The number of words: 2,953

The number of tables: 3

The number of figures: 4

The relationship between common variant schizophrenia liability and number of offspring in the UK Biobank

Valentina Escott-Price[1*], Antonio F. Pardiñas[1], Enrique Santiago[2], James Walters[1], George Kirov[1], Michael J. Owen[1], Michael C. O'Donovan[1]

[1] MRC Centre for Neuropsychiatric Genetics & Genomics, Cardiff University

[2] Departamento de Biología Funcional, Facultad de Biología, Universidad de Oviedo, Oviedo, Spain.

[*] Corresponding author

MRC Centre for Neuropsychiatric Genetics & Genomics, Cardiff University

Hadyn Ellis Building, Maindy Road, Cardiff, CF24 4HQ, UK

E-mail: EscottPriceV@cf.ac.uk

Phone: +44 (0)29 20688429

Abstract

Objective: Schizophrenia is associated with a marked reduction in reproductive success, yet alleles that are common contribute substantially to its liability. Among several possible explanations for this, it has been postulated that those who carry risk alleles but are unaffected are at some reproductive advantage, offsetting the effects of negative selection in those who are affected. Here, we sought to test this, isolating the effects of risk alleles on fecundity from those effects that are contingent on expressing the disorder.

Method: We compared the burden of schizophrenia risk alleles, as indexed by a polygenic risk score (PRS), carried by 139,679 participants in the UK Biobank study who did not have schizophrenia, with the number of offspring of those individuals.

Results: Higher schizophrenia liability in subjects without manifest disorder was weakly but significantly associated with having more children (B=0.006, 95%CI=[0.002, 0.010], p=0.002). The relationship was dependent on gender, with a positive correlation between number of children and liability in females (B=0.011, 95%CI=[0.006, 0.016], $p=7.1x10^{-5}$) whereas in males, higher liability was associated with being childless (OR=0.96, 95%CI=[0.94, 0.98], $p=7.8x10^{-5}$). We also present data showing that the negative effect on number of children associated with schizophrenia itself is 2 to 15-fold greater than the positive effect associated with PRS in unaffected individuals.

Conclusions: We propose that this complex relationship between liability and fecundity is consistent with sexual selection. While the overall pattern of a weak positive correlation with liability may contribute to the persistence of schizophrenia risk alleles, our findings indicate that the negative selection acting on individuals affected by the disorder in the general population is larger than any advantage conferred by genetic loading in unaffected individuals.

Introduction

Due to improvements in medical care and the reduction in prenatal, infant and child mortality in developed countries, the number of offspring can be considered a measure of an individual's reproductive fitness [1]. A number of studies have reported that people with schizophrenia, particularly males, have substantially fewer children than the general population [2-4]. The relationship between an individual and the number of their offspring is, however, a complex phenotype. In addition to variance in aspects of fundamental biology related to fertility and foetal survival, number of offspring is also driven by the mating choices and opportunities of individuals, reflecting that individual's social, cultural, religious, economic and historical environment; it is unclear which of these is primarily responsible for the association between schizophrenia and low fecundity.

Schizophrenia has a substantial heritability [5, 6] to which alleles across the frequency spectrum contribute [7-9]. As expected for a disorder associated with low fecundity, alleles that contribute large effects on risk are rare, their frequencies being dictated by the balance between the rates of new (*de novo*) mutation formation and of removal of those mutations by strong negative selection, which typically occurs within a few generations [10]. However, a substantial component of liability to schizophrenia is conferred by risk alleles that are common and which individually confer small effects on risk [11, 12]. There has been a long-standing debate, from an evolutionary perspective, as to how common risk alleles persist in the population [13]: hypotheses that have been proposed include; schizophrenia alleles confer reproductive advantages in unaffected people depending on their particular genetic background or environmental context (historical or contemporary); risk alleles are maintained by linkage to

positively selected alleles; common risk alleles for schizophrenia are individually too weak to be effectively purged by negative selection; risk alleles occur in regions where the effects of more deleterious variants cause haplotypes to be removed from the gene pool, which reduces genetic diversity, impairs the efficiency of selection, and allows allowing alleles with small deleterious effects to rise in frequency by drift [13]. Of these possible explanations, perhaps the most widely known is the notion that schizophrenia may persist in modern populations because those carrying risk alleles but who do not manifest schizophrenia have psychological traits that increase their fecundity. For example Jarvik and Deckard [14] proposed that schizoid-paranoid personality traits found in non-psychotic relatives of those with schizophrenia, which they termed the Odyssean personality, may be advantageous in a world "plagued by terror, strife and war".

A recent study on reproductive fitness and genetic risk of psychiatric disorders in 93,720 genotyped Icelanders (aged at least 45 years, without a diagnosis of a psychiatric disorder)[15] found no evidence that common schizophrenia risk alleles confer reproductive advantage, although, as expected, rare alleles with large effects sizes in the form of copy number variants (CNVs) did confer reproductive disadvantage, with a greater fecundity reduction in males than females.

