

OPEN

# The genetic relationship between female reproductive traits and six psychiatric disorders

Guiyan Ni<sup>1,2,3</sup>, Azmeraw T. Amare<sup>4,5</sup>, Xuan Zhou<sup>2</sup>, Natalie Mills<sup>5</sup>, Jacob Gratten<sup>1,6</sup> & S. Hong Lee<sup>2</sup>

Female reproductive behaviours have important implications for evolutionary fitness and health of offspring. Here we used the second release of UK Biobank data (N = 220,685) to evaluate the association between five female reproductive traits and polygenic risk scores (PRS) projected from genome-wide association study summary statistics of six psychiatric disorders (N = 429,178). We found that the PRS of attention-deficit/hyperactivity disorder (ADHD) were strongly associated with age at first birth (AFB) (genetic correlation of  $-0.68 \pm 0.03$ ), age at first sexual intercourse (AFS) ( $-0.56 \pm 0.03$ ), number of live births (NLB) ( $0.36 \pm 0.04$ ) and age at menopause ( $-0.27 \pm 0.04$ ). There were also robustly significant associations between the PRS of eating disorder (ED) and AFB ( $0.35 \pm 0.06$ ), ED and AFS ( $0.19 \pm 0.06$ ), major depressive disorder (MDD) and AFB ( $-0.27 \pm 0.07$ ), MDD and AFS ( $-0.27 \pm 0.03$ ) and schizophrenia and AFS ( $-0.10 \pm 0.03$ ). These associations were mostly explained by pleiotropic effects and there was little evidence of causal relationships. Our findings can potentially help improve reproductive health in women, hence better child outcomes. Our findings also lend partial support to the evolutionary hypothesis that causal mutations underlying psychiatric disorders have positive effects on reproductive success.

Female reproductive behaviours, including age at first birth (AFB), age at first sexual intercourse (AFS), age at menarche (AMC), age at menopause (AMP) and number of live births (NLB) have important implications in reproductive health and evolutionary fitness<sup>1,2</sup>. Some of these traits have been shown to associate with the physical and mental health of offspring<sup>3</sup>, and there has been growing evidence that maternal AFB is associated with increased risk of psychiatric disorder<sup>4,5</sup> and behavioural problems in their children<sup>6,7</sup>. For instance, both early and late maternal AFB are associated with an increased risk to schizophrenia (SCZ)<sup>4</sup>, bipolar disorder (BIP)<sup>8</sup>, attention-deficit/hyperactivity disorder (ADHD)<sup>7</sup>, autism spectrum disorder (ASD)<sup>9</sup> and depression<sup>10</sup> in offspring. Age at menopause and menarche also tend to correlate with the risk of adverse mental health outcomes in offspring<sup>11–14</sup>.

The relationship between reproductive behavior and susceptibility to psychiatric disorders is likely complex and bidirectional. Individuals with psychiatric illnesses and their relatives may be more prone to risk taking and impulsive behaviors, which cause early pregnancy and childbirth in women, or they may exhibit poor social interactions that result in delays in key reproductive transitions such as marriage, pregnancy and childbirth<sup>15–20</sup>.

In addition to the epidemiological evidence, the phenotypical relationship between female reproductive traits and psychiatric disorder risk in offspring may have a genetic basis<sup>21–26</sup>. Genome wide association studies (GWAS) that examined reproductive behaviors and psychiatry disorders separately have identified several genome-wide significant genetic loci influencing AFB<sup>21</sup>, NLB<sup>21</sup> and AFS<sup>3</sup>, and ones affecting susceptibility to ADHD<sup>27</sup>, ASD<sup>23</sup>, eating disorders (ED)<sup>28</sup>, BIP<sup>29</sup>, major depressive disorder (MDD)<sup>30</sup>, and SCZ<sup>31</sup>. When reproductive traits and psychiatric disorders are examined together, a complex genetic relationship between the two clusters of traits begins

<sup>1</sup>Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland, 4072, Australia. <sup>2</sup>Australian Centre for Precision Health, University of South Australia Cancer Research Institute, University of South Australia, Adelaide, SA, 5000, Australia. <sup>3</sup>School of Environmental and Rural Science, University of New England, Armidale, NSW, 2351, Australia. <sup>4</sup>South Australian Academic Health Science and Translation centre, South Australian Health and Medical Research Institute (SAHMRI), Adelaide, South Australia, Australia. <sup>5</sup>Discipline of Psychiatry, School of Medicine, University of Adelaide, Adelaide, SA, Australia. <sup>6</sup>Mater Research Institute, University of Queensland, Brisbane, Queensland, 4072, Australia. Correspondence and requests for materials should be addressed to S.H.L. (email: [hong.lee@unisa.edu.au](mailto:hong.lee@unisa.edu.au))

AFB	N	AFS	N	AMC	N	AMP	N	NLB	N
<20	11,414	<16	13,915	<11	8,058	<40	4,055	0	32,805
20 to <25	42,048	16 to <20	86,038	11 to <13	60,029	40 to <46	14,468	1	23,284
25 to <30	46,966	20 to <24	43,054	13 to <15	76,625	46 to <52	42,704	2	80,594
30 to <35	16,880	24 to <28	9,690	15 to <17	25,859	52 to <58	38,011	3	31,182
≥35	4,236	≥28	3,446	≥17	2,285	≥58	3,148	>3	9,879
Total	121,544		156,143		172,856		102,386		177,744

**Table 1.** Sample breakdown by age at first birth, age first had sexual intercourse, age at menarche, age at menopause and number of live births. AFB: Age at first birth. AFS: Age at first sexual intercourse. AMC: Age at menarche. AMP: age at menopause. NLB: Number of live births.

to emerge<sup>32</sup>. For instance, recent studies<sup>24,26</sup> showed a U-shaped relationship between AFB and the polygenic risk score of SCZ<sup>24</sup>, which echoes with the phenotypic association between maternal age and SCZ risk in offspring found in epidemiological studies (e.g. McGrath *et al.*<sup>4</sup>, El-Saadi *et al.*<sup>33</sup>, Byrne *et al.*<sup>34</sup>). However, for most other psychiatric disorders, the latent genetic architecture shared with reproductive traits remains unknown.

In this study, we investigate the genetic association of five specific reproductive traits in women (AFB, AFS, AMC, AMP and NLB) with six common psychiatric disorders (ADHD, ASD, ED, BIP, MDD, and SCZ) using polygenic score and LDSC approach. Causal relationship between the risk of psychiatric disorders and female reproductive traits is further tested using Mendelian randomization. This will shed light on the complex bio-psychosocial risk factors associated with shared genetic effects between female reproductive and psychiatric traits. Furthermore, this study provides insight into whether complex phenotypes of reproductive behaviour can be caused by mental health or rather by pleiotropic genes associated with both traits.

## Materials and Methods

UK Biobank's scientific protocol and operational procedures were reviewed and approved by the North West Multi-centre Research Ethics Committee (MREC), National Information Governance Board for Health & Social Care (NIGB), and Community Health Index Advisory Group (CHIAG). UK Biobank also obtained informed consent from the study participants. Research Ethics approval was obtained from University of South Australia Human Research Ethics Committee (HREC). All methods were performed in accordance with the relevant guidelines and regulations.

**Data.** *UK Biobank sample and quality control.* We used the second release of the UK Biobank dataset that initially included 264,859 women with imputed genotypes for 35.6 million SNPs. The imputation was based on the Haplotype Reference Consortium reference panel<sup>35</sup>. The quality control (QC) criteria for the imputed genotype data included an imputation reliability (INFO score) >0.6<sup>36–38</sup>, minor allele frequency >0.01, p-value for Hardy-Weinberg equilibrium test >10E-7, and SNP missingness <0.05. Women with non-white British ancestry, individual missingness >0.05, individuals with putative sex chromosome aneuploidy or one individual per pair of relatives was excluded. After the QC above, we compared discordance between the first and second release of QCed UK Biobank data for each SNP and individual that were excluded if the discordance rate was above 0.05. In addition, we excluded genome-wide significant SNPs in a GWAS analysis where individuals in the first and second release UK Biobank data were treated as cases and controls. Subsequently, 220,685 individuals and 7,253,311 SNPs remained, from which 1,133,064 Hapmap3 SNPs were selected for the analyses. Traits of interest were AFB, AFS, AMC, AMP, and NLB. Number of observations for each trait are provided in Table 1.

