticism at current sample sizes.

Using SS to stratify MDD in GS:SFHS

MDD cases show significantly higher PRS_{SS} ($p = 2x10^{-3}$) and PRS_D ($p = 1.8x10^{-4}$) than controls. Association between MDD-related traits and stress-sensitivity risk quintiles was assessed on MDD cases in order to identify a subgroup of MDD patients, perhaps defining a characteristic aetiological subtype of MDD. However, stratification analysis failed, and no quintile based on PRS_{SS} nor its interaction with quintiles based on PRS_D showed statistically significant effects on

any trait analyzed. Individuals with high PRS_{SS} were not significantly different from other cases for sex, MDD course, age of onset or episode count, nor neuroticism, mood disorder or schizotypal personality scores ($p > 0.05$; see Supplemental Table S6). Results were unchanged when PRSs were used as continuous variables.

DISCUSSION

Through international research efforts such as the PGC and UK Biobank, there are ever-increasing sample sizes available for understanding the genetics of major depressive disorder. These resources are beginning, and will continue to, identify genome-wide significant loci[57, 58]. However, the acquisition of environmental data, particularly that of stressful life events (both in childhood and adulthood), has been more problematic. We sought to identify the genetic underpinnings of sensitivity to such events and show that even in the absence of such environmental data, such information could improve the prediction of MDD risk.

We performed a GWIS to identify loci showing differential effects on neuroticism scores in individuals with and without MDD. No SNPs reached genome-wide significance but 14 SNPs from 8 loci reached suggestive significance levels. Phenotypic associations of these loci were investigated in the dbGAP repository for GWAS studies. The putative genes for stress-sensitivity have prior evidence of association with a wide range of neuropsychiatric traits such as schizophrenia, bipolar disorder, attention deficit disorder with hyperactivity, mental competency, intuition, sleep or alcohol drinking (Supplemental Table S4: Traits with significant evidence of association with suggestive stress-sensitive genes). Nevertheless, enrichment analysis performed on the loci with suggestive *p* values for stress-sensitivity showed no evidence for specific pathway or tissue enrichment. The top two loci, *PTP4A1-PHF3/EYS* and *ZNF366* have been previously associated with alcohol dependence[59-63], alcohol intake (phs000342) and glucocorticoid receptor function[64-66]. The most significant SNP in this study, rs319924, is an intronic variant in *EYS* that is a potential eQTL for *LGSN*[48], a gene previously associated with male-specific depression[67]. This is of particular interest given previous studies linking alcohol consumption, stress and the risk of depression[68-73]. Evidence of an eQTL effect was also predicted for a lead SNP in *LATS2*, a positive regulator of histone methyltransferase activity[74] a process important in anxiety-related behaviours[75]. The prior association of the top

two loci in this study with alcohol related-phenotypes suggests that genes involved in the sensitivity to environmental stress may mediate the effects of stress on alcohol consumption.

Gene-set analysis using summary statistics from the stress-sensitivity meta-analysis identified association between *ZNF366* and stress-sensitivity at the genome-wide significant level. *ZNF366* (also known as *DC-SCRIPT*) is a corepressor of transcription found in nuclear receptor complexes including the glucocorticoid receptor (GR). *ZNF366* represses GR-mediated transcription in monocyte-derived dendritic cells[64]; and may act through C-terminal binding protein (CTBP1) and histone deacetylases to modulate immune response[65]. There is evidence from a large-scale mRNA display study that *PHF3*, in the region underlying the most significant peak in the single SNP analysis, may also interact, directly or indirectly, with the glucocorticoid receptor (IntAct database[66]) but this has not been confirmed. These results reinforce the hypothesis that we are modelling the genetic architecture of sensitivity to adverse life events.

We estimated a significant lower bound on heritability for stress-sensitivity of 5%. Whilst the known genetic overlap between MDD and neuroticism was detectable, the lack of genetic correlation with SS indicated that there is no significant overlap MDD-dependant stress-sensitivity and MDD or neuroticism. This analysis may be limited by our sample size, however using the largest available meta-analyses of MDD and neuroticism[50, 51] did not decrease the proportion of liability explained by the PRS_{SS}. We note, however, that as such meta-analyses increase in size it is likely, as with the effects of smoking in schizophrenia[76, 77], that the indirect genetic effects of the environment on the risk of depression will be detected by GWAS. However, it is only through studies such as ours that the mechanism for the effect of the risk alleles may be clarified.

The summary results from the GWIS were used to derive a PRS reflecting the genetic difference in stress-sensitivity between MDD and controls. This variable significantly predicted liability to

MDD in GS:SFHS (p = 1.8x10^{-3}) and this was replicated in UKB although with lower statistical significance (p = 0.04). This is consistent with the expectation that the larger the discovery sample (i.e. UKB), the greater the accuracy of the weighting and the more predictive the PRS[78]. Multiple regression models suggest that inclusion of PRS weighted by MDD-dependant stress-sensitivity significantly improves MDD prediction over use of either MDD and/or neuroticism weighted PRS alone (GS:SFHS full model p = 1.7x10^{-3}; UKB full model p = 0.04). However, we were unable to identify a subgroup of MDD cases with higher PRS_{ss}.

A key limitation of this study is the lack of prospective stressful life event data. Future studies will be required to further investigate the effects of adverse life events in individuals with high or low polygenic risk scores for MDD-dependant stress-sensitivity. However, the methodology presented allows for the use of susceptibility to stress as a proxy to study gene-by-stressful life events interaction in the absence of stressful-life events data.

