# Replication of the diathesis-stress model for depression in Generation Scotland.

3

Running title: The diathesis-stress model for depression in Generation Scotland.

4 5 6

Aleix Arnau-Soler<sup>1</sup>, Mark J. Adams<sup>2</sup>, Toni-Kim Clarke<sup>2</sup>, Donald J. MacIntyre<sup>2</sup>, Keith

7 Milburn<sup>3</sup>, Lauren Navrady<sup>2</sup>, Generation Scotland<sup>4</sup>, Major Depressive Disorder Working 8 Group of the Psychiatric Genomics Consortium<sup>5</sup>, Caroline Hayward<sup>6</sup>, Andrew

9 McIntosh<sup>2,7</sup>, Pippa A. Thomson<sup>1,7</sup>

10

<sup>1</sup> Medical Genetics Section, University of Edinburgh, Centre for Genomic and Experimental Medicine and
 MRC Institute of Genetics and Molecular Medicine, Edinburgh, UK.

- 13 <sup>2</sup> Division of Psychiatry, Deanery of Clinical Sciences, University of Edinburgh, Royal Edinburgh Hospital,
- 14 Morningside Park, Edinburgh EH10 5HF, UK.
- 15 <sup>3</sup> Health Informatics Centre, University of Dundee, UK.

<sup>4</sup> A collaboration between the University Medical School and NHS in Aberdeen, Dundee, Edinburgh and
 Glasgow, Scotland, UK.

18  $\,$  ^ 5 For a full list of MDD working group of the PGC investigators, see the Supplementary Material

<sup>6</sup> Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine,
 University of Edinburgh, Edinburgh, UK.

- 21 <sup>7</sup> Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK.
- 22

23 Correspondence should be addressed to: P. Thomson (Pippa.Thomson@ed.ac.uk) or A. Arnau-

- 24 Soler (aleix.arnau.soler@igmm.ed.ac.uk)
- 25
- 26 Centre for Genomic and Experimental Medicine
- 27 Institute of Genetics and Molecular Medicine
- 28 University of Edinburgh
- 29 Western General Hospital
- 30 Crewe Road South
- 31 Edinburgh, EH4 2XU, UK
- 32 Phone: +44(0)1316518756
- 33
- 34

#### 35 ABSTRACT

36 Depression has well-established influences from both genetic and environmental 37 factors. A popular theory of depression aetiology in psychiatry and psychology is the 38 *diathesis-stress* theory, which assumes a multiplicative gene-by-environment 39 interaction (GxE) effect on risk. Recently, *Colodro-Conde et al* empirically tested it, 40 reporting GxE effects additively contributing to liability.

We replicate that study on an independent sample of 4 919 unrelated individuals who reported stressful life events (SLE) over the 6 months immediately before a selfreported measure of depressive symptoms, and test for sex-specific differences.

44 We identified significant but weak positive GxE using the total number of SLE reported in the full cohort ( $R^2 = 0.08\%$ ,  $p = 4.87 \times 10^{-2}$ ) and in women ( $R^2 = 0.19\%$ ,  $p = 1.66 \times 10^{-2}$ ) 45 <sup>2</sup>), but not in men (p > 0.05). We also detected significant GxE effects, but only in 46 47 women, when SLE were split into those in which the respondent may play an active role ("dependent" SLE,  $R^2 = 0.15\%$ ,  $p = 3.85 \times 10^{-2}$ ) or a passive role ("independent" 48 SLE,  $R^2 = 0.16\%$ ,  $p = 3.32 \times 10^{-2}$ ). Further, in women who experienced no SLE, the 49 50 diathesis effect showed a protective effect ( $\beta = -0.061$ , s.e. = 0.029, p = 0.037), 51 suggesting a possible role of genetic plasticity in risk variants.

52 Our study replicates Colodro-Conde *et al.*, reinforcing the presence of additional risk in 53 the aetiology of depression due to GxE effects. Furthermore, these results support 54 possible differences between sexes and effects of SLE subtypes. However, more power 55 is required to robustly replicate these findings.

56

bioRxiv preprint doi: https://doi.org/10.1101/389494; this version posted August 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

# 57 **INTRODUCTION**

58 Adversity faced during stressful life events (SLE) has been consistently recognized as a 59 determinant of depressive symptoms with many studies reporting significant 60 associations between SLE and major depressive disorder  $(MDD)^{1-7}$ . Some studies 61 suggest that severe adversity is present before the onset of illness in over 50% of individuals with depression<sup>8</sup> and may characterize a subtype of cases<sup>9</sup>. However, some 62 individuals facing severe stress never present symptoms of depression<sup>10</sup>. This has led 63 64 to a proposal that interaction between stress and an individual's vulnerability, or 65 diathesis, is a key element in the development of depressive symptoms. Such 66 vulnerability can be conceived as a set of biological factors that predispose to illness. 67 This idea was first conceptualized into the *diathesis-stress* model to explain the development of schizophrenia back in the 1960s<sup>11</sup>. Later in the 1980s it was used to 68 explain the origins of depression<sup>12-14</sup> and since then several *diathesis-stress* models 69 70 have been applied to many psychopathologies<sup>15-19</sup>. The *diathesis-stress* model proposes that a latent *diathesis* may be activated by stress before psychopathological 71 72 symptoms manifest. Some levels of *diathesis* to illness are present in everybody, with a 73 threshold over of which they will present symptoms. Exceeding such illness threshold 74 depends on the interaction between *diathesis* and the degree of adversity faced in SLE, 75 which increase the liability of depression beyond the combined additive effects alone<sup>15</sup>. Inherent genetic risk factors can be conceived as a genetic *diathesis*. Thus, this 76 77 genetically driven effect produced by the *diathesis-stress* interaction can be seen as a 78 gene-by-environment interaction (GxE).

MDD is characterized by a highly polygenic architecture composed of common variants
 with small effect and/or rare variants<sup>20</sup>. Therefore, GxE interactions in depression are

81 also expected to be highly polygenic. In recent years, with the increasing success of 82 GWAS, GxE studies in depression have shifted towards hypothesis-free genome-wide and polygenic approaches that capture liability to depression using molecular data<sup>21-29</sup>. 83 84 Recent advances in genomics and the massive effort from national institutions to 85 collect genetic, clinical and environmental data on large population-based samples 86 now provide an opportunity to empirically test the diathesis-stress model in 87 depression. A novel paradigm to quantify genetic *diathesis* into a single genetic 88 measure, to study GxE effects with more predictive power than any single variant, is the construction of polygenic risk scores (PRS)<sup>30-33</sup>. PRS are genetic indicators of the 89 90 aggregate effect from risk alleles carried by an individual weighted by their allelic 91 effect estimated from Genome-Wide Association Studies (GWAS).

92 This polygenic approach to assess the *diathesis-stress* model for depression has been 93 tested with childhood trauma<sup>21,23,29</sup> and SLE<sup>22,27,29</sup>, as measures of environmental 94 adversity.