Here, we sought to examine the relationship between schizophrenia liability and fecundity in the UK Biobank [16], a powerful resource for testing hypotheses at population level. Specifically, we sought to test the hypothesis that the burden of common risk alleles for schizophrenia is associated with reproductive advantage/disadvantage in subjects without schizophrenia.

Given sex related effects of the disorder on fecundity, we also tested male and female risk allele carriers separately.

Methods

UK Biobank data

UK Biobank is a large prospective cohort of more than half a million residents of the UK for which genetic data and number of children data are available. The phenotype "number of children" is the self-report response to "How many children have you fathered (or mothered)?" If the answer was ">200" then it was discounted by the UK Biobank; if the answer ">15" then the participant was asked to confirm the answer.

The schizophrenia GWAS we used to define risk alleles was primarily of European Ancestry so we restricted the sample to those who self-reported as being of white UK or Irish ancestry (N=140,494). In the first wave of UKBB data, 227 people (79 males and 148 females) have self-declared schizophrenia or been admitted to hospital for an episode of schizophrenia or schizoaffective disorder. These people were removed as we were interested in people not expressing the disorder. We also excluded 558 people with missing information on the numbers of children and thus 139,679 were included in our analyses.

UK Biobank obtained informed consent from all participants and this study was conducted under generic approval from the NHS National Research Ethics Service (approval letter dated 13 May 2016, Ref 16/NW/0274) and under UK Biobank approvals for application #14421.

Measuring Burden of Risk Alleles: Polygenic Risk Scores (PRS)

To define risk alleles, we used the largest available schizophrenia GWAS comprising a meta-analysis of two large studies 1) the latest study of the Psychiatric Genetic consortium (PGC2) [8] and 2) the latest study on individuals with a clinical diagnosis of schizophrenia attending a clozapine clinic and controls (CLOZUK) [13]. The schizophrenia datasets were imputed using SHAPEIT/IMPUTE2 software [17, 18] with a combination of the 1000 Genomes phase 3 (1KGPp3) and UK10K datasets as a reference panel. The total number of subjects in the data included 40,675 schizophrenia cases and 64,643 controls [13]. For constructing PRS in the UK Biobank, we included autosomal imputed genetic data that passed stringent quality control criteria (minor allele frequencies (MAF) ≥0.01) and imputation quality score greater than or equal to 0.9. This resulted in 5,471,613 SNPs. Using the UK Biobank genotypes, we performed clumped pruning using parameters r^2=0.2, a physical distance threshold for clumping SNPs of 1Mb and an association p-value threshold in the PGC2+CLOZUK GWAS study of 0.05. The clumped pruning process retains SNPs that are the most significantly associated with schizophrenia while excluding SNPs at which the genotypes are correlated with the selected SNPs.

We selected markers, based upon the schizophrenia association significance threshold p<0.05, to construct a polygenic score in the UK Biobank data (N markers= 32,576). P<0.05 is the threshold that maximally captures polygenic risk in current studies of schizophrenia [8]. The PRS was calculated from the effect size-weighted sum of associated alleles within each subject. PRS were adjusted for array (the UK Biobank used two different arrays) and the first 8 principal components and then standardised by subtracting the population mean for PRS and dividing by the standard deviation. The first 8 principal components, out of 15 available in the Biobank, were selected after visual inspection of each pair of PCs, taking forward only those that resulted in multiple clusters of individuals (see [19] for details).

Statistical analysis

We used Poisson regression analysis to test number of children for association with schizophrenia PRS, covarying for age and sex. (Age was strongly positively associated with the number of children in all analyses.) We examined interaction between sex and schizophrenia PRS comparing two Poisson regression models 1) PRS , age and sex and 2) PRS, age, sex and interaction PRS x sex. As the distributions of numbers of children born to males and females may differ [15], we also performed this analysis for males and females separately.

As a test of non-linear effects, we tested seven groups according to the numbers of children (0, 1, 2, …, 6+), and performed logistic regression analysis (adjusted for PCs, array, age as before), testing whether childless participants had greater or lower genetic liability to schizophrenia compared with participants with a particular number of children. We used the truncated by 6 children measure for the interaction analysis as the number of UK Biobank participants reporting more than 6 children is small (N=288) and the inclusion of multiple small groups may adversely affect with small numbers the robustness of the analyses.

We report B-coefficients for Poisson regression analyses and odds ratios (OR=exponential(B)) for the logistic regression analyses.

In order to compare the effect on fecundity of the PRS in individuals unaffected by schizophrenia with the effects seen in those with schizophrenia, we used Fisher's Fundamental Theorem of Natural Selection [20, 21], which establishes a relationship between fecundity and strength of selective pressure, based on additive genetic variance. The set of calculations we performed is detailed in the Supplementary Note. Briefly, we retrieved data

from GWAS and epidemiological studies, and used them to calculate the variance in fecundity in the general population (including those with schizophrenia) explained by the PRS. Then, we compared that value with the variance in fecundity that, according to our data, could be explained by the PRS in unaffected individuals. The ratio of both variances allowed us to compare the strength of selective pressures (i.e. negative selection in affected individuals vs. positive selection in unaffected individuals), and infer their effect in the general population.