The existence of genetic variants affecting an individual's risk of depression in response to stress has been predicted previously[15, 16] and is consistent with the departure from a simple additive genetic model seen in twin-studies of recurrent depressive disorder[79]. However, candidate gene studies, particularly those of related to the serotonin system, shown mixed results in identifying SNPs affecting response to adverse life events[80-84]. A GWEIS of depressive symptoms with adverse life events found a single genome-wide significant SNP near *CEP350*, but was unable to replicate it[85]. The polygenic interaction approach used in our study, and paralleling recent developments in GWAS analyses, maximised our power to detect gene-by-environment effects in this heterogeneous disorder, identifying an independent set of risk variants for stress-sensitivity and significantly improving the proportion of liability to MDD predicted by PRS. Here we show that incorporating information across many loci provides clear and replicable evidence for a genetic effect of environmental stress on MDD risk and identifies a potential genetic link with alcohol intake. These results require further study, but may inform treatment of comorbid alcohol dependency and depression.

ACKNOWLEDGMENTS

This research has been conducted using the UK Biobank Resource under UK Biobank main application 4844. The 1st author AAS is funded by University of Edinburgh (www.ed.ac.uk) and MRC for his PhD study at the University of Edinburgh Institute of Genetics and Molecular Medicine (www.ed.ac.uk/igmm). MA is supported by STRADL through a Wellcome Trust Strategic Award (reference 104036/Z/14/Z). Generation Scotland received core support from the Chief Scientist Office of the Scottish Government Health Directorates [CZD/16/6] and the Scottish Funding Council [HR03006]. Genotyping of the GS:SFHS samples was carried out by the Genetics Core Laboratory at the Wellcome Trust Clinical Research Facility, Edinburgh, Scotland and was funded by the Medical Research Council UK and the Wellcome Trust (Wellcome Trust Strategic Award "STratifying Resilience and Depression Longitudinally" (STRADL) Reference 104036/Z/14/Z)."

FINANCIAL DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

1. Porcelli B, Pozza A, Bizzaro N, Fagiolini A, Costantini MC, Terzuoli L, *et al*. Association between stressful life events and autoimmune diseases: A systematic review and meta-analysis of retrospective case-control studies. *Autoimmun Rev* 2016; **15**(4): 325-334.

2. Monroe SM. Modern approaches to conceptualizing and measuring human life stress. *Annu Rev Clin Psychol* 2008; **4:** 33-52.

3. Luciano M, Batty GD, McGilchrist M, Linksted P, Fitzpatrick B, Jackson C, *et al*. Shared genetic aetiology between cognitive ability and cardiovascular disease risk factors: Generation Scotland's Scottish family health study. *Artificial Intelligence* 2010; **38**(3): 304-313.

4. Fernandez-Pujals AM, Adams MJ, Thomson P, McKechanie AG, Blackwood DH, Smith BH, *et al*. Epidemiology and Heritability of Major Depressive Disorder, Stratified by Age of Onset, Sex, and Illness Course in Generation Scotland: Scottish Family Health Study (GS:SFHS). *PLoS One* 2015; **10**(11): e0142197.

5. Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. *Am J Psychiatry* 2000; **157**(10): 1552-1562.

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

7. Demirkan A, Penninx BW, Hek K, Wray NR, Amin N, Aulchenko YS, *et al*. Genetic risk profiles for depression and anxiety in adult and elderly cohorts. *Mol Psychiatry* 2011; **16**(7): 773-783.

8. Peyrot WJ, Milaneschi Y, Abdellaoui A, Sullivan PF, Hottenga JJ, Boomsma DI, *et al*. Effect of polygenic risk scores on depression in childhood trauma. *Br J Psychiatry* 2014; **205**(2): 113-119.

9. Kendler KS, Kuhn J, Prescott CA. The interrelationship of neuroticism, sex, and stressful life events in the prediction of episodes of major depression. *Am J Psychiatry* 2004; **161**(4): 631-636.

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

11. Loret de Mola C, de Franca GV, Quevedo Lde A, Horta BL. Low birth weight, preterm birth and small for gestational age association with adult depression: systematic review and meta-analysis. *Br J Psychiatry* 2014; **205**(5): 340-347.

12. Tennant C. Life events, stress and depression: a review of recent findings. *Aust N Z J Psychiatry* 2002; **36**(2): 173-182.

13. Mullins N, Power RA, Fisher HL, Hanscombe KB, Euesden J, Iniesta R, *et al*. Polygenic interactions with environmental adversity in the aetiology of major depressive disorder. *Psychol Med* 2016; **46**(4): 759-770.

14. Colodro-Conde L, Couvy-Duchesne B, Zhu G, Coventry WL, Byrne EM, Gordon S, *et al*. A direct test of the diathesis-stress model for depression. *Mol Psychiatry* 2017.

15. Silberg J, Rutter M, Neale M, Eaves L. Genetic moderation of environmental risk for depression and anxiety in adolescent girls. *Br J Psychiatry* 2001; **179**: 116-121.

16. Kendler KS, Kessler RC, Walters EE, MacLean C, Neale MC, Heath AC, *et al*. Stressful life events, genetic liability, and onset of an episode of major depression in women. *Am J Psychiatry* 1995; **152**(6): 833-842.

17. Glahn DC, Curran JE, Winkler AM, Carless MA, Kent JW, Jr., Charlesworth JC, *et al*. High dimensional endophenotype ranking in the search for major depression risk genes. *Biological psychiatry* 2012; **71**(1): 6-14.