Recently, Colodro-Conde et al.<sup>27</sup> provided a direct test of the diathesis-stress model for 95 96 recent SLE and depressive symptoms. In this study, Colodro-Conde et al. used PRS 97 weighted by the most recent genome-wide meta-analysis conducted by the Psychiatric 98 Genetics Consortium (PGC; N = 159 601), having a substantially larger sample size than 99 any discovery sample used previously, and measures of three environmental 100 exposures: lack of social support, "personal" SLE (PSLE), and "network" SLE. Colodro-101 Conde et al. reported a significant additive risk on liability to depression due to a GxE 102 effect in individuals who combine a high genetic predisposition to MDD and a high 103 number of reported PSLE, mainly driven by effects in women. No significant interaction 104 was reported in males. They found no significant interaction with "network" SLE or

105 social support. They concluded that the effect of stress on risk of depression was 106 dependent on an individual's *diathesis*, supporting the *diathesis-stress* theory. In 107 addition, they suggested possible sex-specific differences in the aetiology of 108 depression. However, Colodro-Conde et al. findings have not, to our knowledge, been 109 independently replicated. 110 In the present study we aim to replicate Colodro-Conde et al. and assess differences 111 between women and men in an independent sample of 4 919 unrelated white British 112 participants from a further longitudinal assessment from Generation Scotland: Scottish 113 Family Health Study, who along with a self-reported diagnosis of depressive 114 symptoms, self-reported SLE over the preceding 6 months.

bioRxiv preprint doi: https://doi.org/10.1101/389494; this version posted August 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## 115 MATERIALS AND METHODS

#### 116 **Sample description**

117 The present study was conducted using data available on 4 919 unrelated individuals 118 (mean age at questionnaire: 47.2, s.d. = 12.2, range 22-95; *females*: n = 2 990 - 60.8%, 119 mean age 56.1, s.d. = 12.4; males: n = 1 929 - 39.2%, mean age 58.7, s.d. = 11.8) from a 120 further longitudinal assessment from Generation Scotland: Scottish Family Health 121 Study (GS:SFHS; www.ed.ac.uk/generation-scotland/using-resources/scottish-family-122 health-study) funded by a Wellcome Trust Strategic Award "STratifying Resilience and Depression Longitudinally" (STRADL) reference 10436/Z/14/Z. STRADL<sup>34</sup> is a project 123 124 aimed to investigate the aetiology underlying depression by re-contacting participants 125 from GS:SFHS to collect new and updated mental health questionnaires covering a 126 wide range of psychiatric symptoms and SLE measures, among others. Further details on the recruitment procedure and GS:SFHS profile are described in detail elsewhere<sup>35-</sup> 127 128 <sup>39</sup>. In 2014, 21 525 GS:SFHS participants eligible for re-contact were sent self-reported 129 guestionnaires. 9 618 GS:SFHS re-contacted participants (44.7% response rate) agreed to provide new measures to STRADL mental health follow-up<sup>34</sup>. Those participants: 130 131 duplicated, with diagnoses of bipolar disorder or with missing data on reported SLE, 132 population outliers, with sex discrepancies, or with more than 2% missing genotypes, 133 were removed. SNPs with more than 2% genotype missingness, Hardy-Weinberg Equilibrium test  $p < 1 \times 10^{-6}$ , or minor allele frequency lower than 1%, were excluded. 134 135 After applying quality control, individuals were filtered by degree of relatedness (pi-hat < 0.05) using PLINK v1.9<sup>40</sup>, maximizing retention of those participants reporting higher 136 137 numbers of SLE (see phenotype assessment below). The final dataset comprised 4 919 138 unrelated individuals and 560 351 SNPs after quality control. 20 principal components 139 were calculated for the final "full cohort" dataset (N = 4.919), "females" dataset (N = 2140 990) and "males" dataset (N = 1 929). All participants provided written consent. All 141 components of GS:SFHS and STRADL obtained ethical approval from the Tayside 142 Committee on Medical Research Ethics on behalf of the National Health Service (reference 05/s1401/89). GS:SFHS & STRADL data is available to researchers on 143 144 application Scotland to the Generation Access Committee 145 (access@generationscotland.org).

146

#### 147 **Phenotype assessment**

Individual's current depressive symptoms were assessed using the 28-item scaled 148 version of The General Health Questionnaire (GHQ)<sup>41,42</sup>. GHQ is a reliable and validated 149 150 psychometric screening tool to detect common psychiatric and non-psychotic conditions (GHQ Cronbach alpha coefficient: 0.82 - 0.86)<sup>43</sup>. GHQ assesses symptoms 151 152 over the last two weeks through 28 yes/no questions (items). At same time, each 153 symptom/item is rated on a four-point Likert scale from 0 to 3 to assess its degree or 154 severity (e.g. Have you recently felt that life is entirely hopeless? "Not at all", "No more than usual", "Rather more than usual", "Much more than usual"), resulting on an 84-155 156 point scale depression score. Depression score was log transformed to reduce the 157 effect of positive skew/provide a better approximation to a normal distribution. For a 158 better interpretation, *depression* score was scaled to a mean of 0 when required (see 159 Figure 3).

Data from a brief life events questionnaire (BLEQ), based on the List of Threating Experiences (LTE)<sup>44</sup> and self-reported by STRADL participants, was used to construct a measure of common SLE over the previous 6 months. LTE is a reliable psychometric

device to measure psychological "stress"<sup>45,46</sup>. The BLEQ consists of a 12-item 163 164 questionnaire to assess SLE with considerable long-term contextual effects (e.g. Over 165 last 6 months, did you have a serious problem with a close friend, neighbour or 166 relatives?). A final score reflecting the total number of SLE (TSLE) ranging from 0 to 12 167 was constructed by summing yes responses. Additionally, TSLE was split into two 168 categories based on BLEQ-items measuring SLE in which the individual may play and 169 active role, and therefore in which the SLE is influenced by genetic factors and thus 170 subjected to be "dependent" on an individual's own behaviour or symptoms (DSLE; 6 171 BLEQ-items, e.g. a serious problem with a close friend, neighbour or relatives may be 172 subject to a respondent's own behaviour), or SLE that are not influenced by genetic 173 factors, likely to be "independent" on a participant's own behaviour (ISLE; 5 BLEQ-174 items, e.g. a serious illness, injury or assault happening to a close relative is potentially independent of a respondent's own behaviour)<sup>44,47</sup>. The BLEQ-item "Did you/your wife 175 176 or partner give birth?" was excluded from this categorization. In addition, for each SLE 177 measure, individuals were categorized in three categories based on the amount of SLE 178 experienced (O SLE = "missing", 1 or 2 SLE = "low", and 3 or more SLE = "high").