Results

PRS and number of children

For males and females (without a diagnosis of schizophrenia) combined, the number of children was weakly but significantly positively associated with schizophrenia PRS (B=0.006, 95%CI=[0.002,0.010], p=0.0017), see summary in Table 1. The effect size was very similar to the one reported previously in the Icelandic population [15] (B=0.006, p=0.160), although in that study, the association was not statistically significant. The quadratic test for non-linear relationships was not significant.

Fecundity effects in healthy and affected individuals

The observed positive association between schizophrenia PRS and fecundity applies to the subset of the general population which is unaffected by schizophrenia. When this effect is extrapolated to the general population, it can be compared with the known negative association between fecundity and onset of schizophrenia[2-4]. After both effects are considered, in the light of the schizophrenia population prevalence, the negative effect contributed by the disorder to the population is substantially (2 to 15-fold (95%CI)) greater than the positive effect associated with the PRS (Supplementary Note).

PRS and number of children in males and females separately

The percentages of males and females with each category of number of children born to males and females are shown in Figure 1. Polygenic risk of schizophrenia was not associated with the number of offspring for males (B=0.001, 95%CI=[-0.005,0.007], p=0.745) but it was positively associated with that for females, (B=0.011, 95%CI=[0.006, 0.016], $p=7.1x10^{-5}$), see also Table 1.

We further explored the patterns of association in males and females, comparing the PRS in people with specific numbers of offspring (Tables 2 and 3) with the PRS in people without children; OR is derived as the exponential of the logistic regression Beta coefficient estimate, and represents the change in the odds of membership in the target group, defined by the exact numbers of children, with an increase of one standard deviation in PRS. In Figure 2, we show the distribution for the PRS of females and males by number of children. The boundaries of the box are the first and third quartiles, the horizontal line inside the box shows the median and "whiskers" above and below the box show the locations of the minimum and maximum. Across all categories of numbers of children, PRS is positively associated with the numbers offspring of females, that is, the higher the liability to schizophrenia, the more children they have (Table 2, Figure 2.A). In males, the effect was less consistent. Males with 1 or 2 children had lower schizophrenia PRS compared to those with no children. Indeed, schizophrenia PRS was associated with a relative absence of children born to fathers OR=0.96, 95%CI=[0.94, 0.99], p=0.008, and OR=0.94, 95%CI=[0.92, 0.96], $p=2.6x10^{-9}$, compared with PRS for 1 and 2 children, respectively; see also first two lines of Table 2 and Figure 2.B). Moreover, there was a non-significant trend for schizophrenia PRS to be elevated in those with atypically large

numbers of offspring (5 or more). Comparing PRS for males with and without children, higher genetic liability to schizophrenia was associated with having no offspring in males (OR=0.96, 95%CI=[0.94, 0.98], $p=8.01x10^{-5}$).

PRS x sex interaction

We tested for an interaction effect between number of offspring, liability to schizophrenia (PRS) and sex using Poisson regression. The outcome variable was the number of children and the predictors were schizophrenia PRS, gender and an interaction term (schizophrenia PRS x gender). The interaction term was significant (B=-0.011, SE=0.004, p=0.009) over and above schizophrenia PRS (B=0.01, SE=0.003, $p=8.8x10^{-5}$), gender (B=-0.021, SE=0.004, $p=2.9x10^{-7}$) and age (B=0.015, SE=0.0003, $p<10^{-16}$, respectively). The negative sign of the interaction term B coefficient indicates that the increase in the numbers of children in women with higher genetic loading to schizophrenia was greater than that in men (in the UK Biobank women are coded as "0" and men are coded as "1") and was more pronounced when we compared people with and without offspring (B=-0.08, SE=0.014, $p=5.5x10^{-10}$) using logistic regression. The interaction of sex and PRS on number of offspring is depicted in Figure 3.

Given the non linear relationship between PRS and fecundity in males (Table 3), we further explored the different patterns of relationship between genetic liability to schizophrenia and the total numbers of children of males and females. In Figure 4 we present these as ratios of the mean numbers of offspring born to males and females in 50%, 40%, ..., 10%, 5% and 1% percentiles of the schizophrenia PRS distribution to the total population mean. Relatively more children are born to females with higher PRS, and this increases towards the PRS extreme, plateauing at the 1%-ile. In males, a more complex pattern is observed, showing an

initial decrease in the relative fecundity as PRS increases followed by an increase and then at the extreme end of the distribution, a sharp decrease in relative fecundity.

Discussion

We sought to test the hypothesis that the burden of common risk alleles for schizophrenia is associated with reproductive advantage in subjects without schizophrenia, and that this might explain the persistence of common risk alleles in the population in the face of negative selection in those affected by schizophrenia. Our findings suggest that at a population level, higher common variant genetic loading to schizophrenia is associated with a small reproductive advantage in females. The effect is not seen in males and there is evidence that those with the highest risk scores have fewer offspring. When males and females are considered together there is a net reproductive advantage associated with high PRS. However, the effects of PRS are not sufficient to explain the persistence and allele frequencies of schizophrenia risk alleles in contemporary populations as they are small relative to the effects of schizophrenia itself on fecundity, which at a population level make a stronger impact.