18. van den Berg SM, de Moor MH, McGue M, Pettersson E, Terracciano A, Verweij KJ, *et al*. Harmonization of Neuroticism and Extraversion phenotypes across inventories and cohorts in the Genetics of Personality Consortium: an application of Item Response Theory. *Behav Genet* 2014; **44**(4): 295-313.

19. Wray NR, Birley AJ, Sullivan PF, Visscher PM, Martin NG. Genetic and phenotypic stability of measures of neuroticism over 22 years. *Twin Res Hum Genet* 2007; **10**(5): 695-702.

20. Hirschfeld RM, Klerman GL. Personality attributes and affective disorders. *Am J Psychiatry* 1979; **136**(1): 67-70.

21. Hirschfeld RM, Klerman GL, Clayton PJ, Keller MB, McDonald-Scott P, Larkin BH. Assessing personality: effects of the depressive state on trait measurement. *Am J Psychiatry* 1983; **140**(6): 695-699.

22. Levinson DF. The genetics of depression: a review. *Biol Psychiatry* 2006; **60**(2): 84-92.

23. Middeldorp CM, Cath DC, Van Dyck R, Boomsma DI. The co-morbidity of anxiety and depression in the perspective of genetic epidemiology. A review of twin and family studies. *Psychol Med* 2005; **35**(5): 611-624.

24. Smits DJM, Boeck PD. From BIS/BAS to the big five. *European Journal of Personality* 2006; **20**(4): 255-270.

25. Tak LM, Kingma EM, van Ockenburg SL, Ormel J, Rosmalen JG. Age- and sex-specific associations between adverse life events and functional bodily symptoms in the general population. *J Psychosom Res* 2015; **79**(2): 112-116.

26. Hovens JG, Giltay EJ, van Hemert AM, Penninx BW. Childhood Maltreatment and the Course of Depressive and Anxiety Disorders: The Contribution of Personality Characteristics. *Depress Anxiety* 2016; **33**(1): 27-34.

27. Bazana PG, Stelmack RM. Chapter 8 - Stability of Personality Across the Life Span: A Meta-Analysis. *On the Psychobiology of Personality*. Elsevier: Oxford, 2004, pp 113-144.

28. McGue M, Bacon S, Lykken DT. Personality Stability and Change in Early Adulthood. *Developmental Psychology* 1993; **29**(1): 96-109.

29. Jeronimus BF, Ormel J, Aleman A, Penninx BW, Riese H. Negative and positive life events are associated with small but lasting change in neuroticism. *Psychol Med* 2013; **43**(11): 2403-2415.

30. Jeronimus BF, Riese H, Sanderman R, Ormel J. Mutual reinforcement between neuroticism and life experiences: a five-wave, 16-year study to test reciprocal causation. *Journal of personality and social psychology* 2014; **107**(4): 751-764.

31. Riese H, Snieder H, Jeronimus BF, Korhonen T, Rose RJ, Kaprio J, *et al*. Timing of Stressful Life Events Affects Stability and Change of Neuroticism. *European Journal of Personality* 2014; **28**(2): 193-200.

32. Boals A, Southard-Dobbs S, Blumenthal H. Adverse events in emerging adulthood are associated with increases in neuroticism. *J Pers* 2015; **83**(2): 202-211.

33. Lockenhoff CE, Terracciano A, Patriciu NS, Eaton WW, Costa PT, Jr. Self-reported extremely adverse life events and longitudinal changes in five-factor model personality traits in an urban sample. *J Trauma Stress* 2009; **22**(1): 53-59.

34. Poulton RG, Andrews G. Personality as a cause of adverse life events. *Acta Psychiatr Scand* 1992; **85**(1): 35-38.

35. Rijsdijk FV, Sham PC, Sterne A, Purcell S, McGuffin P, Farmer A, *et al*. Life events and depression in a community sample of siblings. *Psychol Med* 2001; **31**(3): 401-410.

36. Beard JR, Heathcote K, Brooks R, Earnest A, Kelly B. Predictors of mental disorders and their outcome in a community based cohort. *Soc Psychiatry Psychiatr Epidemiol* 2007; **42**(8): 623-630.

37. Rietschel L, Zhu G, Kirschbaum C, Strohmaier J, Wust S, Rietschel M, *et al.* Perceived stress has genetic influences distinct from neuroticism and depression. *Behav Genet* 2014; **44**(6): 639-645.

38. Smith DJ, Nicholl BI, Cullen B, Martin D, Ul-Haq Z, Evans 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* 2013; **8**(11): e75362.

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

40. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010; **26**(17): 2190-2191.

41. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* 2015; **11**(4): e1004219.

42. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. FUMA: Functional mapping and annotation of genetic associations. *bioRxiv* 2017.

43. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C, *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015; **47**(3): 291-295.

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

45. Pers TH, Karjalainen JM, Chan Y, Westra HJ, Wood AR, Yang J, *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun* 2015; **6**: 5890.

46. Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, *et al.* The NCBI dbGaP database of genotypes and phenotypes. *Nat Genet* 2007; **39**(10): 1181-1186.

47. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012; **22**(9): 1790-1797.

48. Consortium GT. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015; **348**(6235): 648-660.

49. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. *Bioinformatics* 2015; **31**(9): 1466-1468.

50. Wray NR, Sullivan PF. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *bioRxiv* 2017.

51. Genetics of Personality C, de Moor MH, van den Berg SM, Verweij KJ, Krueger RF, Luciano M, *et al*. Meta-analysis of Genome-wide Association Studies for Neuroticism, and the Polygenic Association With Major Depressive Disorder. *JAMA Psychiatry* 2015; **72**(7): 642-650.

52. Team RC. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, Austria, 2015.

53. Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* 2011; **88**(3): 294-305.