179

# 180 **Polygenic profiling & statistical analysis**

Polygenic risk scores (PRS) were generated by PRSice<sup>48</sup>, whose functionality relies entirely in PLINK v1.9<sup>40</sup>, and calculated using the genotype data of STRADL participants (i.e. target sample) and summary statistics for MDD from the PGC-MDD2 GWAS release (July 2016, discovery sample) used by Colodro-Conde *et al.* with the added contribution from QIMR cohort and the exclusion of Generation Scotland participants, resulting in summary statistics for MDD derived from a sample of 50 455 cases and 105

# 187 411 controls.

Briefly, PRSice removed strand-ambiguous SNPs and clump-based pruned ( $r^2 = 0.1$ , 188 189 within a 10Mb window) our target sample to obtain the most significant independent 190 SNPs (n) in approximate linkage equilibrium. Independent risk alleles were then 191 weighted by the allelic effect sizes estimated in the independent discovery sample 192 were aggregated into PRS. PRS were generated for eight p thresholds (p thresholds: < 5  $x 10^{-8}$ ,  $< 1 \times 10^{-5}$ , < 0.001, < 0.01, < 0.05, < 0.1, < 0.5, <=1) determined by the discovery 193 sample. (See Supplementary Table 1 for summary of PRS). 194 Following Colodro-Conde *et al.*, covariates (i.e. age, age<sup>2</sup>, sex, age-by-sex and age<sup>2</sup>-by-195

sex interactions, and 20 principal components) were regressed from PRS (PRS') and SLE
 scores (i.e. TSLE, DSLE and ISLE; SLEs') before fitting models in GCTA 1.26.0<sup>49</sup> to guard
 against confounding influences on the PRS-by-SLEs interactions<sup>50</sup>. PRS' and SLEs' were

standardized to a mean of 0 and a standard deviation of 1. A genetic relatedness

200 matrix (GRM) was calculated for each dataset (i.e. *full cohort, males* and *females*) using

201 GCTA.

202 Mixed linear models using the GRM were used to estimate the variance in *depression* 

score explained by PRS', SLEs' and their interaction, and stratified by sex.

204 The mixed linear model used to assess the effects of PRS' is as follows:  $Depression = \beta_0 + \beta_1 PRS' + GRM + Covariates$ 

205 Mixed linear models used to assess the effect of the stressors (=SLEs') are as follows:  $Depression = \beta_0 + \beta_1 TSLE' + GRM + Covariates$   $Depression = \beta_0 + \beta_1 DSLE' + GRM + Covariates$   $Depression = \beta_0 + \beta_1 ISLE' + GRM + Covariates$ 

- 206 The Mixed linear models (i.e. the diathesis-stress model) used to assess PRS'
- 207 interactions with SLEs' are as follows:

Depression =  $\beta_0 + \beta_1 PRS' + \beta_2 TSLE' + \beta_3 PRS' xTSLE' + GRM + Covariate \square$ Depression =  $\beta_0 + \beta_1 PRS' + \beta_2 DSLE' + \beta_3 PRS' xDSLE' + GRM + Covariate \square$  $Depression = \beta_0 + \beta_1 PRS' + \beta_2 ISLE' + \beta_3 PRS' xISLE' + GRM + Covariate \square$ Covariates fitted in the models above were age, age<sup>2</sup>, sex, age-by-sex, age<sup>2</sup>-by-sex and 208 20 principal components. Sex and its interactions (age-by-sex and age<sup>2</sup>-by-sex) were 209 210 not included when stratifying by sex. All parameters from the models were estimated 211 using GCTA and the significance of the effect ( $\beta$ ) from fixed effects assessed using 212 Student's *t*-test at *p*-value threshold = 0.05. 213 Using linear regressions we applied a least squares approach to assess PRS' effects on 214 depression score in each SLE category (i.e. "missing stress", "low stress" "high stress") 215 where significant GxE were detected (significance at p-value < 0.05).

216

## 217 **RESULTS**

Depression PRS' significantly predicted the *depression* score in the whole sample ( $\beta$  = 218 0.080, s.e. = 0.014,  $p = 7.53 \times 10^{-9}$ ) explaining 0.64% of the variance at its best p 219 220 threshold (= 0.1; see Figure 1a). Stratifying by sex, PRS' significantly predicted the 221 depression score in both sexes, explaining 0.59% in men and 0.67% in women (men: pthreshold = 0.1,  $\beta$  = 0.077, s.e. = 0.022, p = 2.09 x 10<sup>-4</sup>; women: p-threshold = 0.1,  $\beta$  = 222 0.082, s.e. = 0.018,  $p = 4.93 \times 10^{-6}$ ; see Figure 1a). Self-reported SLE over the last 6 223 224 months (TSLE, mean = 1.3 SLE), significantly predicted symptoms of depression 225 (depression score) in the whole sample and stratified by sex: full cohort: variance explained = 4.91%,  $\beta$  = 0.222, s.e. = 0.014, p = 9.98 x 10<sup>-59</sup>; men: 4.19%,  $\beta$  = 0.205, s.e. = 226 0.021,  $p = 2.23 \times 10^{-22}$ ; women: 5.33%,  $\beta = 0.231$ , s.e. = 0.018,  $p = 7.48 \times 10^{-38}$  (Figure 227 1b). However, the variance in *depression* score explained by the TSLE appears lower 228 229 than the variance explained by the measure of PSLE used in Colodro-Conde et al. 230 (12.9%). There was no significant difference in the direct effect of TSLE between 231 women and men. Although questions about "dependent" SLE (DSLE, mean = 0.4 SLE) 232 represented over 28% of the TSLE-items reported, the main effect of DSLE explained 233 approximately 93% of the amount of variance explained by TSLE (full cohort: variance explained = 4.56%,  $\beta$  = 0.212, s.e. = 0.014, p = 1.73 x 10<sup>-54</sup>; men: 3.74%,  $\beta$  = 0.193, s.e. = 234 0.021.  $p = 9.66 \times 10^{-21}$ ; women: 5.07%,  $\beta = 0.225$ , s.e. = 0.018.  $p = 8.09 \times 10^{-35}$ ; see 235 236 Figure 1b). "Independent" SLE (ISLE, mean = 0.85 SLE), which represented over 69% of 237 TSLE-items, explained approximately 57% of the amount of variance explained by TSLE (*full cohort*: variance explained = 2.80%,  $\beta$  = 0.167, s.e. = 0.014, *p* = 1.32 x 10<sup>-33</sup>; *men*: 238 2.44%,  $\beta = 0.156$ , s.e. = 0.022,  $p = 2.88 \times 10^{-13}$ ; women: 3.02%,  $\beta = 0.174$ , s.e. = 0.018, p239 = 5.20 x  $10^{-22}$ ; see Figure 1b). When DSLE and ISLE were combined together in a single 240

model, DSLE explained 3.34% of the variance of depressive score compared to 1.45%
of the variance being explained by ISLE. This suggests that DSLE have a greater effect
on liability to depressive symptoms than ISLE.