In considering the effects we have seen of PRS in unaffected individuals, we should bear in mind that schizophrenia liability is highly pleiotropic [22], that is, alleles that increase risk for schizophrenia also increase liability to other clinical, cognitive and behavioural traits, some of which may be associated with higher fecundity. For example, people with relatively poorer cognitive ability, a phenotype which is associated with higher schizophrenia polygenic risk [23], may be less motivated, have less opportunity, or less ability to control fertility. It is important to stress that links between fecundity on the one hand and behavioural and cognitive traits

on the other are highly dependent on the contemporary environmental context and that the present results apply to the UK population, one of the wealthiest populations in the world; whether they generalize globally and historically will depend on whether the effects are mediated by primarily by direct impacts on biology or, as we speculate above, an interplay between behavioural and personal traits and social, cultural, religious, and economic norms that are highly variable in history or in different societies. Therefore there is a need to test this hypothesis in different countries and cultures.

It has previously been documented that the effect of schizophrenia on fecundity is more pronounced among males [2, 4, 24]. Our population-based study demonstrates that higher genetic liability to schizophrenia is associated with childlessness in males. We have also shown a significant statistical interaction between gender and genetic liability to schizophrenia. At the extreme on the liability scale of the PRS distribution (1%), males have fewer children than in other PRS %-tiles (see Figure 4).

The sex related differences between number of children and schizophrenia liability is consistent with a family based study [25] in which the sisters of people with schizophrenia had a significantly increased number of children, while brothers of affected individuals had significantly fewer children. Our own and those findings are consistent with the idea of sexual selection [26], whereby number of offspring is substantially the result of female choice. Thus, the association between high PRS and childlessness in males (but not females) is consistent with female mating choices favouring males with characteristics related to lower schizophrenia liability (one might speculate for example better social cognition), whereas male mating choice is less relevant. However, other explanations are possible. Schizophrenia

liability may differentially influence the reproductive biological fitness of males and females, or behavioural and cognitive traits which would be selected against by both sexes may be differentially affected by schizophrenia liability in males and females.

The strengths of our study are 1) the use of the largest European ancestry schizophrenia dataset to date to identify genetic risk loci and 2) the largest available genotyped population cohorts to generate schizophrenia PRS, and the ability to measure liability to schizophrenia directly at a molecular level. These allow us to test the hypothesis with extremely high power; including the ability to test the effects on males and females separately. Limitations are 1) our analysis is based upon common SNPs and the effects on fecundity of rare variants are not tested 2) we are unable to test mechanism, i.e. whether the genetic effects are mediated by behaviour or by reproductive biology 3) as noted, the conclusions are specific to place (UK) and time (participants largely of reproductive age in the latter half of the 20th century).

Acknowledgements

This research was conducted using the UK Biobank resource. UK Biobank was established by the Wellcome Trust, Medical Research Council, Department of Health, Scottish Government and Northwest Regional Development Agency. UK Biobank has also had funding from the Welsh Assembly Government and the British Heart Foundation. Data collection was funded by UK Biobank. The work at Cardiff University was supported by Medical Research Council (MRC) Centre (MR/L010305/1) and Program Grants (G0800509). The CLOZUK sample was genotyped with funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 279227

(CRESTAR Consortium; http://www.crestar-project.eu/). JW is supported by a JMAS Sim Fellowship from the Royal College of Physicians of Edinburgh.

Authors contributions: VEP and ES performed statistical analyses; AP and JW processed the genetic data; JW and GK discussed the results and commented on the manuscript; VEP, MJO and MOD designed the study and wrote the paper with input from all authors.

Disclosures: The authors have nothing to disclose.

Table 1. Association of schizophrenia PRS with number of children in unaffected subjects.

	B	Low 95%CI	Upper 95%CI	p	R^{2*}
All sample	0.006	0.002	0.010	0.0017	$8.4x10^{-5}$
Males	0.0009	-0.005	0.007	0.745	$3.9x10^{-6}$
Females	0.011	0.006	0.016	$7.1x10^{-5}$	0.0005

* Estimated with the linear regression of the number of children of unaffected individuals on the PRS.

Table 2. Schizophrenia PRS and number of children of females.

N of children	Number of females with N* children	OR	L95	U95	p	R^2
0	13863	1	-	-	-	-
1	9926	1.071	1.044	1.099	2.1E-07	0.0010
2	32732	1.031	1.011	1.052	0.0028	0.0002
3	13044	1.042	1.017	1.068	0.0011	0.0003
4	3379	1.085	1.044	1.128	3.3E-05	0.0012
5	742	1.131	1.050	1.219	0.0012	0.0019
6+	317	1.196	1.069	1.336	0.0017	0.0034

* as defined in column 1.

OR is derived as the exponential of the logistic regression Beta coefficient estimate, and represents the change in the odds of membership in the target group, defined by the exact numbers of children with every increase of one standard deviation in PRS.

Table 3. Schizophrenia PRS and number of children of males.