54. Chen GB. Estimating heritability of complex traits from genome-wide association studies using IBS-based Haseman-Elston regression. *Front Genet* 2014; **5**: 107.

55. Hirschfeld RM, Williams JB, Spitzer RL, Calabrese JR, Flynn L, Keck PE, Jr., *et al*. Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire. *Am J Psychiatry* 2000; **157**(11): 1873-1875.

56. Raine A. The SPQ: a scale for the assessment of schizotypal personality based on DSM-III-R criteria. *Schizophr Bull* 1991; **17**(4): 555-564.

57. consortium C. Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* 2015; **523**(7562): 588-591.

58. Hyde CL, Nagle MW, Tian C, Chen X, Paciga SA, Wendland JR, *et al*. Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat Genet* 2016; **48**(9): 1031-1036.

59. Zuo L, Tan Y, Zhang X, Wang X, Krystal J, Tabakoff B, *et al*. A New Genomewide Association Meta-Analysis of Alcohol Dependence. *Alcohol Clin Exp Res* 2015; **39**(8): 1388-1395.

60. Cook MN, Baker JA, Heldt SA, Williams RW, Hamre KM, Lu L. Identification of candidate genes that underlie the QTL on chromosome 1 that mediates genetic differences in stress-ethanol interactions. *Physiol Genomics* 2015; **47**(8): 308-317.

61. Zuo L, Wang K, Wang G, Pan X, Zhang X, Zhang H, *et al*. Common PTP4A1-PHF3-EYS variants are specific for alcohol dependence. *Am J Addict* 2014; **23**(4): 411-414.

62. Zuo L, Zhang X, Deng HW, Luo X. Association of rare PTP4A1-PHF3-EYS variants with alcohol dependence. *J Hum Genet* 2013; **58**(3): 178-179.

63. Zuo L, Zhang CK, Wang F, Li CS, Zhao H, Lu L, *et al*. A novel, functional and replicable risk gene region for alcohol dependence identified by genome-wide association study. *PLoS One* 2011; **6**(11): e26726.

64. Hontelez S, Karthaus N, Looman MW, Ansems M, Adema GJ. DC-SCRIPT regulates glucocorticoid receptor function and expression of its target GILZ in dendritic cells. *J Immunol* 2013; **190**(7): 3172-3179.

65. Lopez-Garcia J, Periyasamy M, Thomas RS, Christian M, Leao M, Jat P, *et al*. ZNF366 is an estrogen receptor corepressor that acts through CtBP and histone deacetylases. *Nucleic Acids Res* 2006; **34**(21): 6126-6136.

66. Miyamoto-Sato E, Fujimori S, Ishizaka M, Hirai N, Masuoka K, Saito R, *et al*. A comprehensive resource of interacting protein regions for refining human transcription factor networks. *PLoS One* 2010; **5**(2): e9289.

67. Aragam N, Wang KS, Pan Y. Genome-wide association analysis of gender differences in major depressive disorder in the Netherlands NESDA and NTR population-based samples. *J Affect Disord* 2011; **133**(3): 516-521.

68. Fidalgo TM, da Silveira ED, da Silveira DX. Psychiatric comorbidity related to alcohol use among adolescents. *Am J Drug Alcohol Abuse* 2008; **34**(1): 83-89.

69. Lopez B, Turner RJ, Saavedra LM. Anxiety and risk for substance dependence among late adolescents/young adults. *J Anxiety Disord* 2005; **19**(3): 275-294.

70. Low NC, Lee SS, Johnson JG, Williams JB, Harris ES. The association between anxiety and alcohol versus cannabis abuse disorders among adolescents in primary care settings. *Fam Pract* 2008; **25**(5): 321-327.

71. Rutledge PC, Sher KJ. Heavy drinking from the freshman year into early young adulthood: the roles of stress, tension-reduction drinking motives, gender and personality. *J Stud Alcohol* 2001; **62**(4): 457-466.

72. Schmidt NB, Buckner JD, Keough ME. Anxiety sensitivity as a prospective predictor of alcohol use disorders. *Behav Modif* 2007; **31**(2): 202-219.

73. Varlinskaya EI, Kim EU, Spear LP. Chronic intermittent ethanol exposure during adolescence: Effects on stress-induced social alterations and social drinking in adulthood. *Brain Res* 2017; **1654**(Pt B): 145-156.

74. Torigata K, Daisuke O, Mukai S, Hatanaka A, Ohka F, Motooka D, *et al*. LATS2 Positively Regulates Polycomb Repressive Complex 2. *PLoS One* 2016; **11**(7): e0158562.

75. Shen EY, Jiang Y, Javidfar B, Kassim B, Loh YE, Ma Q, *et al*. Neuronal Deletion of Kmt2a/Mll1 Histone Methyltransferase in Ventral Striatum is Associated with Defective Spike-Timing-Dependent Striatal Synaptic Plasticity, Altered Response to Dopaminergic Drugs, and Increased Anxiety. *Neuropsychopharmacology* 2016; **41**(13): 3103-3113.

76. Gage SH, Jones HJ, Taylor AE, Burgess S, Zammit S, Munafo MR. Investigating causality in associations between smoking initiation and schizophrenia using Mendelian randomization. *Sci Rep* 2017; **7**: 40653.

77. Hartz SM, Horton AC, Hancock DB, Baker TB, Caporaso NE, Chen LS, *et al*. Genetic correlation between smoking behaviors and schizophrenia. *Schizophr Res* 2017.