244 We detected significant, albeit weak, GxE effects on depression score as 245 conceptualized in the *diathesis-stress* model (see figure 2). The PRS interaction with TSLE was significant in the full cohort ( $\beta = 0.028$ , s.e. = 0.014,  $R^2 = 0.08\%$ ,  $p = 4.87 \times 10^{-10}$ 246 <sup>2</sup>) and slightly stronger in women ( $\beta = 0.044$ , s.e. = 0.018,  $R^2 = 0.19\%$ ,  $\rho = 1.66 \times 10^{-2}$ ; 247 see figure 2a). The best-fit threshold was much lower in women (*p*-threshold =  $1 \times 10^{-5}$ ) 248 249 compared to the full cohort (*p*-threshold = 0.01). GxE effects estimated in women and men at p-threshold = 1 x  $10^{-5}$  where significantly different (GxE\*sex p = 0.017), but not 250 251 at the best p-threshold in the full cohort (p-threshold = 0.01, GxE\*sex p = 0.32). In women, GxE effect with DSLE predicted *depression* score (*p*-threshold = 1 x  $10^{-5}$ ;  $\beta$  = 252 0.039, s.e. = 0.019,  $R^2 = 0.15\%$ ,  $p = 3.85 \times 10^{-2}$ ; see figure 2b), as did the GxE effect with 253 ISLE (*p*-threshold = 1 x  $10^{-5}$ ;  $\beta$  = 0.040, s.e. = 0.019, R<sup>2</sup> = 0.16%, *p* = 3.32 x  $10^{-2}$ ; see 254 255 figure 2c). No significant interaction was detected in men (best-fit p-threshold = 0.1) using either TSLE ( $\beta$  = 0.039, s.e. = 0.022, R<sup>2</sup> = 0.15%, p = 7.19 x 10<sup>-2</sup>; see figure 2a), 256 257 DSLE ( $\beta$  = 0.024, s.e. = 0.022, R<sup>2</sup> = 0.06%, p = 0.28; see figure 2b) or ISLE ( $\beta$  = 0.043, s.e. = 0.022,  $R^2$  = 0.18%, p = 5.47 x 10<sup>-2</sup>; see figure 2c). 258

Examination of the significant GxE detected by SLE categories (i.e. "missing", "low" and "high") suggests that in participants reporting TSLE the risk of depressive symptoms was higher when greater numbers of SLE were reported; whereas in participants who reported no SLE over the preceding 6 months, the risk of depressive symptoms was the same regardless of their *diathesis* risk ("missing": PRS'  $\beta$  = 0.021, s.e. = 0.022, *p* = 0.339; "low": PRS'  $\beta$  = 0.043, s.e. = 0.021, *p* = 0.039; "high": PRS'  $\beta$  = 0.142, s.e. = 0.039,

265	$p = 2.86 \times 10^{-4}$ ; see Table 1 and figure 3a). Although a similar pattern by increasing SLE
266	was observed in women (see Table 1, figure 3b and Supplementary figure 2), we only
267	detected a significant diathesis effect on depression score in those women who did not
268	experienced TSLE over the last 6 months. This effect was negative (PRS' $\beta$ = -0.061, s.e.
269	= 0.029, $p$ = 0.037, $R^2$ = 0.3%; see figure 3b), suggesting a protective effect of
270	increasing PRS' in those women reporting no SLE, and suggesting that the contributing
271	alleles may make an individual sensitive to both positive and negative environmental
272	effects (i.e. "plasticity alleles" rather than "risk alleles") <sup>51,52</sup> .

## 273 **DISCUSSION**

The findings reported in this study replicate those of Colodro-Conde *et al.* results in an independent sample from Generation Scotland of similar sample size and study design, and suggest possible sex-specific differences in genetic risk of MDD in response to "dependent" SLE.

We identified significant, albeit weak, GxE effect in liability to depression at the 278 population level ( $p = 4.87 \times 10^{-2}$ ) and in women ( $p = p = 1.66 \times 10^{-2}$ ), but not in men ( $p = 1.66 \times 10^{-2}$ ) 279 280 7.19 x  $10^{-2}$ ). Both Colodro-Conde *et al.* and our studies suggest that individuals with an 281 inherent genetic predisposition to MDD, reporting high number of recent SLE, are at 282 additional risk of depressive symptoms due to GxE effects, supporting the *diathesis*-283 stress theory. However, these interactions did not survive multiple testing correction and the power of these studies to draw robust conclusions remains limited<sup>53</sup>. With 284 285 increased power these studies could more accurately determine both the presence 286 and magnitude of a gene-by-environment interaction in depression.

In the full cohort, the total variance of depressive score explained by the PRS' main 287 288 effect and the GxE effect jointly was 0.34%, of which 0.07% was attributed to the GxE effect (p-threshold = 0.01; PRS  $p = 1.19 \times 10^4$ , GxE  $p = 4.87 \times 10^2$ ; both derived from 289 290 the full diathesis-model with TSLE); lower than the proportion of variance attributed to common SNPs (8.9%) in the full PGC-MDD analysis<sup>20</sup>. As Colodro-Conde et al. noted, 291 292 this result aligns with estimates from experimental organisms suggesting that around 20% of the heritability may be typically attributed to the effects of  $GxE^{54}$ , although it is 293 294 inconsistent with the majority of human traits with the potential exception of depression<sup>55</sup>. 295

We saw concordance between our results and those of Colodro-Conde *et al.*<sup>27</sup> 296 297 Consistent with PRS predicting PSLE in Colodro-Conde et al., PRS for MDD predicted 298 SLE in our study (see Supplementary Figure 1), although not at the p-threshold at 299 which significant GxE effects were detected. *Depression* score was positive correlated with self-reported SLE scores (TSLE:  $r^2 = 0.214$ , DSLE:  $r^2 = 0.211$ , ISLE:  $r^2 = 0.169$ ; all p < 100300  $2.2 \times 10^{-16}$ ). Genetic and phenotypic correlations, added to the usage of self-reported 301 302 screening to construct the measures fitted into the analysis, suggest possible 303 confounding i.e. self-report bias. Genetic correlation between SLE and MDD may be 304 driven by genetic factors which results in greater exposure to highly stressful/risky 305 environments, or via personality traits, such as neuroticism, which shows positive 306 associations with negative response to and greater reporting of negative life 307 events<sup>56,57</sup>. This genetic correlation implies that the SLE effect is unlikely to act solely 308 as an environmental risk factor, and genetic factors predisposing to MDD may also 309 increase exposure to/reporting of SLE. This hinders to interpret our findings as pure 310 GxE effects. To solve this limitation and assess this aspect, Colodro-Conde et al. broke 311 down the PSLE measure into SLE in which the individual may played an active role (PSLE-active) or a passive role (PSLE-passive)<sup>47,58</sup>. Equivalently, we split the TSLE 312 313 measure of 12-items into SLE that are potentially either SLE "dependent" (DSLE; 6-314 items) on a participant's own behaviour (i.e. therefore potentially driven by genetic factors) or not ("independent" SLE, ISLE; 5-items)<sup>44,47</sup>. PSLE-active and DSLE are 315 316 reported to be more heritable and have stronger associations with MDD than PSLEpassive or ISLE<sup>47,58,59</sup>. Thus, if the GxE detected were driven by ISLE it would point 317 318 towards a more pure GxE rather than a subtle genotype-by-genotype (GxG) interaction 319 or genotype-by-genotype-by-environment (GxGxE) interaction. We observed in our

study that although DSLE was significantly heritable ( $h_{SNP}^2 = 0.131$ , s.e. = 0.071, p =320 321 0.029) and reporting ISLE was not significantly heritable ( $h_{SNP}^2 = 0.000$ , s.e. = 0.072, p =322 0.5), in women significant GxE was seen for both. In Colodro-Conde et al., PSLE-active 323 explained most of the variance explained by PSLE. PSLE-passive score explained a 324 marginal amount of variance (0.77%) compared to the PSLE measure (10.5%). In 325 addition, conversely to our findings in which we identified significant GxE effects 326 within women using all three measures of SLE, and explaining similar amount of 327 variances (0.15% – 0.19%; see figure 2), Colodro-Conde et al. did not identify 328 significant GxE using PSLE-passive.