N of children	Number of males with N* children	OR**	L95	U95	p	R2
0	13787	1	-	-	-	-
1	8346	0.967	0.941	0.994	0.015	0.0002
2	27945	0.942	0.923	0.962	2.19E-08	0.0008
3	11089	0.982	0.957	1.007	0.159	4.2E-05
4	3182	0.989	0.951	1.029	0.577	1.3E-07
5	878	1.041	0.972	1.115	0.251	0.0003
6+	454	1.115	1.015	1.225	0.024	0.0015

* as defined in column 1.

** OR is derived as the exponential of the logistic regression Beta coefficient estimate, and represents the change in the odds of membership in the target group, defined by the exact numbers of children with every increase of one standard deviation in PRS.

Figure 1. The percentages of males and females with each category of number of children born to males and females. Number of children is constrained to 6+.

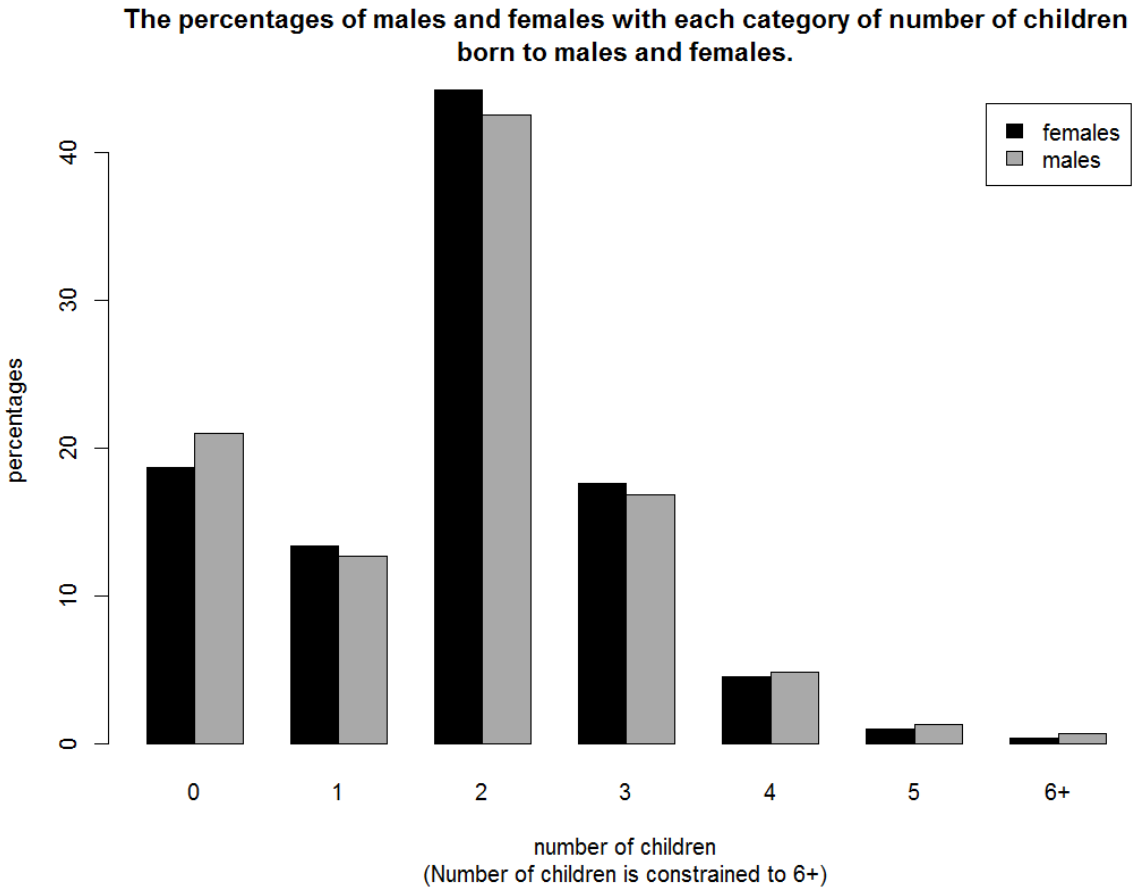


Figure 2. Distribution of schizophrenia PRS and numbers of children of females and males. Number of children is constrained to 6+.