78. Dudbridge F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet* 2013; **9**(3): e1003348.

79. Polderman TJ, Benyamin B, de Leeuw CA, Sullivan PF, van Bochoven A, Visscher PM, *et al*. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat Genet* 2015; **47**(7): 702-709.

80. Laucht M, Treutlein J, Blomeyer D, Buchmann AF, Schmid B, Becker K, *et al*. Interaction between the 5-HTTLPR serotonin transporter polymorphism and environmental adversity for mood and anxiety psychopathology: evidence from a high-risk community sample of young adults. *Int J Neuropsychopharmacol* 2009; **12**(6): 737-747.

81. Avery BM, Vrshek-Schallhorn S. Nonsynonymous HTR2C polymorphism predicts cortisol response to psychosocial stress I: Effects in males and females. *Psychoneuroendocrinology* 2016; **70**: 134-141.

82. Vrshek-Schallhorn S, Stroud CB, Mineka S, Zinbarg RE, Adam EK, Redei EE, *et al*. Additive genetic risk from five serotonin system polymorphisms interacts with interpersonal stress to predict depression. *J Abnorm Psychol* 2015; **124**(4): 776-790.

83. Doornbos B, Dijck-Brouwer DA, Kema IP, Tanke MA, van Goor SA, Muskiet FA*, et al*. The development of peripartum depressive symptoms is associated with gene polymorphisms of MAOA, 5-HTT and COMT. *Prog Neuropsychopharmacol Biol Psychiatry* 2009; **33**(7): 1250-1254.

84. Shimasaki A, Kondo K, Saito T, Esaki K, Otsuka Y, Mano K*, et al.* A genetic variant in 12q13, a possible risk factor for bipolar disorder, is associated with depressive state, accounting for stressful life events. *PLoS One* 2014; **9**(12): e115135.

85. Dunn EC, Wiste A, Radmanesh F, Almli LM, Gogarten SM, Sofer T*, et al.* Genome-Wide Association Study (Gwas) and Genome-Wide by Environment Interaction Study (Gweis) of Depressive Symptoms in African American and Hispanic/Latina Women. *Depress Anxiety* 2016; **33**(4): 265-280.

FIGURE LEGENDS

Figure 1. Manhattan plots showing SS associations. Manhattan plots of the GWIS from **A)** UKB, **B)** GS:SFHS and **C)** sample size weighted meta-analysis of UKB and GS:SFHS. The x-axis is chromosomal position and y-axis is the *p* value (-log10 *p* value) of association with SS-effect. Suggestive genome-wide significance threshold ($p = 1x10^{-5}$) is shown by solid line at y = 5. Genes or closest gene up- and down-stream from SNP position (/) are annotated. "-": No gene within 100kb of the SNP.

Figure 2. Miami plots showing comparison between association profile between SS and MDD main additive effects. Miami plots from meta-analysis filter at $p = 1x10^{-3}$: **A)** filtering for SS *p* values (•), **B)** filtering for MDD *p* values (×). The x-axis is chromosomal position and y-axis is the *p* value (-log10 *p* value) of association with SS (up; red dots) and MDD *p* value (down; blue crosses). Dot line: genome-wide suggestive threshold ($p = 1x10^{-5}$) at the filtered effect; dashed lines: $p = 0.01$ and 0.05 at unfiltered effect.

Figure 3. MDD is best predicted using multiple PRS. MDD risk explained (R^2 coefficient (%); top bar values) on the liability scale by each PRS; weighted by main additive and SS-effects independently and combined. **A)** GS:SFHS: there is an increment on MDD risk prediction from adding PRS_{SS} to PRS_{MDD} model of 64.3% and 35.8% when adding PRS_{SS} to both MDD and neuroticism PRS together. **B)** UKB: there is an increment on MDD risk prediction from adding PRS_{SS} to PRS_{MDD} model of 17.1% and 11.2% when combining PRS_{SS} with both MDD and neuroticism PRS. **C & D)** Replication of A & B fitting PRS weighted by main additive effects for MDD and neuroticism using summary statistics from worldwide consortiums (i.e. PGC &GPC). Significance codes: *p* values *** < 0.001 < ** < 0.01 < * < 0.05.

Table 1. Top 25 SNPs from meta-analysis of GWISs.