*Psychiatric genomics consortium (PGC) GWAS summary statistics results.* The GWAS summary results of six psychiatric disorders were obtained from the PGC (<http://www.med.unc.edu/pgc/>), including ADHD<sup>27</sup>, ASD<sup>23</sup>, ED<sup>28</sup>, BIP<sup>29</sup>, MDD<sup>30</sup>, and SCZ<sup>31</sup>. The number of cases and controls in these studies are provided in Table S1.

**Statistical analyses.** *Estimation of polygenic risk scores (PRS).* For each psychiatric disorder, PRS were computed for the UK Biobank sample as the sum of the risk alleles weighted by the estimated SNP effects from the PGC GWAS that are in public domain (<http://www.med.unc.edu/pgc/>) and not likely to include UK Biobank sample. To compute PRS for the UK Biobank sample we used approximately 1,133,064 HapMap 3 SNPs from each study. The projection of the SNP effects onto the UK Biobank data was conducted using MTG2<sup>39</sup>.

*Mean difference of PRS across the five categories.* For each reproductive trait, we first divided UK Biobank samples into five groups according to their values on the trait (see Table 1 for sample breakdowns). We then did pairwise comparisons between group means for each trait using two-tailed t-tests.

*Linear prediction.* Using a linear or polynomial regression model, we assessed, for each disorder, if PRS significantly predicted each of the phenotypes of the reproductive traits. In the prediction model, each of the reproductive traits was adjusted for age at interview, year of birth, study center, genotype batch, and the first 15 principal components (PCs) provided by the UK Biobank. For sensitivity analyses, we repeated the prediction using additional variables including educational attainment, income level, and smoking and alcohol consumption status.

**Genetic correlation.** Genetic correlation is a classic population parameter used to infer the geometric mean of trait variance (i.e. the additive genetic covariance between two traits scaled by the square root of the product of the genetic variance for each trait). Pairwise genetic correlations were estimated between the reproductive traits and psychiatric disorders using linkage disequilibrium score regression (LDSC) based on GWAS summary statistics<sup>40,41</sup>. GWAS was carried out and SNP effects were estimated for each of the reproductive traits using PLINK 1.9<sup>42</sup>. In the GWAS, phenotypes were adjusted for age at interview, year of birth, assessment center at which participant consented, genotype batch and the first 15 PCs. We also used additional covariates such as educational attainment, income level, smoking and alcohol consumption status to further adjust the phenotypes in sensitivity analyses.

**Analysis of overlapping samples between PGC and UK Biobank data.** The intercept of LDSC is estimated as  $\rho N_s / \sqrt{N_1 N_2}$ , where  $\rho$  is the phenotypic correlation among  $N_s$  overlapping individuals in the two studies with sample sizes of  $N_1$  and  $N_2$ , respectively<sup>40</sup>. To check if there were any overlapping samples between UK Biobank and PGC, we explicitly estimated the intercept when using LDSC to estimate genetic correlations between the reproductive traits and psychiatric disorders

**Causal effects of psychiatric disorder PRS on female reproduction traits.** We performed two-sample Mendelian randomization (MR) analyses to test whether there were causal relationships between the risk of psychiatric disorders and female reproductive traits. We used psychiatric disorder PRS as exposure data and female reproductive traits as outcomes. Genome-wide significant SNPs ( $p < 5E-08$ ) from the PGC GWAS summary statistic results were used as instruments in the MR analyses. Associations between the instruments and the outcomes were tested<sup>42</sup>. Prior to the association test, we adjusted the outcomes, the phenotypes of the female reproductive traits, for age at interview, year of birth, assessment centre at which participant consented, genotype batch and the first 15 PCs. We used an inverse-variance weighted (IVW) regression<sup>43</sup> as the primary analysis to determine a causal relationship and used MR-Egger<sup>44</sup>, weighted median<sup>45</sup> and weighted mode<sup>46</sup>, which rely on different assumptions regarding the validity of instruments, for sensitivity analyses. Details of those assumptions are described elsewhere<sup>43–46</sup>. MR analyses were performed using a R package, TwoSampleMR<sup>47</sup>, and P-values were obtained using a two-tailed t-test with the respective degree of freedom. The instruments were available for ADHD, BIP and SCZ only, therefore ASD, ED, MDD were not tested for the MR analyses.

## Results

**Basic characteristics.** In this study, we analyzed genetic data from 220,685 women who were part of the UK Biobank. The average age of the study participants at recruitment was 57, with a range from 40 to 71. The number of observations for each reproductive trait is provided in Table 1. The total sample of women was divided into five groups according to their AFB, AFS, AMC, AMP or NLB status (Table 1) to detect the mean difference of PRS across the five categories.

**Mean difference of PRS across the five categories.** First, we computed PRS of the six psychiatric disorders for the UK Biobank sample and estimated the mean of the PRS for each age category of each reproductive trait (Figs 1–5). Then, the mean difference of the PRS was tested across the five categories for each of the reproductive traits (Tables S2–S6). A statistically significant difference in the PRS was determined after Bonferroni correction for multiple testing (i.e. significance level divided by the number of tests,  $0.05/300 = 1.7E-4$ ) and the results that passed the significance threshold were highlighted in bold (see Tables S2–S6).

In Fig. 1, the direction of association between the PRS and AFB was positive (the older the AFB, the higher the risk) for ED, ASD and BIP, but negative (the younger the AFB, the higher the risk) for ADHD, MDD and SCZ although some traits (MDD and SCZ) showed non-linear (U-shaped) associations. A U-shaped relationship was previously shown between AFB and SCZ in an independent study<sup>24</sup>. The mean difference in PRS was statistically significant for most pairwise comparisons of age classes for ADHD, ED and MDD (Table S2). The significance was particularly pronounced for ADHD, with a P-value of  $2.0E-184$ , for example, for the difference between women with  $AFB < 20$  and women with  $30 \leq AFB < 35$  (Table S2). For SCZ, most significant signals came from younger AFB groups (Fig. 1 and Table S2).

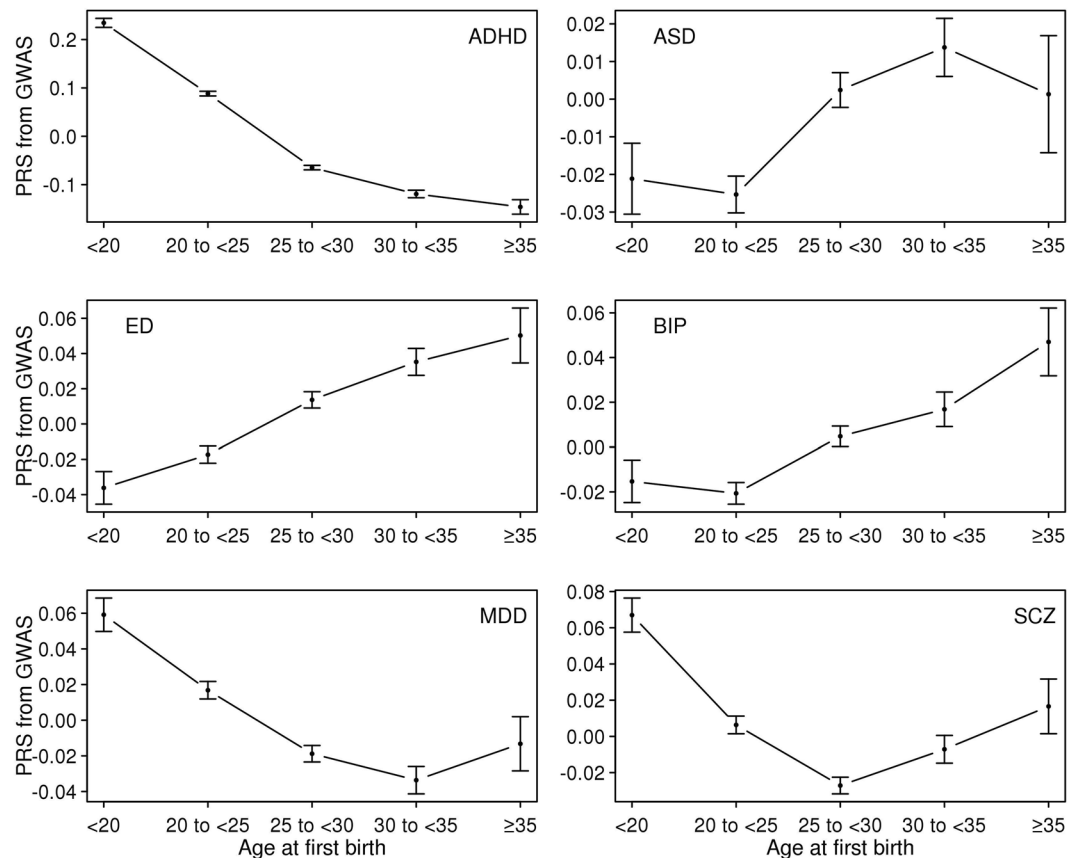
The results for AFS were similar to those for AFB, consistent with the strong correlation between these traits (Fig. 6), although some signals were reduced and not significant after the Bonferroni correction (ASD, BIP and MDD) (Fig. 2 and Table S3).