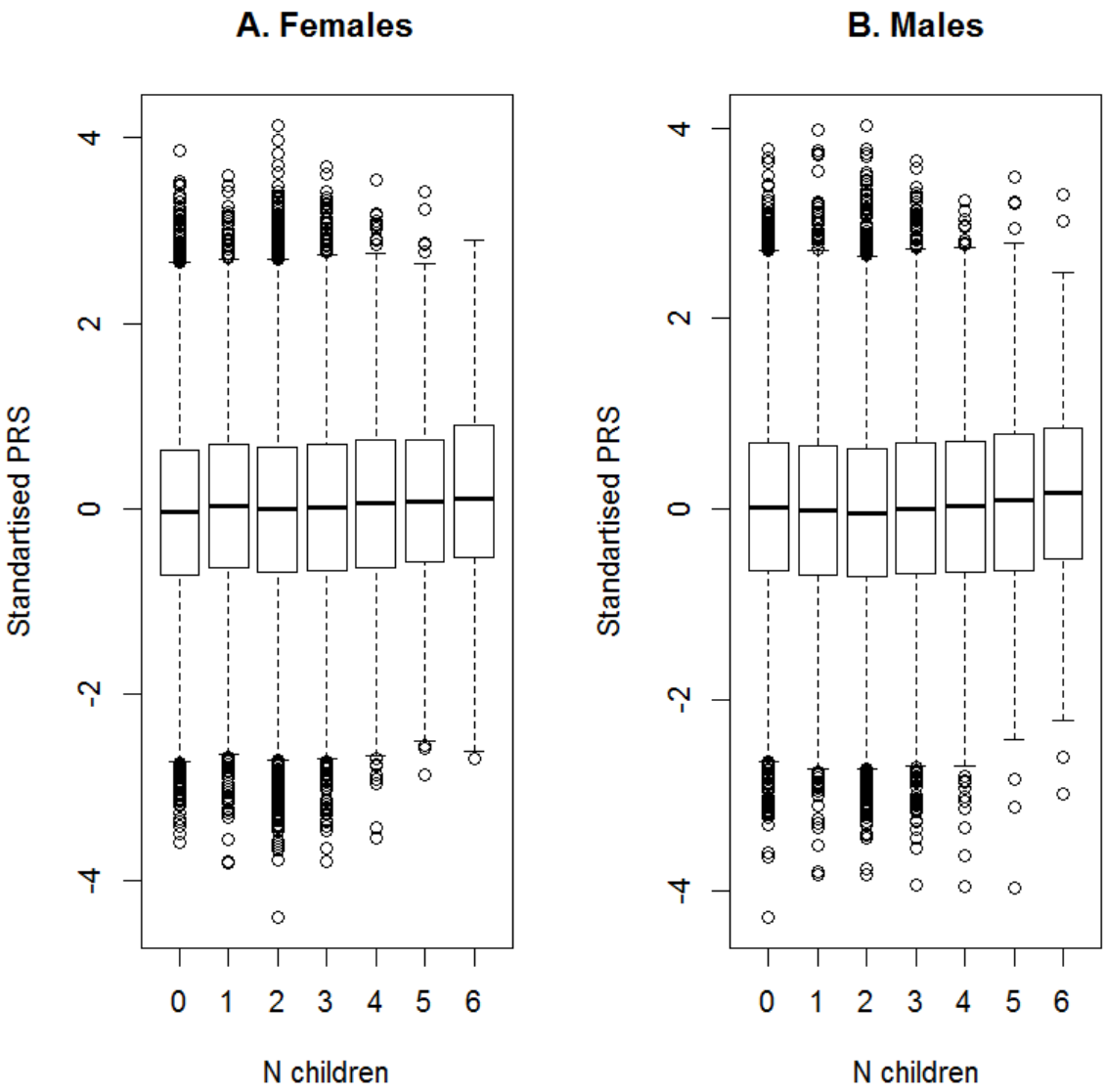
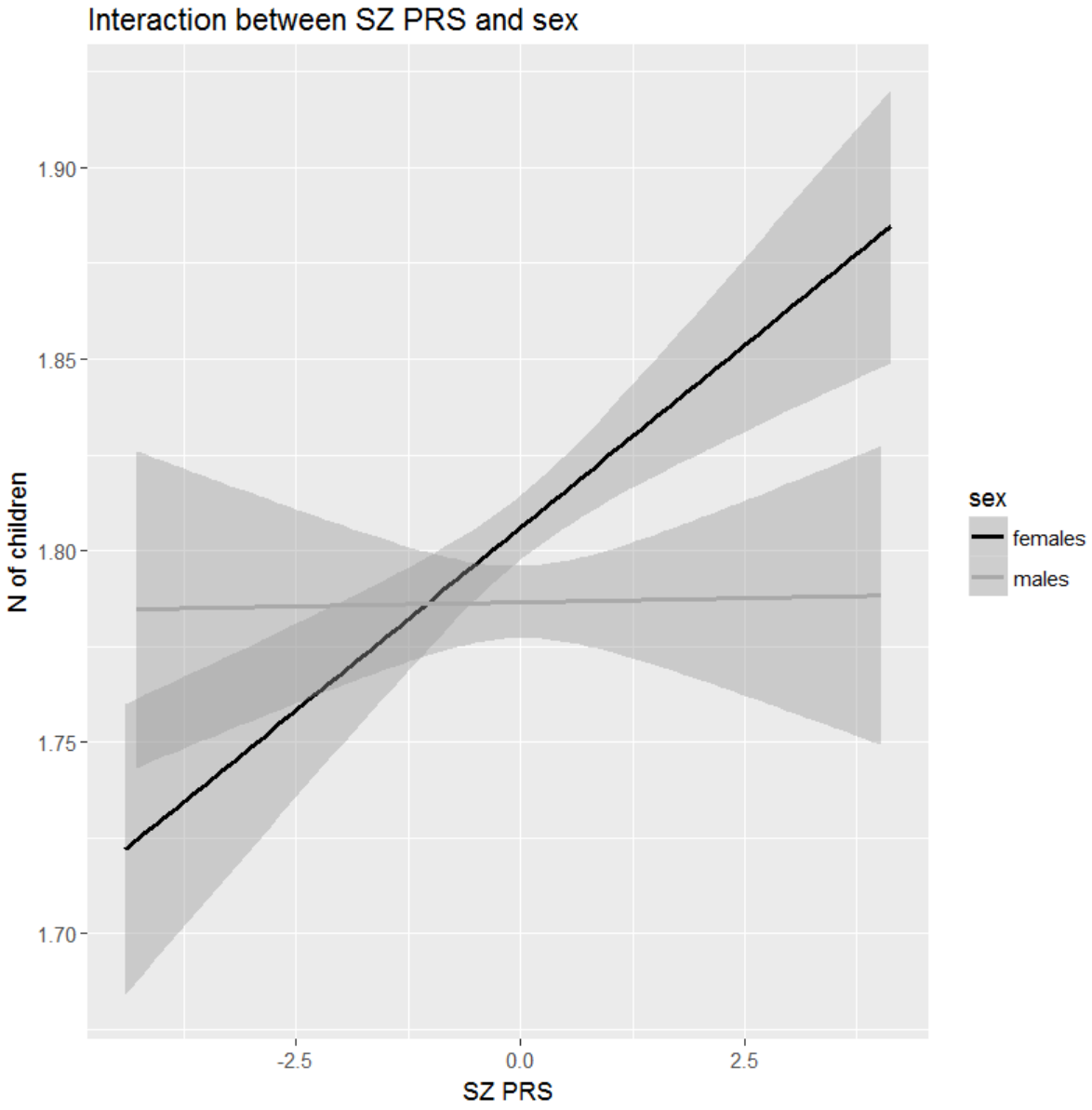