329 Despite the replicated findings supporting the *diathesis-stress* model, there are a few 330 differences between our study and Colodro-Conde et al. to consider. First, differences 331 in PRS profiling may have affected replication power. We used the same equivalent 332 PGC-MDD2 GWAS as discovery sample. However, whereas Colodro-Conde et al. 333 generated PRS in their target sample containing over 9.5M imputed SNP, in this study 334 we generated PRS in a target sample of over 560K genotyped SNPs (see 335 Supplementary table 1 for comparison). This potentially results in a less informative 336 PRS in our study, with less predictive power, although the variance explained by our 337 PRS' was slightly larger (0.64% vs. 0.46%). The size of the discovery sample is key to 338 constructing an accurate predictive PRS, but to exploit the most of the variants available may be an asset<sup>53</sup>. 339

Secondly, different screening tools were used to measure both current depression and recent environmental stressors across the two studies. Both studies transformed their data, using item response theory or by log-transformation, to improve the data distribution. However, neither study used depression scores that were normally

344 distributed; and although both screening methods have been validated and applied to 345 detect depressive symptoms, different questions may cover and emphasise different features of the illness and may result in different results<sup>54,60</sup>. The same applies to the 346 347 measurement of environmental stressors. We used a measure of 12-items adapted 348 from the List of Threating Experiences (LTE) to assess SLE over the past 6 months. 349 However, Colodro-Conde et al. used a self-reported PSLE measure over the last 12 350 months, thus covering a longer time-period of exposure to SLE prior to the assessment 351 of current depressive symptoms than in our study. PSLE included the same 12-item LTE 352 we assessed. However, they added seven items concerning serious problems getting 353 along with others (i.e. spouse, someone living with you, a close friend, etc), a range of 354 incidents highly subject to an individual's own behaviour. Thus, PSLE contains a greater 355 number of DSLE/PSLE-active-like items than ISLE/PSLE-passive-like items compared to 356 the original 12-item LTE (and TSLE). Although the overall number of DSLE-items 357 reported in our study represented less than 30% of the overall number of TSLE-items 358 reported, DSLE explained over 90% of the equivalent variance explained by TSLE (see 359 figure 2b). However, ISLE explained (~57%) far more than the remaining variance. Both 360 DSLE and ISLE effects jointly assessed in a single model showed that DSLE explained 361 70% of the variance of depressive score explained by TSLE, compared to only 30% of 362 such variance being explained by ISLE. Therefore, both covering longer time-period 363 and upweighting by DSLE/PSLE-active-like items may explain the increased explanatory power of PSLE (12.9%) used by Colodro-Conde et al. to predict depression score 364 365 compared to our TSLE measure (4.91%).

Finally, the LTE used to construct stress measures in both studies, despite being avalidated and robust screening tool, does not cover a wide range of life events with

368 marked long-term, or mild to moderate contextual stress, that could impact on the 369 final adversity faced by an individual. These unmeasured aspects of the environmental 370 exposure or its impact may also contribute to lack of stronger replication and positive 371 findings. However, these sources of bias (e.g. self-reporting bias) may be solved soon 372 by extracting data directly from national population-based electronic records (i.e. 373 medical, criminal, financial, historical, etc). Not only time-point measures but also 374 longitudinal scores may be constructed over lifespan to study GxE under a life-course 375 approach. Some studies suggests that the presence of environment-by-environment 376 (ExE) and gene-by-environment-by-environment (GxExE) interactions over lifespan are responsible of the development of mental disorders<sup>61,62</sup>. Therefore, a life-course 377 378 approach using linkable datasets is likely to help to better detect GxE effects, while 379 reducing bias, and thus provide a better understanding of both genetics and 380 environmental factors involved into the genetic-environment interplay underlying 381 depression, and mental illness in general.

In women (*p*-threshold = 1 x  $10^{-5}$ , *p* = 1.66 x  $10^{-2}$ ), the GxE effect estimates were 382 383 significantly higher (p = 0.017) than that in men, suggesting possible differences in the 384 aetiology of depression between sexes. However, such difference was not significant 385 (p = 0.32) at p-threshold (0.01) where significant GxE effect was detected in the full 386 cohort. This may explain previous differences seen between sexes such as in: 387 associated risk (i.e. approximately 1.5 - 2-fold higher in women), symptoms reported 388 and/or coping strategies (e.g. whereas women tend to cope through verbal and emotional strategies, men tend to cope by doing sport and consuming alcohol)<sup>63-67</sup>. 389 390 This also aligns with an increased risk associated with a lack of social support seen in women compared to men<sup>27</sup>. Furthermore, although we do not know whether 391

392 participants experienced recent events with positive effects, some genetic variants 393 associated with MDD may operate as 'plasticity alleles' and not just as 'risk alleles', 394 which could provide a protective effect in those women who did not experienced 395 recent SLE (and who may or may not have experienced positive environments; p = $(0.037)^{51,52}$ . This would be consistent with a differential-susceptibility model<sup>68,69</sup> of 396 397 depression. This has been suggested for the interaction effects seen between the 398 serotonin transporter linked promoter region gene (5-HTTLPR) locus and family support and liability to adolescent depression in boys<sup>70</sup>. However, our result is only 399 400 nominally significant and will require replication in larger samples. Conversely, in the 401 full cohort, our findings were consistent with a latent *diathesis* being activated by the presence of stress to manifest symptoms of depression as proposed by the *diathesis*-402 403 stress model. Future GxE studies of depression should assess the full range of life 404 events (i.e. positive and negative).

405 Empirically demonstrating the *diathesis-stress* theory for depression would validate recent<sup>24-26,28</sup> and future studies using a genome-wide approach to identify genetic 406 407 mechanisms and interactive pathways involved in GxE underpinning the causative 408 effect of "stress" in the development of depressive symptoms and mental illness in 409 general. Furthermore, sex-specific differences may improve personalized treatments 410 and therapies such as better coping strategies. This study adds to our understanding of 411 gene-by-environment interactions, although larger samples will be required to confirm 412 the suggested differences in *diathesis-stress* effects between men and women.