For AMC and AMP, there were relatively few statistically significant differences between age groups for the six psychiatric disorders, with the exception of ADHD (Tables S4 and S5). The relationship between the PRS of ADHD and AMC was non-linear (Fig. 3), with earlier or later AMC tending to have significantly higher ADHD risk than intermediate AMC (Table S4). A similar relationship was observed between the ADHD PRS and AMP in that earlier or later AMP was associated with higher ADHD risk than moderate AMP (Fig. 4 and Table S5).

Since NLB was expected to be negatively correlated with AFB (see Fig. 6), the pattern of the mean PRS of ADHD was inversely correlated between the groups classified according to NLB and AFB (Figs 1 and 5). Table S6 shows that the ADHD PRS was strongly associated with NLB. In addition, there were a number of significant association signals of NLB for MDD and SCZ (Table S6).

**Linear prediction.** We used a regression method to assess the significance of genetic association between the reproductive behavior traits and six psychiatric disorders (see Methods).

We regressed each of pre-adjusted reproductive traits (AFB, AFS, AMC, AMP and NLB) on the PRS of each of the psychiatric disorders. Using the linear regression prediction modeling, we showed that the PRS of each six psychiatric disorder were significantly associated with at least one of the five reproductive traits, confirming



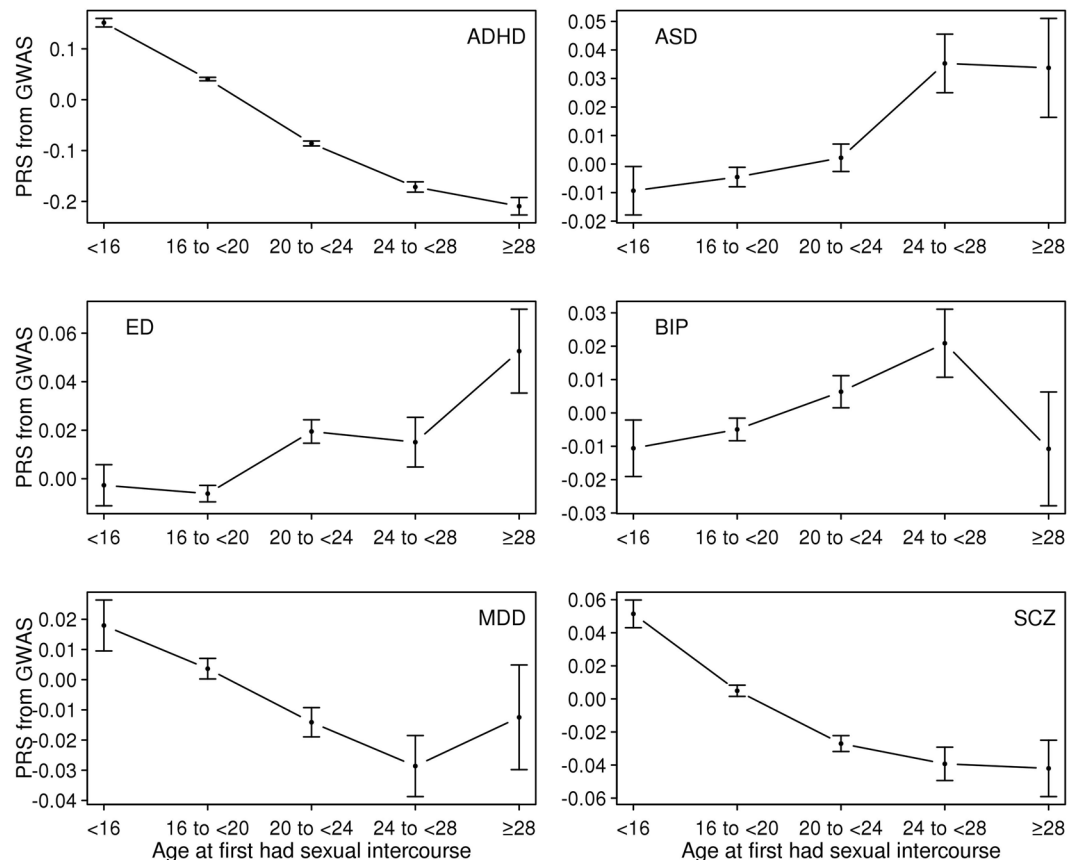
**Figure 1.** Means and standard errors of PRSs for the six psychiatric disorders, by age at first birth, in the UK Biobank sample.

some of the robust associations from the earlier analyses of mean difference of PRS across the five age categories (Tables S2–S6). Of all the five reproductive traits, AFB was the trait best predicted by the PRS of the six psychiatric disorders (Fig. 7, the corresponding R-squared and P-values are in Table S7). Because AFB and AFS are highly correlated traits, the results for AFS were similar to those for AFB except that there was no significant association between PRS of AFS and BIP, which is consistent with the analyses of mean difference of PRS above (Tables S2 and S3). NLB was significantly predicted by the PRS of ADHD, ASD, MDD, and SCZ. In addition, AMC and AMP were only associated with the PRS of ASD and ADHD, respectively. We noted that the majority of significance was explained by linear predictions, but not by quadratic polynomial predictions for all of the associations (Table S8). There were marginal significances for quadratic polynomial associations only for a few pairwise comparisons (AFB vs SCZ and AFS vs ED).

We conducted a sensitivity analysis in which dependent variables were further adjusted for educational attainment, income level, smoking and alcohol consumption status (Fig. S1 and Table S9). Most of the association signals remained significant with some exceptions. For example, the association of AFB, AFS, and NLB with ASD disappeared, as did the association between AFB and BIP. Conversely, ADHD became significantly associated with AMC after correcting for the additional covariates.

When using GWAS p-value thresholds to filter SNPs, the significance of linear prediction decreased for most of the association analyses between the reproductive traits and six psychiatric disorders (Table S10). This indicates that the associations between the reproductive traits and six psychiatric disorders were probably due to many genes, not due to a few major genes.

**Genetic correlations.** We used LDSC to estimate genetic correlations between the reproductive behavior traits and six psychiatric disorders (see Methods). We estimated genetic correlations between the five reproductive traits to reveal the shared genetic architecture of the traits (Fig. 6). As expected, the genetic correlation between AFB and AFS was very high ( $0.821 \pm 0.018$ ) and that between AFB and NLB was high ( $-0.594 \pm 0.032$ ). However, the genetic correlation between AMC and AMP was relatively low and they also have moderate or low genetic correlations with other reproductive traits, e.g.  $0.346 \pm 0.036$  between AFB and AMP,  $0.303 \pm 0.039$  between AFS and AMP and  $\sim 0.1$  between AMC and other traits (Fig. 6). These results explain the observations in the analyses of mean difference in PRS (Figs 1–5) and linear predictors (Fig. 7), where the results between AFB and AFS were similar, and those between AFB (AFS) and NLB were reciprocally similar for the association with ADHD PRS. The estimated genetic correlations among those five reproductive traits remained similar after



**Figure 2.** Means and standard errors of PRS for the six psychiatric disorders, by age at first sexual intercourse, in the UK Biobank sample.