Rank	CHR	SNP	BP	A1	A2	N	Z-score	*p*	GENE	POSITION
1	6	rs319924	64487247	A	G	30178	5.024	$5.05x10^{-7}$	EYS	Intronic
2	5	rs246565	71809247	A	C	30149	-4.684	$2.82x10^{-6}$	ZNF366	5998bp 5'
3	10	rs2265265	24854876	A	C	29753	4.604	$4.15x10^{-6}$	KIAA1217 / ARHGAP21	18104bp 3' / 17662bp 3'
4	6	rs1057530	64427095	A	G	30111	-4.556	$5.21x10^{-6}$	PHF3/EYS	1677bp 3' / 2781bp 3'
5	16	rs7199110	78790765	A	G	29545	-4.553	$5.29x10^{-6}$	WWOX	Intronic
6	6	rs10485358	64386060	A	G	30030	-4.546	$5.46x10^{-6}$	PHF3	Intronic
7	12	rs10778077	101193988	A	G	30031	4.54	$5.62x10^{-6}$	ANO4	Intronic
8	5	rs13358894	71803446	A	G	30239	4.527	$5.99x10^{-6}$	ZNF366	197bp 5'
9	10	rs2256220	24856314	A	G	29172	-4.524	$6.06x10^{-6}$	KIAA1217 / ARHGAP21	19542bp 3' / 16224bp 3'
10	10	rs3762096	98136250	A	G	29848	-4.521	$6.15x10^{-6}$	TLL2	Intronic
11	11	rs2221540	132716369	A	G	30247	-4.492	$7.05x10^{-6}$	OPCML	Intronic
12	5	rs10043659	71781839	A	G	29972	4.483	$7.37x10^{-6}$	ZNF366	Intronic
13	12	rs10778078	101195088	A	G	30074	-4.45	$8.58x10^{-6}$	ANO4	Intronic
14	9	rs10987199	128968987	A	G	30164	-4.442	$8.91x10^{-6}$	LOC101929116	63416bp 3'
15	5	rs10042132	71789021	A	G	30097	-4.416	$1.01x10^{-5}$	ZNF366	Intronic
16	11	rs10894606	132671611	A	C	30136	-4.404	$1.06x10^{-5}$	OPCML	Intronic
17	12	rs7295089	2440464	A	G	29966	4.372	$1.23x10^{-5}$	CACNA1C	Intronic
18	5	rs9293292	71696942	A	G	27739	-4.351	$1.36x10^{-5}$	PTCD2/ZNF366	41762bp 3' / 42292bp 3'
19	15	rs3097437	27872136	A	C	30237	4.346	$1.38x10^{-5}$	GABRG3	93762bp 3'
20	9	rs1999377	11919732	A	G	30164	4.344	$1.40x10^{-5}$	-	Intragenic
21	5	rs6862221	71754962	A	G	30158	4.342	$1.41x10^{-5}$	ZNF366	Intronic
22	5	rs9293289	71683885	A	G	29118	-4.323	$1.54x10^{-5}$	PTCD2/ZNF366	28705bp 3' / 55349bp 3'
23	11	rs4575282	132719646	A	G	30247	-4.313	$1.61x10^{-5}$	OPCML	Intronic
24	9	rs2417008	128970219	A	G	30215	-4.3	$1.71x10^{-5}$	LOC101929116	62184bp 3'
25	9	rs7021461	128972210	A	C	30214	4.299	$1.72x10^{-5}$	LOC101929116	60193bp 3'

Position position of the SNP respect to closest gene transcripts within 100kb (including UTRs) from 5 prime (5') or 3prime (3').

0

Table 2. MDD risk prediction at best fits.

UKB predicting on GS:SFHS

Weighted effect	Best fit threshold	# SNPs	R^2 (%)[a]	R^2 (%)[b]	*p*
SS	0.004	1 313	0.175	0.243	$1.8x10^{-3}$
MDD	0.125	25 833	0.278	0.386	$8.3x10^{-5}$
EPQN	$1x10^{-3}$	863	0.245	0.34	$2.2x10^{-4}$
PGC2 & GPC predicting on GS:SFHS					
PGC2 MDD	0.227	41 917	0.888	1.233	$1.9x10^{-12}$
GPC EPQN	0.013	4 234	0.101	0.14	0.018