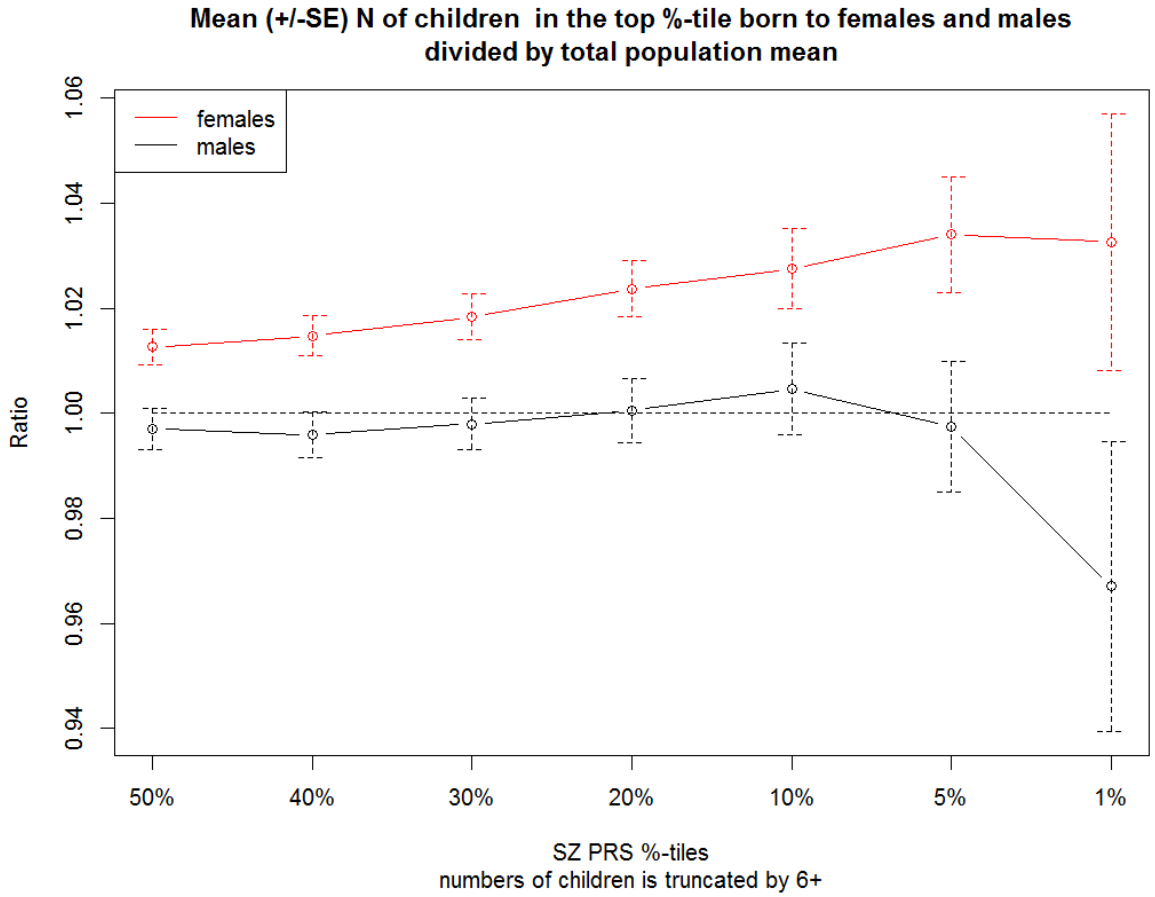