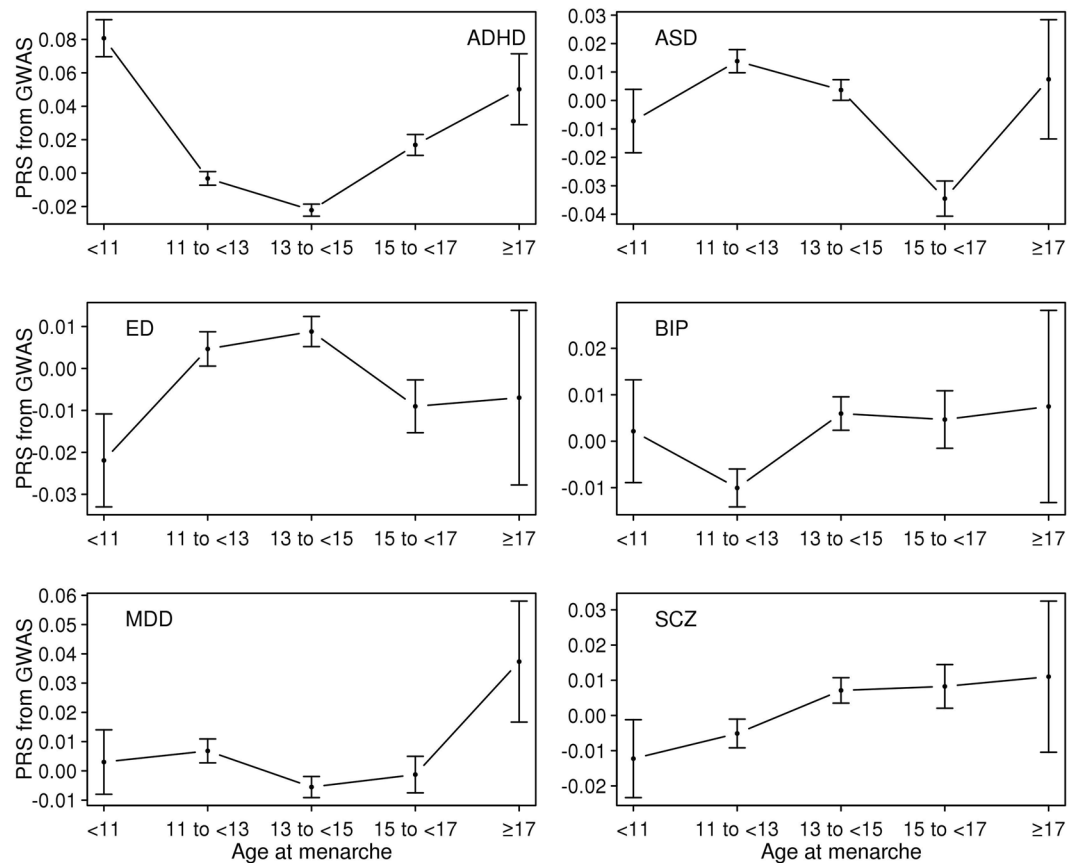
additional adjustment of the dependent variables for educational attainment, income levels, and smoking and alcohol consumption status (Fig. 6 vs. Fig. S2).

Figure 8 shows the estimated genetic correlation from LDSC for each pair of five reproductive traits and six psychiatric disorders. The detailed genetic correlations and P-value are in Table S11. Nine pairs of genetic correlations out of 30 were significantly different from zero after Bonferroni correction. For AFB analyses, the estimated genetic correlations were greater than zero (positive association) between AFB and ED ( $0.349 \pm 0.061$ ), and lower than zero (negative association) between AFB and ADHD ( $-0.677 \pm 0.034$ ) and AFB and MDD ( $-0.273 \pm 0.069$ ). Similarly, AFS was inversely correlated with ADHD ( $-0.563 \pm 0.034$ ), MDD ( $-0.265 \pm 0.066$ ) and SCZ ( $-0.100 \pm 0.030$ ) and positively correlated with ED ( $0.189 \pm 0.055$ ). For AMC and AMP, there was no significant genetic correlation except that between AMP and ADHD ( $-0.272 \pm 0.038$ ). NLB showed positive genetic correlation with ADHD ( $0.356 \pm 0.042$ ) and was non-significant for other pairs of traits. These results agreed with the analyses of mean difference of PRS (Figs 1–5) and linear prediction above (Fig. 7).

In the analyses where dependent variables were further adjusted for education, income levels, smoking and alcohol consumption status, the estimated genetic correlations between reproductive traits and psychiatric disorders were not substantially changed (Fig. S3 and Table S11), compared to those depicted in Fig. 8 and Table S11.

**Analysis of overlapping samples between PGC and UK biobank data.** The intercepts between the reproductive traits and six psychiatric disorders were not significantly different from zero, indicating little overlaps between UK biobank and PGC samples (Figs S4 and S5). The sole exception was that the intercept between AFB and ED was significantly lower than zero ( $-0.017 \pm \text{SE } 0.006$ ), which may be due to sampling errors or excessive heterogeneity<sup>40,41</sup>. Most of the intercepts from the cross-trait LDSC analyses of the five reproductive traits were significantly different from zero, as expected (Figs S6 and S7). We note that the intercepts of AMP-AMC and AMC-NLB were not different from zero, which was because of the small phenotypic correlation between these pairs of traits.

**Causal effects of psychiatric disorder PRS on female reproduction traits.** In the analysis using IVW regression, we found weak evidence (not significant after multiple testing correction) for a causal relationship between ADHD and AFB ( $\beta = -0.521$ , 95% confidence interval:  $-0.969$  to  $-0.072$ , p-value = 0.031), and between ADHD and AFS ( $\beta = -0.407$ , 95% CI:  $-0.667$  to  $-0.147$ , p-value = 0.010). None of the sensitivity analyses confirmed the significance except the weighted median approach, i.e. p-value = 0.042 for the causal



**Figure 3.** Means and standard errors of PRS for the six psychiatric disorders by age at menarche in the UK Biobank sample.