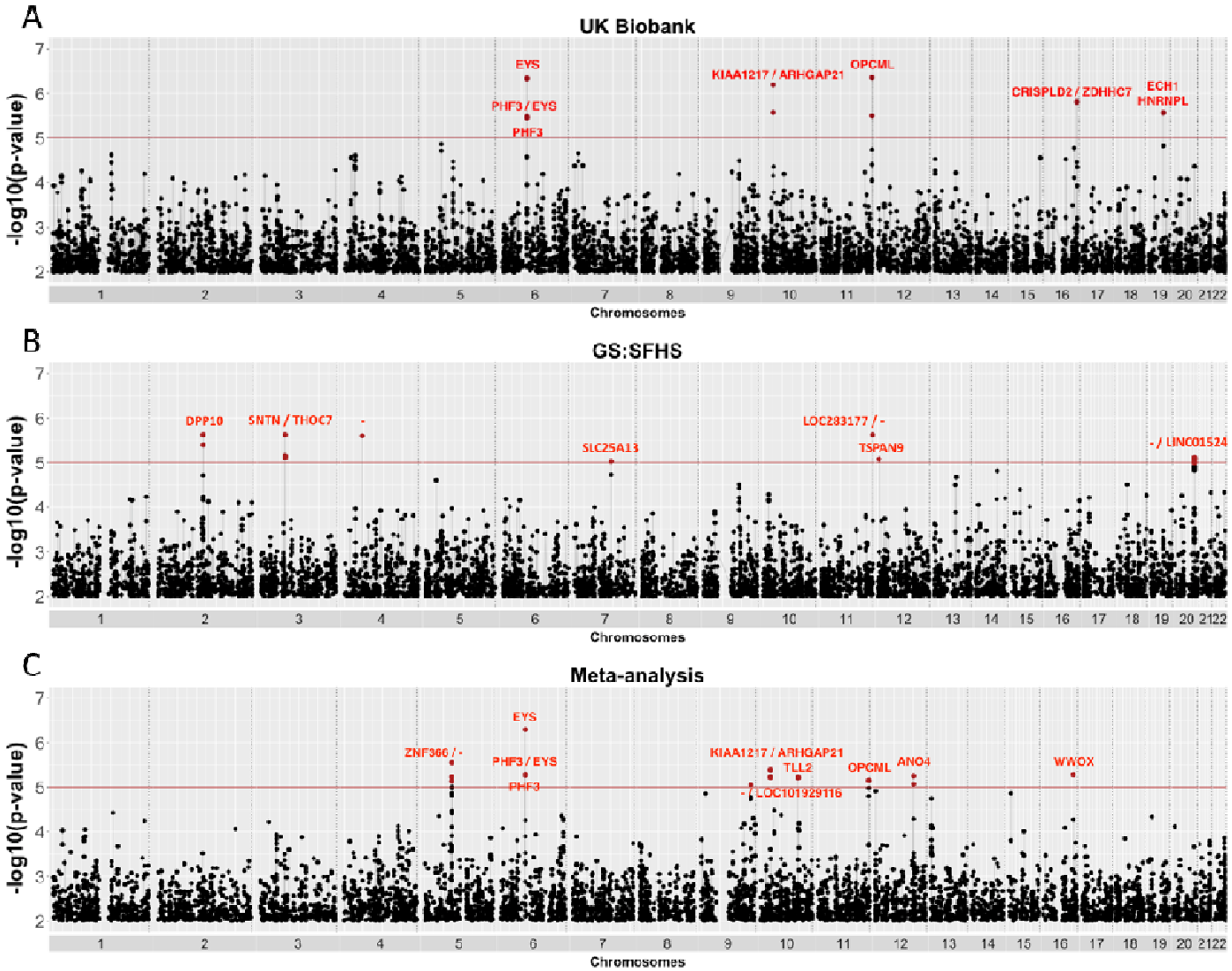
GS:SFHS predicting on UKB

Weighted effect	Best fit threshold	# SNPs	R^2 (%)[a]	R^2 (%)[b]	*p*
SS	$1.15x10^{-4}$	42	0.024	0.037	0.04
MDD	0.033	8 135	0.135	0.211	$1.1x10^{-6}$
EPQN	0.07	15 613	0.114	0.178	$7.3x10^{-6}$
PGC2 & GPC predicting on UKB					
PGC2 MDD	0.095	22 005	0.818	1.281	$< 2.2x10^{-16}$
GPC EPQN	$6x10^{-3}$	2 182	0.067	0.105	$5.6x10^{-4}$

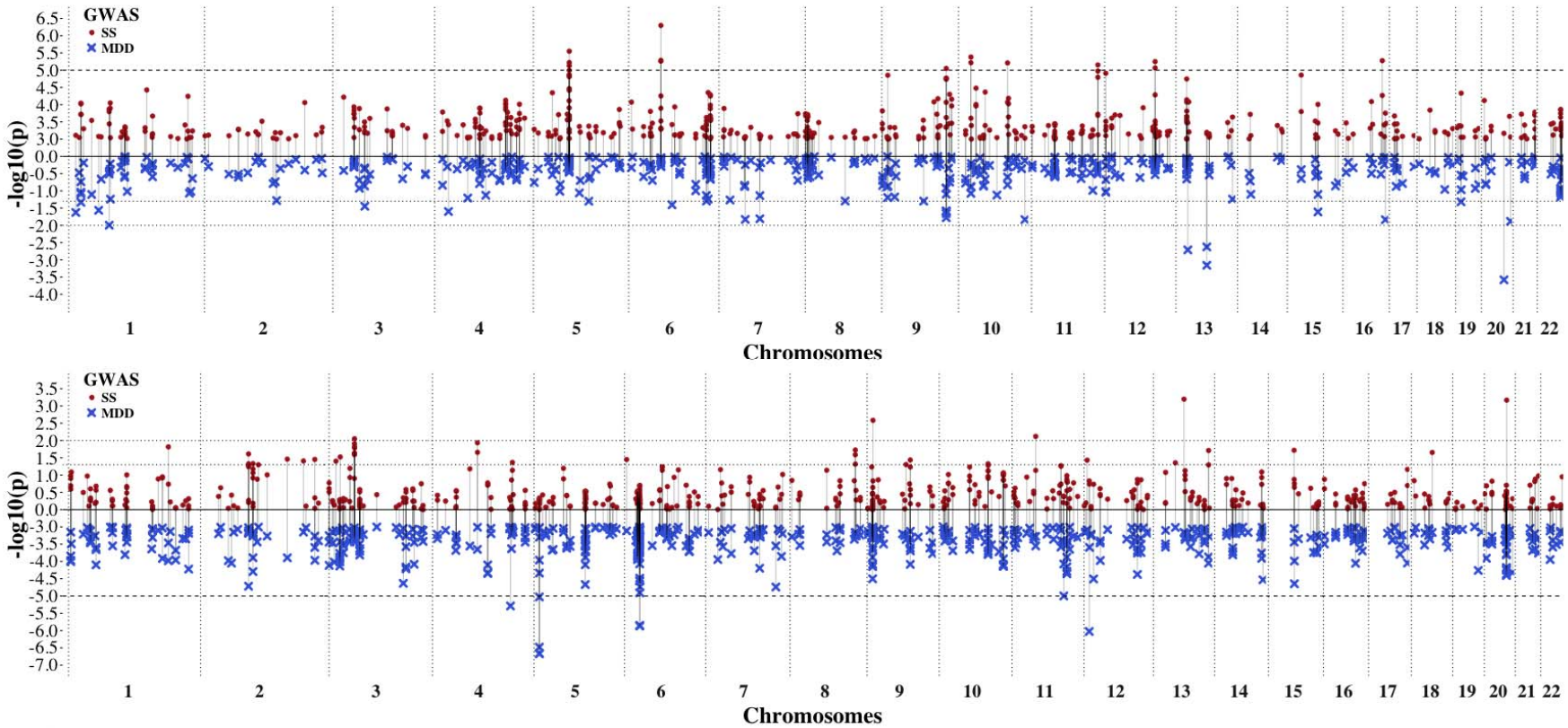
SS stress-sensitivity, *MDD* major depressive disorder, *EPQN* neuroticism score, [a]Nagelkerke's R^2 at observed scale, [b]R^2 on the liability scale