# 413 **ACKNOWLEDGMENTS**

414 The 1<sup>st</sup> author AAS is funded by University of Edinburgh (www.ed.ac.uk) and MRC for 415 his PhD study at the University of Edinburgh Institute of Genetics and Molecular 416 Medicine (www.ed.ac.uk/igmm). DJM acknowledges the financial support of NHS Research Scotland (NRS), through NHS Lothian. MA is supported by STRADL through a 417 418 Wellcome Trust Strategic Award (reference 104036/Z/14/Z). Generation Scotland 419 received core support from the Chief Scientist Office of the Scottish Government 420 Health Directorates [CZD/16/6] and the Scottish Funding Council [HR03006]. 421 Genotyping of the GS:SFHS samples was carried out by the Genetics Core Laboratory at 422 the Wellcome Trust Clinical Research Facility, Edinburgh, Scotland and was funded by 423 the Medical Research Council UK and the Wellcome Trust (Wellcome Trust Strategic 424 Award "STratifying Resilience and Depression Longitudinally" (STRADL) Reference 425 104036/Z/14/Z). The Psychiatric Genomics Consortium MDD Working Group depends 426 on the contributions of many parties.

427

## 428 **FINANCIAL DISCLOSURE**

429 The authors declare no conflict of interest.

# 430 **REFERENCES**

431 1. Hammen, C. Stress and depression. <i>Annu Rev Clin Psychol</i> 2	L, 293-319
432 (2003). 433 2 Kossler P.C. The effects of stressful life events on depressi	on Annu Pau
$A_{34}$ Psychol <b>A8</b> 101-214 (1007)	
435 3 Kendler KS Karkowski I M & Prescott C A Causal relati	onshin hetween
436 stressful life events and the onset of major depression Am	I Psychiatry <b>156</b>
437 837.41 (1999)	<i>j</i> 1 <i>sycillati y</i> <b>130</b> ,
438 4 Paykel ES Life events and affective disorders <i>Acta Psych</i>	iatr Scand Sunnl
439 61-6 (2003)	ad beana bappi,
440 5 Stroud CB Davila I & Mover A The relationship betwee	n stress and
441 depression in first onsets versus recurrences; a meta-anal	vtic review I
442 Abnorm Psychol <b>117</b> . 206-13 (2008).	, ele re re re re r
443 6. Ensel. W.M., Peek, M.K., Lin, N. & Lai, G. Stress in the life co	urse: a life historv
444 approach. <i>J Aging Health</i> <b>8</b> , 389-416 (1996).	y
445 7. Kendler, K.S., Karkowski, L.M. & Prescott, C.A. Stressful life	events and major
446 depression: risk period, long-term contextual threat, and c	liagnostic
447 specificity. J Nerv Ment Dis <b>186</b> , 661-9 (1998).	0
448 8. Mazure, C.M. Life Stressors as Risk Factors in Depression.	Clinical
449 <i>Psychology: Science and Practice</i> <b>5</b> , 291-313 (1998).	
450 9. Lichtenberg, P. & Belmaker, R.H. Subtyping major depress	ive disorder.
451 <i>Psychother Psychosom</i> <b>79</b> , 131-5 (2010).	
452 10. Elisei, S., Sciarma, T., Verdolini, N. & Anastasi, S. Resilience	and depressive
disorders. <i>Psychiatr Danub</i> <b>25 Suppl 2</b> , S263-7 (2013).	
454 11. Bleuler, M. Conception of Schizophrenia Within the Last Fi	fty Years and
455 Today [Abridged]. <i>Proc R Soc Med</i> <b>56</b> , 945-52 (1963).	
456 12. Bebbington, P. Misery and beyond: the pursuit of disease t	heories of
457 depression. <i>Int J Soc Psychiatry</i> <b>33</b> , 13-20 (1987).	
458 13. McGuffin, P., Katz, R. & Bebbington, P. The Camberwell Col	laborative
459 Depression Study. III. Depression and adversity in the rela	tives of
460 depressed probands. <i>Br J Psychiatry</i> <b>152</b> , 775-82 (1988).	
461 14. Robins, C.J. & Block, P. Cognitive theories of depression vie	ewed from a
462 diathesis-stress perspective: Evaluations of the models of	Beck and of
463 Abramson, Seligman, and Teasdale. <i>Cognitive Therapy and</i>	Research <b>13</b> ,
464 297-313 (1989).	
465 15. Monroe, S.M. & Simons, A.D. Diathesis-stress theories in th	e context of life
466 stress research: implications for the depressive disorders.	Psychol Bull <b>110</b> ,
467 406-25 (1991).	
468 16. Vogel, F. Schizophrenia genesis: The origins of madness. A	merican Journal of
469 Human Genetics <b>48</b> , 1218-1218 (1991).	
470 17. Mann, J.J., Waternaux, C., Haas, G.L. & Malone, K.M. Toward	
471 Of Suicidal Denavior in psychiatric patients. Am J Psychiatry	<b>/ 150</b> , 181-9
4/2 (1999). 4/2 19 Diamann D <i>et al</i> The hyperproved model of incomplete a	ovious of the
47.5 10. Riemann, <i>D. et ul.</i> The hyperatousal model of insomnia, a f 47.4 concept and its evidence. Sheen Med Pay <b>14</b> , 10, 21 (2010)	
475 19 Rolt M & Helming I M & Tintle NI The Accorditions has	
T/5 I/. Doit, M.A., Heiming, L.M. & Hitue, N.L. The Associations De	twaan Salf-
476 Reported Exposure to the Chernohyl Nuclear Disaster 7 on	tween Self- e and Mental