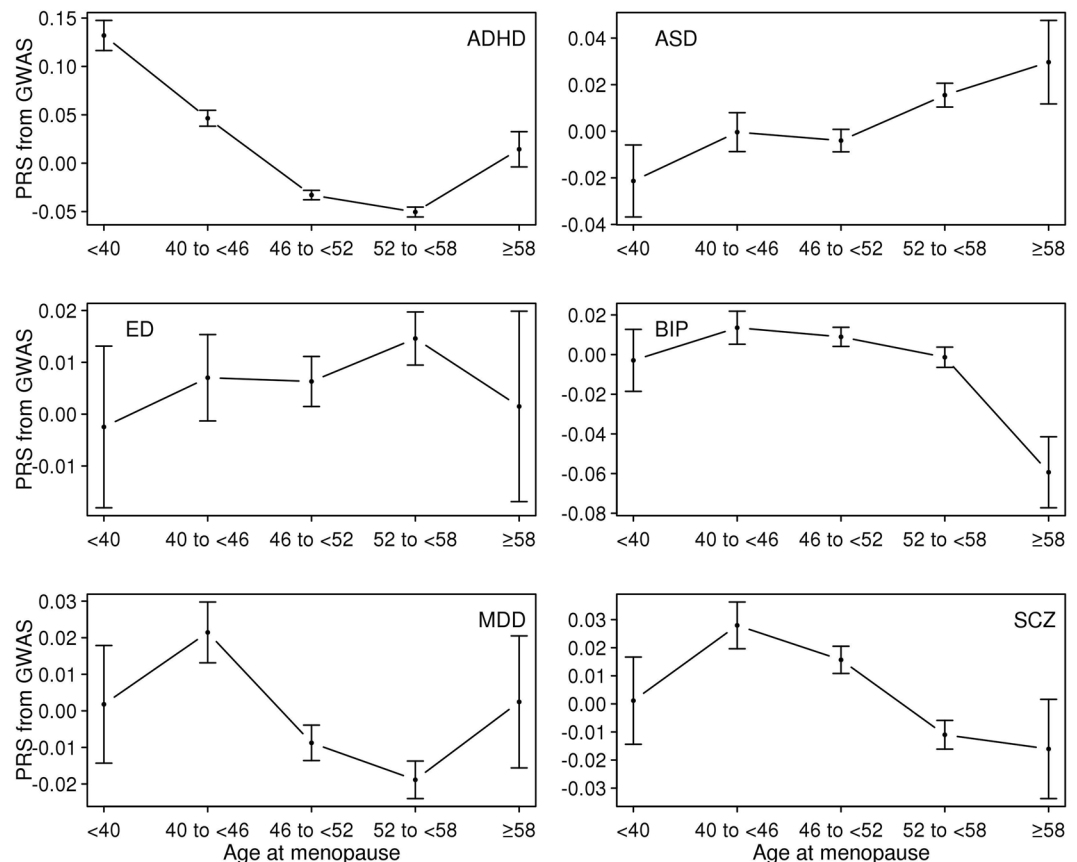
relationship between ADHD and AFB and  $p$ -value = 0.016 for that between ADHD and AFS (which was also not significant after multiple testing correction) (Table 2). It was noted that the MR-Egger intercept estimate for the causal relationship between ADHD and AFS was significantly different from zero ( $p$ -value = 0.041), indicating that the signal of causal relationship might be due to pleiotropic effects. There was no notable evidence for causal relationship between any pair of psychiatric disorder PRS and reproduction traits in the analyses. The  $I^2$  statistics<sup>48</sup> of MR-Egger were 0.96, 0.69 and 0.97 for ADHD, BIP and SCZ, mostly satisfying the ‘no measurement error’ assumption<sup>48</sup> except BIP that had a weaker association signal than ADHD or SCZ (Figs 7 and 8).

## Discussion

We revealed the complex psychosocial genetic risk architecture underpinning the association between reproductive traits and six psychiatric disorders using the HapMap3 SNPs that are known to be reliable<sup>1,37,38,40</sup>. The strong genetic associations between ADHD and most reproductive traits (except AMC) were supported by three different approaches, the analyses of mean difference of PRS across five different (age) categories, linear prediction and genetic correlation using LDSC. Genetic associations between two of the reproductive traits (AFB and AFS) and both MDD and ED were consistently significant across the various analyses. The association between SCZ and AFB (and NLB) was significant in the analyses of mean difference of PRS and linear prediction although the estimated genetic correlation from LDSC analysis was not significant.

It is known that modelling genetic correlation between traits can increase accuracy significantly in predicting individual genetic risk for the traits<sup>49,50</sup>. Our findings of significant associations between female reproductive traits and between these traits and psychiatric disorders may be useful in improving female reproductive health, hence better child outcomes. For example, the high-risk group for ADHD could be informed about features of ADHD, such as impulsive behaviour and possible consequences of impulsivity. This intervention may lead to prevent them giving birth at an immature age, which can improve their reproductive health and the maternal environment for their baby. Furthermore, this information of genetic predisposition, combined with information on psychosocial factors, can be recorded as a form of family medical history, and used to monitor health of offspring.

Psychiatric disorders are highly heritable traits (e.g. ADHD or SCZ) while they persist in the population in a stable prevalence rate. It is questioned why natural selection has not excluded the causal mutations underlying psychiatric disorders. Three plausible hypotheses have been proposed. First, the impact of natural selection on the removal of existing causal mutations may be slower than the addition of new genetic mutations causing, for



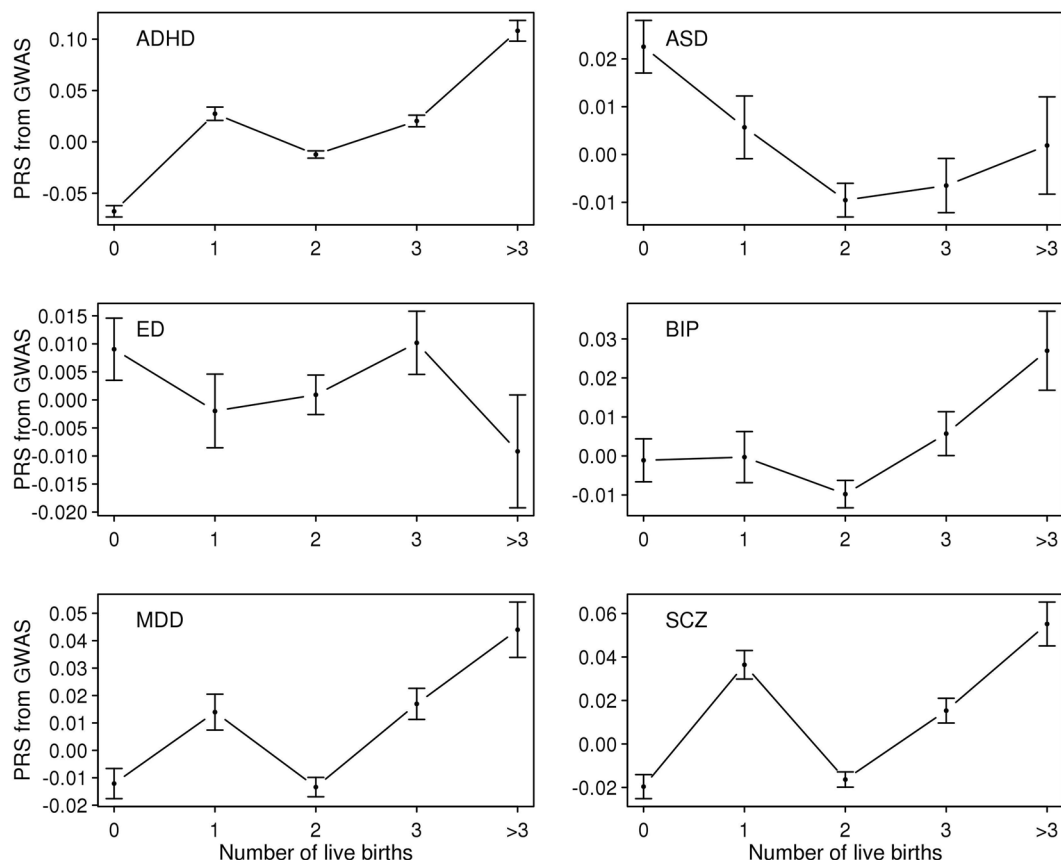
**Figure 4.** Means and standard errors of PRS for the six psychiatric disorders by age at menopause in the UK Biobank sample.

example, ADHD or SCZ. Second, pre-existing neutral mutations may interact with environments that have been new exposures for the population, and cause the disorders<sup>51</sup>. Third, causal mutations underlying psychiatric disorders have positive effects on reproductive success. The last hypothesis can be supported by our observation that NLB and ADHD have a strong positive genetic correlation (Figs 7 and 8). For SCZ and MDD, there is a suggestive signal for a positive genetic association with NLB (Figs 7 and 8), which agrees with a recent comprehensive study for SCZ<sup>52</sup>. However, this hypothesis is not supported by the results for ASD, BIP and ED.