1 **Figure 1**

2

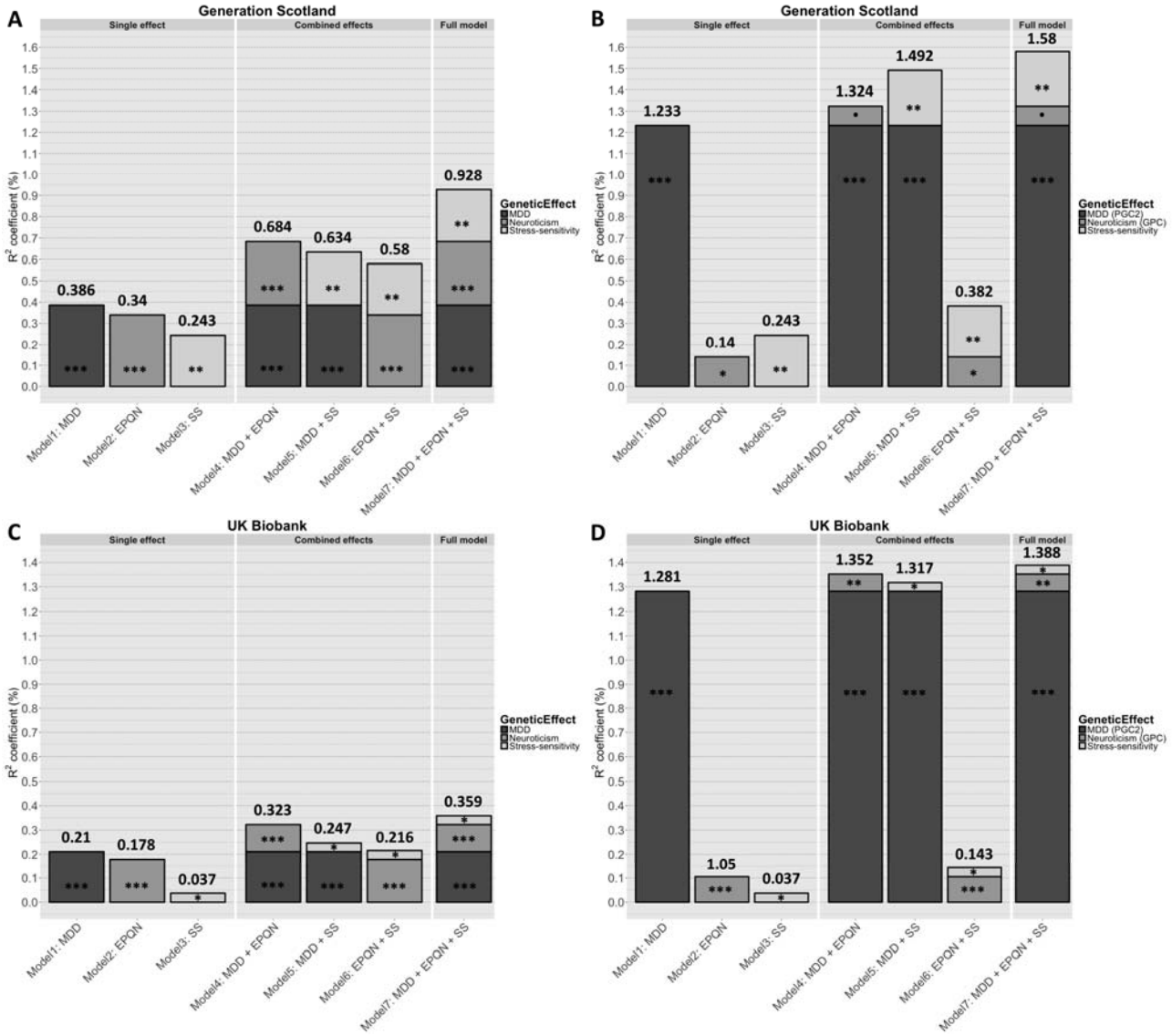

A
UK Biobank
EYS
PHF3 / EYS
PHF3
KIAA1217 / ARHGAP21
OPCML
CRISPLD2 / ZDHHC7
ECH1
HNRNPL
-log10(p-value)
Chromosomes
B
GS:SFHS
DPP10
SNTN / THOC7
SLC25A13
LOC283177 / -
TSPAN9
- / LINC01524
-log10(p-value)
Chromosomes
C
Meta-analysis
EYS
ZNF366 / -
PHF3 / EYS
PHF3
KIAA1217 / ARHGAP21
TLL2
- / LOC101929116
OPCML
ANO4
WWOX
-log10(p-value)
Chromosomes

3 **Figure 2**


GWAS
SS
MDD
-log10(p)
Chromosomes
GWAS
SS
MDD
-log10(p)
Chromosomes

4

5 **Figure 3**


A
Generation Scotland
Single effect
Combined effects
Full model
R² coefficient (%)
0.386
0.34
0.243
0.684
0.634
0.58
0.928
GeneticEffect
MDD
Neuroticism
Stress-sensitivity
Model1: MDD
Model2: EPQN
Model3: SS
Model4: MDD + EPQN
Model5: MDD + SS
Model6: EPQN + SS
Model7: MDD + EPQN + SS
B
Generation Scotland
1.233
0.14
0.243
1.324
1.492
0.382
1.58
GeneticEffect
MDD (PGC2)
Neuroticism (GPC)
Stress-sensitivity
C
UK Biobank
0.21
0.178
0.037
0.323
0.247
0.216
0.359
D
UK Biobank
1.281
1.05
0.037
1.352
1.317
0.143
1.388