478	20.	Wray, N.R. & Sullivan, P.F. Genome-wide association analyses identify 44
479		risk variants and refine the genetic architecture of major depression.
480		<i>bioRxiv</i> (2017).
481	21.	Peyrot, W.J. <i>et al.</i> Effect of polygenic risk scores on depression in childhood
482	22	trauma. Br J Psychiatry <b>205</b> , 113-9 (2014).
483	22.	Musliner, K.L. <i>et al.</i> Polygenic risk, stressful life events and depressive
484		symptoms in older adults: a polygenic score analysis. <i>Psychol Med</i> <b>45</b> , 1709-
485	22	20 (2015). Descrit MU stal Desc Childhead Turun Madamta Daharania Dialafar
486	23.	Peyrot, W.J. et al. Does Childhood Trauma Moderate Polygenic Risk for
487		Conomica Concertium, Riel Reachigtry (2017)
400	24	Dunn E C at al Conomo Wide Acceptation Study (Curee) and Conomo Wide
409	24.	by Environment Interaction Study (Curvic) of Depressive Symptome in
490		African American and Hignanic /Latina Woman Depressive Symptoms in
491		(2016)
492	25	(2010). Otowa, T. <i>et al.</i> The First Pilot Genome-Wide Gene-Environment Study of
493 101	23.	Depression in the Japanese Population <i>PLoS One</i> <b>11</b> e0160823 (2016)
495	26	Ikeda M <i>et al</i> Conome-wide environment interaction between depressive
496	20.	state and stressful life events <i>I Clin Psychiatry</i> <b>77</b> e29-30 (2016)
497	27	Colodro-Conde L <i>et al</i> $A$ direct test of the disthesis-stress model for
498	27.	depression Mol Psychiatry (2017)
499	28	Coleman IRI Fley TC & Breen G Genome-wide gene-environment
500	20.	analyses of depression and reported lifetime traumatic experiences in IIK
501		Biohank <i>bioRxiv</i> (2018)
502	29.	Mullins, N. <i>et al.</i> Polygenic interactions with environmental adversity in the
503		aetiology of major depressive disorder. <i>Psychol Med</i> <b>46</b> , 759-70 (2016).
504	30.	Ivegbe, C., Campbell, D., Butler, A., Ainakina, O. & Sham, P. The emerging
505		molecular architecture of schizophrenia, polygenic risk scores and the
506		clinical implications for GxE research. Soc Psychiatry Psychiatr Epidemiol <b>49</b> ,
507		169-82 (2014).
508	31.	McGrath, J.J., Mortensen, P.B., Visscher, P.M. & Wray, N.R. Where GWAS and
509		epidemiology meet: opportunities for the simultaneous study of genetic and
510		environmental risk factors in schizophrenia. <i>Schizophr Bull</i> <b>39</b> , 955-9
511		(2013).
512	32.	Plomin, R. Commentary: missing heritability, polygenic scores, and gene-
513		environment correlation. J Child Psychol Psychiatry 54, 1147-9 (2013).
514	33.	Wray, N.R. et al. Research review: Polygenic methods and their application
515		to psychiatric traits. <i>J Child Psychol Psychiatry</i> <b>55</b> , 1068-87 (2014).
516	34.	Navrady, L.B. et al. Cohort Profile: Stratifying Resilience and Depression
517		Longitudinally (STRADL): a questionnaire follow-up of Generation Scotland:
518		Scottish Family Health Study (GS:SFHS). Int J Epidemiol (2017).
519	35.	Smith, B.H. <i>et al.</i> Generation Scotland: the Scottish Family Health Study; a
520		new resource for researching genes and heritability. <i>BMC Med Genet</i> <b>7</b> , 74
521		(2006).
522	36.	Fernandez-Pujals, A.M. <i>et al.</i> Epidemiology and Heritability of Major
523		Depressive Disorder, Stratified by Age of Onset, Sex, and Illness Course in
524		Generation Scotland: Scottish Family Health Study (GS:SFHS). <i>PLoS One</i> <b>10</b> ,
525		e0142197 (2015).

526 527	37.	Smith, B.H. <i>et al.</i> Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic
528	20	research on health and illness. Int J Epidemiol 42, 689-700 (2013).
529 530	38.	Amador, C. <i>et al.</i> Recent genomic heritage in Scotland. <i>BMC Genomics</i> <b>16</b> , 437 (2015)
530	39	Kerr S M <i>et al</i> Pedigree and genotyning quality analyses of over 10 000
532	57.	DNA samples from the Generation Scotland, Scottish Family Health Study
533		BMC Med Genet 14, 38 (2013).
534	40.	Purcell, S. <i>et al.</i> PLINK: a tool set for whole-genome association and
535		population-based linkage analyses. Am J Hum Genet 81, 559-75 (2007).
536	41.	Goldberg, D.P. & Hillier, V.F. A scaled version of the General Health
537		Ouestionnaire. <i>Psychol Med</i> 9, 139-45 (1979).
538	42.	Sterling, M. General Health Questionnaire - 28 (GHO-28). J Physiother 57,
539		259 (2011).
540	43.	Goldberg, D.P. <i>et al.</i> The validity of two versions of the GHQ in the WHO
541		study of mental illness in general health care. <i>Psychol Med</i> 27, 191-7 (1997).
542	44.	Brugha, T., Bebbington, P., Tennant, C. & Hurry, J. The List of Threatening
543		Experiences: a subset of 12 life event categories with considerable long-
544		term contextual threat. <i>Psychol Med</i> <b>15</b> , 189-94 (1985).
545	45.	Brugha, T.S. & Cragg, D. The List of Threatening Experiences: the reliability
546		and validity of a brief life events questionnaire. Acta Psychiatr Scand 82, 77-
547		81 (1990).
548	46.	Motrico, E. <i>et al.</i> Psychometric properties of the List of Threatening
549		ExperiencesLTE and its association with psychosocial factors and mental
550		disorders according to different scoring methods. J Affect Disord <b>150</b> , 931-
551		40 (2013).
552	47.	Kendler, K.S., Karkowski, L.M. & Prescott, C.A. The assessment of
553		dependence in the study of stressful life events: validation using a twin
554		design. <i>Psychol Med</i> <b>29</b> , 1455-60 (1999).
555	48.	Euesden, J., Lewis, C.M. & O'Reilly, P.F. PRSice: Polygenic Risk Score
556		software. Bioinformatics 31, 1466-8 (2015).
557	49.	Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-
558		wide complex trait analysis. Am J Hum Genet 88, 76-82 (2011).
559	50.	Keller, M.C. Gene x environment interaction studies have not properly
560		controlled for potential confounders: the problem and the (simple) solution.
561		Biol Psychiatry <b>75</b> , 18-24 (2014).
562	51.	Belsky, J. & Beaver, K.M. Cumulative-genetic plasticity, parenting and
563		adolescent self-regulation. J Child Psychol Psychiatry 52, 619-26 (2011).
564	52.	Belsky, J. <i>et al.</i> Vulnerability genes or plasticity genes? <i>Mol Psychiatry</i> <b>14</b> ,
565		746-54 (2009).
566	53.	Dudbridge, F. Power and predictive accuracy of polygenic risk scores. <i>PLoS</i>
567		Genet <b>9</b> , e1003348 (2013).
568	54.	Eaves, L.J., Last, K., Martin, N.G. & Jinks, J.L. A progressive approach to non-
569		additivity and genotype-environmental covariance in the analysis of human
570		differences. British Journal of Mathematical and Statistical Psychology 30, 1-
571		42 (1977).
572	55.	Polderman, T.J. <i>et al.</i> Meta-analysis of the heritability of human traits based
573		on fifty years of twin studies. <i>Nat Genet</i> <b>47</b> , 702-9 (2015).