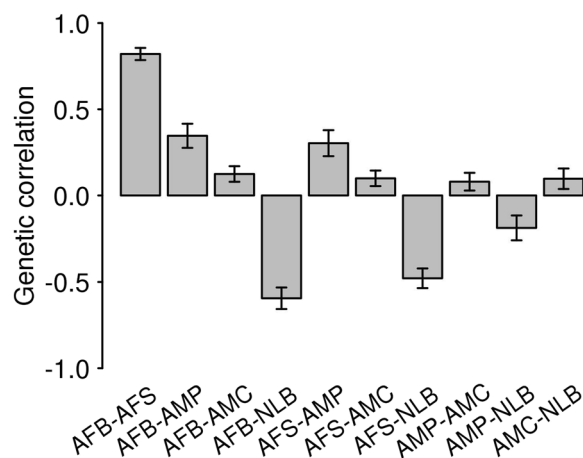
ADHD is shown to be genetically associated with a set of reproductive traits such as AFB or AFS. However, we found no clear evidence that higher genetic risk of ADHD decreased AFB or AFS, from the MR analyses. The negative association between ADHD PRS and AFB or between ADHD PRS and AFS found in linear prediction and genetic correlation might be mostly because of pleiotropic effects. MR-Egger result also showed that pleiotropic effects explained the relationship between ADHD PRS and AFS. We did not find any evidence to support a causal relationship for any pair of psychiatric disorder PRS and reproduction traits, which agree with a recent study<sup>53</sup> reporting that there is no evidence of a causal effect of genetic liability for schizophrenia on NLB or AFB. It was noted that we used a small number of SNPs as instruments in the MR analyses especially for ADHD and BIP. Therefore, a further study with a large number of instruments should be warranted when more genome-wide SNPs become available.

We report a novel shared genetic architecture between female reproductive traits, i.e. including strong positive genetic correlation between AFB and AFS, high negative genetic correlation between AFB and NLB, and AFS and NLB, moderate positive genetic correlation between AMP and AFB (AFS), moderate negative genetic correlation between AMP and NLB, and negligible genetic correlation between AMC and other reproductive traits. These reported relationships among the reproductive traits may help shed light on how causal genetic variants are involved and shared in the regulation of reproductive mechanisms at the genome-wide level. The strong positive and negative associations between AFB and AFS, and AFB and NLB are intuitive and well expected. However, other relationships (e.g. moderate positive association between AFB and AMP and negligible association between AMC and AMP) should be further investigated in a future study to identify the exact mechanism of menarche and menopause in relation to reproductive success<sup>54</sup>.

There are a number of limitations in this study. Firstly, we used the PRS or LDSC approach using GWAS summary statistics instead of individual-level genotype data for psychiatric disorders, for which we do not have access permissions. The power and accuracy of detection and estimation in the analyses could be increased when using individual-level genotype data<sup>55</sup>. Secondly, we tested causal relationship between the PRS of ADHD, BIP and SCZ and five reproductive traits in women and found no evidence of causality. A relatively small number



**Figure 5.** Means and standard errors of PRS for the six psychiatric disorders by number of live births in the UK Biobank sample.



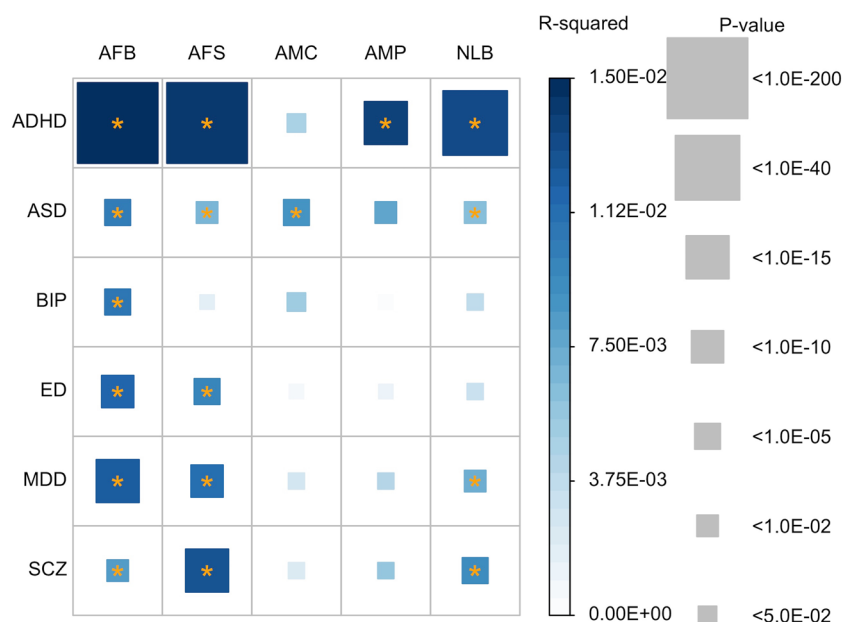
**Figure 6.** Genetic correlations among the five reproductive traits estimated using the base model. In the base model, the reproductive traits were adjusted for age at interview, year of birth, study center, genotype batch, and the first 15 principal components. Error Bars are 95% confidence intervals. AFB: Age at first birth. AFS: Age at first sexual intercourse. AMC: Age at menarche. AMP: age at menopause. NLB: Number of live births.

of instruments might result in lack of power to detect such causal relationship if there is any. Thirdly, we only focused on reproductive traits in women because data on related male traits (e.g. AFB, AFS) are not fully available in the UKB. The collection of data on reproductive traits such as AFB in men, that would enable the study of genetic associations between psychiatric disorders and these reproductive traits should be a priority for the field. Fourthly, for the female reproductive traits, we used self-reported data and ascertained sample that is less representative of the general population. A fifth limitation is the possible confounding from factors such as a history of early childhood adversity and socioeconomic status (SES). A history of childhood adversity has been associated



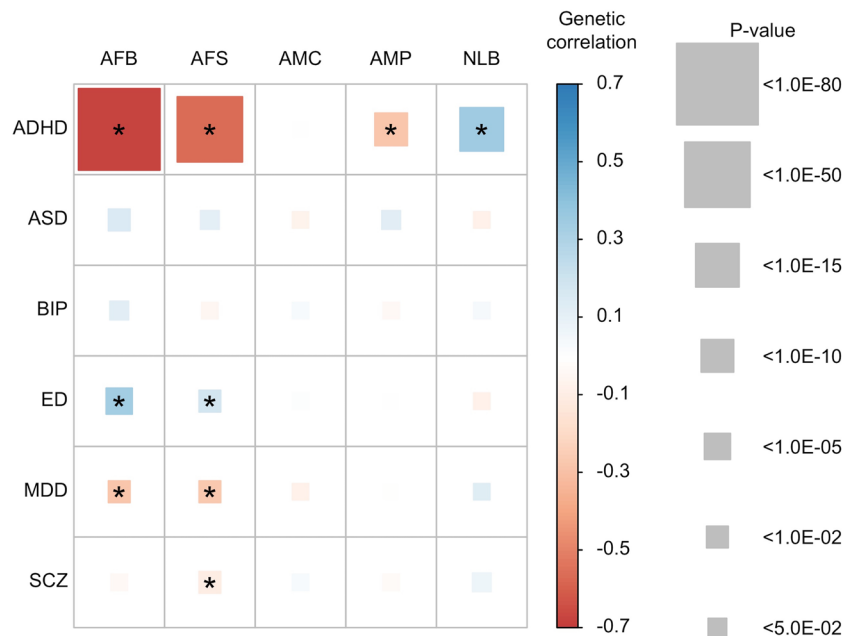
	Method	#Instruments	DF	AFB		AFS		AMC		AMP		NLB	
				Beta	p-value	Beta	p-value	Beta	p-value	Beta	p-value	Beta	p-value
ADHD	Inverse variance weighted	6	5	-0.521	0.031	-0.407	0.010	-0.029	0.553	-0.209	0.144	0.028	0.292
ADHD	MR Egger Intercept	6	4	-0.080	0.267	-0.073	0.041	-0.025	0.161	-0.044	0.393	-0.004	0.679
ADHD	MR Egger	6	4	0.290	0.679	0.336	0.264	0.223	0.215	0.236	0.649	0.071	0.521
ADHD	Weighted median	6	5	-0.414	0.042	-0.352	0.016	-0.038	0.371	-0.244	0.175	0.038	0.220
ADHD	Weighted mode	6	5	-0.429	0.143	-0.309	0.114	-0.040	0.461	-0.250	0.305	0.067	0.194
BIP	Inverse variance weighted	4	3	0.083	0.297	0.087	0.246	-0.029	0.493	-0.137	0.197	-0.001	0.933
BIP	MR Egger Intercept	4	2	0.008	0.881	0.048	0.272	-0.026	0.359	-0.069	0.318	0.002	0.846
BIP	MR Egger	4	2	0.035	0.918	-0.198	0.420	0.125	0.452	0.272	0.486	-0.013	0.836
BIP	Weighted median	4	3	0.090	0.324	0.052	0.408	0.003	0.916	-0.177	0.176	0.003	0.862
BIP	Weighted mode	4	3	0.147	0.305	0.044	0.615	0.015	0.649	-0.210	0.242	0.005	0.839
SCZ	Inverse variance weighted	76	75	-0.056	0.268	-0.062	0.061	0.024	0.208	0.001	0.974	0.006	0.424
SCZ	MR Egger Intercept	76	74	-0.003	0.842	-0.006	0.584	0.002	0.811	0.008	0.569	0.003	0.275
SCZ	MR Egger	76	74	-0.011	0.960	0.018	0.907	0.004	0.965	-0.107	0.583	-0.032	0.374
SCZ	Weighted median	76	75	0.011	0.830	-0.066	0.055	0.010	0.516	0.012	0.828	0.014	0.166
SCZ	Weighted mode	76	75	0.091	0.449	-0.112	0.295	-0.015	0.713	-0.008	0.953	0.026	0.333

**Table 2.** Two sample MR of causal association of psychiatric disorders on female reproduction traits of interest. AFB: Age at first birth. AFS: Age at first sexual intercourse. AMC: Age at menarche. AMP: age at menopause. NLB: Number of live births. ADHD: Attention-Deficit/Hyperactivity Disorder. ASD: Autism spectrum disorder. ED: Eating disorder. BIP: Bipolar disorder. MDD: Major depressive disorder. SCZ: Schizophrenia.



**Figure 7.** Coefficient of determination (R<sup>2</sup>) and p-values for its significance based on a linear prediction model. Color of each box represents the level of R-squared, and the size of squares represents its significance (p-value). R-squared that are significantly different from zero after Bonferroni correction (0.05/30) are marked with an asterisk. Dependent variables were adjusted for age at interview, year of birth, assessment centre at which the participant consented, genotype batch, and the first 15 principal components. The number of records used for the analyses was 121,544 for AFB, 156,143 for AFS, 102,386 for AMP, and 172,856 for AMC and 177,744 for NLB. AFB: age at first birth. AFS: age at first sexual intercourse. AMP: age at menopause. AMC: age at menarche. NLB: number of live births. ADHD: Attention-Deficit/Hyperactivity Disorder. ASD: Autism spectrum disorder. ED: Eating disorder. BIP: Bipolar disorder. MDD: Major depressive disorder. SCZ: Schizophrenia.

with a range of adverse mental health outcomes, including psychosis<sup>56</sup>, BIP<sup>57</sup>, and MDD<sup>58</sup>. With regard to SES, children living in low-income households have been found to have an increased risk for ADHD<sup>59</sup>. Low access to economic resources have been associated both with an earlier age of menarche, and in both females and males,



**Figure 8.** Genetic correlations between the five reproductive traits and the six psychiatric disorders estimated using the base model. Color of each box represents the level of estimated genetic correlation (blue for positive and red for negative correlation), and the size of squares represents its significance (p-value). Estimated genetic correlations that are significantly different from zero after Bonferroni correction (0.05/30) are marked with an asterisk. AFB: Age at first birth. AFS: Age at first sexual intercourse. AMC: Age at menarche. AMP: age at menopause. NLB: Number of live births. ADHD: Attention-Deficit/Hyperactivity Disorder. ASD: Autism spectrum disorder. ED: Eating disorder. BIP: Bipolar disorder. MDD: Major depressive disorder. SCZ: Schizophrenia. In the base model, the reproductive traits were adjusted for age at interview, year of birth, study center, genotype batch, and the first 15 principal components.

a larger number of offspring at an earlier age<sup>60</sup>. In particular, a strong link has been observed between lower SES during childhood and parenthood during adolescence<sup>61</sup>. Therefore, in sensitivity analyses, we repeated the prediction using income level (as a measure of SES) as an additional variable.

In conclusion, we revealed the latent genetic architecture between the five reproductive traits in women and six psychiatric disorders. In particular, ADHD PRS and a set of reproductive traits were substantially associated through common genetic factors. There were also a number of robust associations between reproductive traits and ED, SCZ and MDD PRS. There was no evidence that the risk of any psychiatric disorder modulated the phenotypes of the female reproductive traits. Therefore, the associations found in this study were mostly due to genetic pleiotropic effects. Our findings can have potential to help improve reproductive health in women and their child outcomes. These findings also can help address an evolutionary hypothesis that causal mutations underlying ADHD, MDD and SCZ have positive effects on reproductive success.

**URLS.** PGC: <http://www.med.unc.edu/pgc/>.

MTG2: <https://sites.google.com/site/honglee0707/mtg2>.

LDSC: <https://github.com/bulik/ldsc>.

### Data Availability

The genome-wide analysis summary statistics of Psychiatric genomics consortium results can be downloaded from the public domain (<http://www.med.unc.edu/pgc/>). The individual genotype data of UK Biobank sample can be available via application to UK Biobank (<http://www.ukbiobank.ac.uk/register-apply/>).

### References

1. Tropf, F. C. *et al.* Hidden heritability due to heterogeneity across seven populations. *Nature Human Behaviour* **1**, 757–765 (2017).
2. Tropf, F. C. *et al.* Human fertility, molecular genetics, and natural selection in modern societies. *PLoS ONE* **10**, 1–14 (2015).
3. Day, F. R. *et al.* Physical and neurobehavioral determinants of reproductive onset and success. *Nature genetics* **48**, 617 (2016).
4. McGrath, J. J. *et al.* A comprehensive assessment of parental age and psychiatric disorders. *JAMA psychiatry* **71**, 301–9 (2014).
5. Croen, L. A., Najjar, D. V., Fireman, B. & Grether, J. K. Maternal and paternal age and risk of autism spectrum disorders. *Archives of pediatrics & adolescent medicine* **161**, 334–340 (2007).
6. Barclay, K. & Myrskylä, M. Maternal age and offspring health and health behaviours in late adolescence in Sweden. *SSM-population health* **2**, 68–76 (2016).
7. Chang, Z. *et al.* Maternal age at childbirth and risk for ADHD in offspring: A population-based cohort study. *International Journal of Epidemiology* **43**, 1815–1824 (2014).
8. Menezes, P. R. *et al.* Paternal and maternal ages at conception and risk of bipolar affective disorder in their offspring. *Psychological medicine* **40**, 477–485 (2010).