574 575 576	56.	Kendler, K.S., Kuhn, J. & Prescott, C.A. The interrelationship of neuroticism, sex, and stressful life events in the prediction of episodes of major
570 577 578	57.	Magnus, K., Diener, E., Fujita, F. & Pavot, W. Extraversion and neuroticism as predictors of objective life events: a longitudinal analysis. <i>J Pers Soc Psychol</i>
579	58	05, 1040-55 (1995). Plomin P. Lichtenstein P. Pedersen N.L. McClearn C.F. & Nesselroade I.P.
581 582	50.	Genetic influence on life events during the last half of the life span. <i>Psychol</i>
583	59	Aging 5, 23-30 (1990). Clarke TK <i>et al</i> Genetic and environmental determinants of stressful life
584	57.	events and their overlap with depression and neuroticism [version 1:
585		referees: awaiting peer review], in <i>Wellcome Open Research</i> Vol. 3 %M
586		10.12688/wellcomeopenres.13893.1 (2018).
587	60.	Kang, SM. & G. Waller, N. Moderated Multiple Regression, Spurious
588		Interaction Effects, and IRT, 87-105 (2005).
589	61.	Keers, R. & Pluess, M. Childhood quality influences genetic sensitivity to
590		environmental influences across adulthood: A life-course Gene x
591		Environment interaction study. <i>Dev Psychopathol</i> <b>29</b> , 1921-1933 (2017).
592	62.	Assary, E., Vincent, J.P., Keers, R. & Pluess, M. Gene-environment interaction
593		and psychiatric disorders: Review and future directions. <i>Semin Cell Dev Biol</i>
594		(2017).
595	63.	Weissman, M.M. <i>et al.</i> Sex differences in rates of depression: cross-national
596		perspectives. J Affect Disord <b>29</b> , 77-84 (1993).
597	64.	Van de Velde, S., Bracke, P. & Levecque, K. Gender differences in depression
598		in 23 European countries. Cross-national variation in the gender gap in
599		depression. <i>Soc Sci Med</i> <b>71</b> , 305-13 (2010).
600	65.	Labonte, B. <i>et al</i> . Sex-specific transcriptional signatures in human
601		depression. <i>Nat Med</i> (2017).
602	66.	Angst, J. et al. Gender differences in depression. Epidemiological findings
603		from the European DEPRES I and II studies. Eur Arch Psychiatry Clin
604		Neurosci <b>252</b> , 201-9 (2002).
605	67.	Piccinelli, M. & Wilkinson, G. Gender differences in depression. Critical
606		review. Br J Psychiatry <b>177</b> , 486-92 (2000).
607	68.	Belsky, J. & Pluess, M. Beyond diathesis stress: differential susceptibility to
608		environmental influences. <i>Psychol Bull</i> <b>135</b> , 885-908 (2009).
609	69.	Belsky, J., Bakermans-Kranenburg, M.J. & van Ijzendoorn, M.H. For Better
610		and for Worse: Differential Susceptibility to Environmental Influences.
611		Current Directions in Psychological Science <b>16</b> , 300-304 (2007).
612	70.	Li, J.J., Berk, M.S. & Lee, S.S. Differential susceptibility in longitudinal models
613		of gene-environment interaction for adolescent depression. <i>Dev</i>
614		Psychopathol <b>25</b> , 991-1003 (2013).
615		
616		

# 617 **FIGURE LEGENDS**

618 Figure 1. a) Association between polygenic risk scores (PRS) and depressive score 619 (main effects, one-sided tests). PRS were generated at 8 p-threshold levels using 620 summary statistics from the Psychiatric Genetic Consortium MDD GWAS (released July 621 2016) with the exclusion of Generation Scotland participants. Depression score was 622 derived from The General Health Questionnaire. Y-axis represents the % of variance of 623 depression score explained by PRS main effects. Full cohort (yellow) was split into men 624 (blue) and women (red). b) Association between number of SLE reported and 625 depression score (main effect, one-sided tests, results expressed in % of depression 626 score explained). SLE were self-reported through a brief life events questionnaire 627 based on the List of Threating Experiences and categorized into: total number of SLE 628 reported (TSLE), "dependent" SLE (DSLE) or "independent" SLE (ISLE). Full cohort 629 (yellow) was split into men (blue) and women (red).

630 Figure 2. Association between GxE effect and depression score. Results represent 631 percentage of depression score explained by the interaction term (two-sided tests) 632 fitted in linear mixed models to empirically test the *diathesis-stress* model. Red 633 numbers show significant interactions p-values. Full cohort (yellow) was split into men 634 (blue) and women (red). PRS were generated at 8 p-threshold levels using summary 635 statistics from the Psychiatric Genetic Consortium MDD GWAS (released July 2016) 636 with the exclusion of Generation Scotland participants. The interaction effect was 637 tested with a) the number of SLE reported (TSLE), b) "dependent" SLE (DSLE) and c) 638 "independent" SLE (ISLE).

Figure 3. Scatterplot of significant *diathesis-stress* interactions on the risk of depressive
symptoms a) in full cohort and b) in women. X-axis represents the direct effect of PRS

641	(standard deviation from the mean) based on <b>a</b> ) $p$ -threshold = 0.01 and <b>b</b> ) $p$ -threshold
642	= 1 x $10^{-5}$ , using the total number of SLE reported (TSLE) by each participant (dot) as
643	environmental exposures at three SLE levels represented by colours. Blue: O SLE,
644	"missing stress", n = 1 833/1 041; green: 1 or 2 SLE, "low stress", n = 2 311/1 459; red:
645	3 or more SLE, "high stress", n = 775/490 in the full cohort and women, respectively. Y-
646	axis reflects log transformed depression score standardized to mean of 0 and standard
647	deviation of 1. Lines represent the increment of risk of depression under a certain
648	degree of "stress" dependent on a genetic predisposition (= <i>diathesis</i> ).
649	

Sample	FULL COHORT			WOMEN								
diathesis	PRS at <i>p</i> v	hold = 0.01	PRS at <i>p</i> value threshold = 10-5									
SLE	E TSLE			TSLE			DSLE			ISLE		
SLE category	missing	low	high	missing	low	high	missing	low	high	missing	low	high
N	1833	2311	755	1041	1459	490	2111	816	63	1393	1338	259
Effect (β)	0.021	0.043	0.142	-0.061	0.014	0.078	-0.022	0.061	-0.05	-0.044	0.027	0.079
s.e.	0.022	0.021	0.039	0.029	0.027	0.049	0.021	0.038	0.147	0.026	0.027	0.068
t	0.957	2.07	3.644	-2.086	0.541	1.609	-1.055	1.605	-0.337	-1.654	0.978	1.16
<i>p</i> value	0.339	0.039	2.86 x 10 <sup>-4</sup>	0.037	0.589	0.108	0.292	0.109	0.738	0.098	0.328	0.247
CI (95%)	-0.022 0.065	0.002 0.084	0.065 0.218	-0.119 -0.004	-0.038 0.066	-0.017 0.174	-0.063 0.019	-0.014 0.136	-0.344 0.245	-0.095 0.008	-0.027 0.080	-0.055 0.214

Table 1. *Diathesis* effect under 3 SLE categories at those significant GxE effect detected in the full cohort and in women.

SLE categories (amount of SLE experienced): 0 SLE = "missing", 1 or 2 SLE = "low", and 3 or more SLE = "high". TSLE, DSLE and ISLE stand for "total", "dependent" and "independent" SLE reported.