9. Idring, S. *et al.* Parental age and the risk of autism spectrum disorders: findings from a Swedish population-based cohort. *International journal of epidemiology* **43**, 107–115 (2014).
10. Tearne, J. E. *et al.* Older maternal age is associated with depression, anxiety, and stress symptoms in young adult female offspring. *Journal of abnormal psychology* **125**, 1 (2016).
11. Leibenluft, E. Women with bipolar illness: clinical and research issues. *American Journal of Psychiatry* **153**, 163 (1996).
12. Ballinger, C. B. Psychiatric aspects of the menopause. *The British Journal of Psychiatry* **156**, 773–787 (1990).
13. Tondo, L., Pinna, M., Serra, G., De Chiara, L. & Baldessarini, R. J. Age at menarche predicts age at onset of major affective and anxiety disorders. *European Psychiatry* **39**, 80–85 (2017).
14. Platt, J. M., Colich, N. L., McLaughlin, K. A., Gary, D. & Keyes, K. M. Transdiagnostic psychiatric disorder risk associated with early age of menarche: A latent modeling approach. *Comprehensive Psychiatry* (2017).
15. Özcan, N. K., Boyacıoğlu, N. E., Enginkaya, S., Dinç, H. & Bilgin, H. Reproductive health in women with serious mental illnesses. *Journal of clinical nursing* **23**, 1283–1291 (2014).
16. Matevosyan, N. R. Reproductive health in women with serious mental illnesses: a review. *Sexuality and Disability* **27**, 109–118 (2009).
17. Bellack, A. S., Morrison, R. L., Wixted, J. T. & Mueser, K. T. An analysis of social competence in schizophrenia. *The British Journal of Psychiatry* **156**, 809–818 (1990).
18. Giordano, G. N., Ohlsson, H., Sundquist, K., Sundquist, J. & Kendler, K. S. The association between cannabis abuse and subsequent schizophrenia: a Swedish national co-relative control study. *Psychological medicine* **45**, 407–414 (2015).
19. Salas-Wright, C. P., Vaughn, M. G., Ugalde, J. & Todic, J. Substance use and teen pregnancy in the United States: evidence from the NSDUH 2002–2012. *Addictive behaviors* **45**, 218–225 (2015).
20. Reddy, L. F. *et al.* Impulsivity and risk taking in bipolar disorder and schizophrenia. *Neuropsychopharmacology* **39**, 456 (2014).
21. Barban, N. *et al.* Genome-wide analysis identifies 12 loci influencing human reproductive behavior. *Nature genetics* **48**, 1462 (2016).
22. Perry, J. R. B. *et al.* Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature* **514**, 92 (2014).
23. The Autism Spectrum Disorders Working Group of The Psychiatric Genomics, C. Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia. *Molecular Autism* **8**, 21 (2017).
24. Mehta, D. *et al.* Evidence for genetic overlap between schizophrenia and age at first birth in women. *JAMA Psychiatry* **73**, 497–505 (2016).
25. Gratten, J. *et al.* Risk of psychiatric illness from advanced paternal age is not predominantly from de novo mutations. *Nature Genetics* **48**, 718–724 (2016).
26. Ni, G., Gratten, J., Wray, N. R. & Lee, S. H. Age at first birth in women is genetically associated with increased risk of schizophrenia. *Scientific reports* **8**, 10168 (2018).
27. Demontis, D. *et al.* Discovery Of The First Genome-Wide Significant Risk Loci For ADHD. *bioRxiv* (2017).
28. Duncan, L. *et al.* Significant locus and metabolic genetic correlations revealed in genome-wide association study of anorexia nervosa. *American Journal of Psychiatry* **174**, 850–858 (2017).
29. Psychiatric Gwas Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nature genetics* **43**, 977–983 (2011).
30. Ripke, S. *et al.* A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular psychiatry* **18**, 497–511 (2013).
31. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
32. Mullins, N. *et al.* Reproductive fitness and genetic risk of psychiatric disorders in the general population. *Nature Communications* **8**, 15833 (2017).
33. El-Saadi, O. *et al.* Paternal and maternal age as risk factors for psychosis: findings from Denmark, Sweden and Australia. *Schizophrenia Research* **67**, 227–236 (2004).
34. Byrne, M., Agerbo, E., Ewald, H., Eaton, W. W. & Mortensen, P. B. Parental age and risk of schizophrenia: a case-control study. *Archives of general psychiatry* **60**, 673–678 (2003).
35. Loh, P.-R. *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nature genetics* **48**, 1443 (2016).
36. Border, R. *et al.* Imputation of behavioral candidate gene repeat variants in 486,551 publicly-available UK Biobank individuals. *European Journal of Human Genetics*, **1** (2019).
37. Lee, S. H., Weerasinghe, W. M. S. P. & van der Werf, J. H. J. Genotype-environment interaction on human cognitive function conditioned on the status of breastfeeding and maternal smoking around birth. *Scientific Reports* **7**, 6087 (2017).
38. Peyrot, W. J. *et al.* Does childhood trauma moderate polygenic risk for depression? A meta-analysis of 5765 subjects from the psychiatric genomics consortium. *Biological psychiatry* **84**, 138–147 (2018).
39. Lee, S. H. & van der Werf, J. MTG2: An efficient algorithm for multivariate linear mixed model analysis based on genomic information. *Bioinformatics* **32**, 1420–1422 (2016).
40. Bulik-Sullivan, B. *et al.* An Atlas of Genetic Correlations across Human Diseases and Traits. *Nature Genetics* **47**, 1236–1241 (2015).
41. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics* **47**, 291–295 (2015).
42. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4** (2015).
43. Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genetic epidemiology* **37**, 658–665 (2013).
44. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *International journal of epidemiology* **44**, 512–525 (2015).
45. Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genetic epidemiology* **40**, 304–314 (2016).
46. Hartwig, F. P., Davey Smith, G. & Bowden, J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *International journal of epidemiology* **46**, 1985–1998 (2017).
47. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human genome. *eLife* **7**, e34408 (2018).
48. Bowden, J. *et al.* Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I<sup>2</sup> statistic. *International journal of epidemiology* **45**, 1961–1974 (2016).
49. Maier, R. *et al.* Joint Analysis of Psychiatric Disorders Increases Accuracy of Risk Prediction for Schizophrenia, Bipolar Disorder, and Major Depressive Disorder. *American Journal of Human Genetics* **96**, 283–294 (2015).
50. Maier, R. M. *et al.* Improving genetic prediction by leveraging genetic correlations among human diseases and traits. *Nature Communications* **9**, 989 (2018).
51. Durisko, Z., Mulsant, B. H., McKenzie, K. & Andrews, P. W. Using evolutionary theory to guide mental health research. *The Canadian Journal of Psychiatry* **61**, 159–165 (2016).
52. Escott-Price, V. *et al.* The Relationship Between Common Variant Schizophrenia Liability and Number of Offspring in the UK Biobank. *American Journal of Psychiatry*, appi-ajp (2019).
53. Lawn, R. B. *et al.* Schizophrenia risk and reproductive success: a Mendelian randomization study. *Royal Society Open Science* **6**, 181049 (2019).

54. Mishra, G. D. *et al.* Early menarche, nulliparity and the risk for premature and early natural menopause. *Human Reproduction* **32**, 679–686 (2017).
55. Ni, G. *et al.* Estimation of Genetic Correlation via Linkage Disequilibrium Score Regression and Genomic Restricted Maximum Likelihood. *American Journal of Human Genetics* **102** (2018).
56. Varese, F. *et al.* Childhood adversities increase the risk of psychosis: a meta-analysis of patient-control, prospective- and cross-sectional cohort studies. *Schizophr Bull* **38**, 661–71 (2012).
57. Palmier-Claus, J. E., Berry, K., Bucci, S., Mansell, W. & Varese, F. Relationship between childhood adversity and bipolar affective disorder: systematic review and meta-analysis. *Br J Psychiatry* **209**, 454–459 (2016).
58. Chapman, D. P. *et al.* Adverse childhood experiences and the risk of depressive disorders in adulthood. *J Affect Disord* **82**, 217–25 (2004).
59. Choi, Y., Shin, J., Cho, K. H. & Park, E. C. Change in household income and risk for attention deficit hyperactivity disorder during childhood: A nationwide population-based cohort study. *J Epidemiol* **27**, 56–62 (2017).
60. Amir, D., Jordan, M. R. & Bribiescas, R. G. A Longitudinal Assessment of Associations between Adolescent Environment, Adversity Perception, and Economic Status on Fertility and Age of Menarche. *PLoS One* **11**, e0155883 (2016).
61. van Roode, T., Sharples, K., Dickson, N. & Paul, C. Life-Course Relationship between Socioeconomic Circumstances and Timing of First Birth in a Birth Cohort. *PLoS One* **12**, e0170170 (2017).

## Acknowledgements

This research is supported by the Australian National Health and Medical Research Council (1080157, 1087889), and the Australian Research Council (DP160102126, DP190100766, FT160100229). This research has been conducted using the UK Biobank Resource. UK Biobank (<http://www.ukbiobank.ac.uk>) Research Ethics Committee (REC) Approval Number is 11/NW/0382. Our reference number approved by UK Biobank is 14575. This research was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian Government.

## Author Contributions

S.H.L. conceived the idea and directed the study. G.N. and S.H.L. performed the analyses and interpreted the results. J.G. and N.M. provided critical feedback and key elements in interpreting the results. S.H.L., G.N., J.G. and X.Z. drafted the manuscript. With support from S.H.L., A.A. made interpretation of the analyses output, organised the scientific theory of the work and drafted the manuscript. All authors contributed to editing and approval of the final manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-48403-x>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019