Figure 3. Relationship between schizophrenia PRS of males and females and number of children. The effects of schizophrenia PRS were estimated by logistic regression, with schizophrenia PRS and sex as main effects, including the schizophrenia PRS x sex interaction term, adjusting for age.



Legend: The figure plots the conditional coefficients ("marginal effects") of variables included in multiplicative interaction term in the regression model (i.e. the changes in the coefficient of SZ PRS conditional on sex).

Figure 4. The mean number of children born to males and females in the top %-tiles of the schizophrenia PRS distribution (x-axis) divided by the mean number of children in the whole sample. The numbers of children is truncated by 6+.



Legend: A ratio greater than one (dashed horizontal line) means participants in the corresponding PRS %-ile have a higher number of offspring than the population average.

References

1. Stearns SC, Byars SG, Govindaraju DR, Ewbank D. Measuring selection in contemporary human populations. *Nat Rev Genet.* Sep 2010;11(9):611-622.

2. Haukka J, Suvisaari J, Lonnqvist J. Fertility of patients with schizophrenia, their siblings, and the general population: a cohort study from 1950 to 1959 in Finland. *The American journal of psychiatry.* Mar 2003;160(3):460-463.

3. McGrath JJ, Hearle J, Jenner L, Plant K, Drummond A, Barkla JM. The fertility and fecundity of patients with psychoses. *Acta Psychiatr Scand.* Jun 1999;99(6):441-446.

4. Nanko S, Moridaira J. Reproductive rates in schizophrenic outpatients. *Acta Psychiatr Scand.* Jun 1993;87(6):400-404.

5. Cardno AG, Gottesman, II. Twin studies of schizophrenia: from bow-and-arrow concordances to star wars Mx and functional genomics. *Am J Med Genet.* Spring 2000;97(1):12-17.

6. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry.* Dec 2003;60(12):1187-1192.

7. Rees E, Walters JT, Georgieva L, et al. Analysis of copy number variations at 15 schizophrenia-associated loci. *Br J Psychiatry.* Feb 2014;204(2):108-114.

8. Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* 07/24/print 2014;511(7510):421-427.

9. Singh T, Kurki MI, Curtis D, et al. Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. *Nat Neurosci.* Apr 2016;19(4):571-577.

10. Kirov G, Rees E, Walters JT, et al. The penetrance of copy number variations for schizophrenia and developmental delay. *Biological psychiatry.* Mar 01 2014;75(5):378-385.

11. International Schizophrenia C, Purcell SM, Wray NR, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature.* Aug 6 2009;460(7256):748-752.

12. Purcell SM, Moran JL, Fromer M, et al. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature.* Feb 13 2014;506(7487):185-190.

13. Pardinas AF, Holmans, P., Pocklington, A.J., Escott-Price, V., et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and maintained by background selection. *Nature genetics.* 2017;in press.

14. Jarvik LF, Deckard BS. The Odyssean personality. A survival advantage for carriers of genes predisposing to schizophrenia? *Neuropsychobiology.* 1977;3(2-3):179-191.

15. Mullins N, Ingason A, Porter H, et al. Reproductive fitness and genetic risk of psychiatric disorders in the general population. *Nat Commun.* Jun 13 2017;8:15833.

16. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* Mar 2015;12(3):e1001779.

17. Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods.* Jan 2013;10(1):5-6.

18. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nature genetics.* Jul 22 2012;44(8):955-959.

19. Smith DJ, Escott-Price V, Davies G, et al. Genome-wide analysis of over 106 000 individuals identifies 9 neuroticism-associated loci. *Molecular psychiatry.* Nov 2016;21(11):1644.

20. Fisher RA. The genetical theory of natural selection. Oxford: The Clarendon press; 1930.

21. Edwards AWF. R.A. Fisher's gene-centred view of evolution and the Fundamental Theorem of Natural Selection. *Biological Reviews.* 2014;89(1):135-147.

22. O'Donovan MC, Owen MJ. The implications of the shared genetics of psychiatric disorders. *Nat Med.* Nov 2016;22(11):1214-1219.

23. Hubbard L, Tansey KE, Rai D, et al. Evidence of Common Genetic Overlap Between Schizophrenia and Cognition. *Schizophr Bull.* May 2016;42(3):832-842.
24. Howard LM, Kumar C, Leese M, Thornicroft G. The general fertility rate in women with psychotic disorders. *The American journal of psychiatry.* Jun 2002;159(6):991-997.
25. Power RA, Kyaga S, Uher R, et al. Fecundity of patients with schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, or substance abuse vs their unaffected siblings. *JAMA Psychiatry.* Jan 2013;70(1):22-30.
26. Jones AG, Ratterman NL. Mate choice and sexual selection: what have we learned since Darwin? *Proc Natl Acad Sci U S A.* Jun 16 2009;106 Suppl 1:10001-10008.

SUPPLEMENTARY NOTE:
Fecundity effects on schizophrenia liability in the general population

DATA:

1- The heritability (h^2) of liability for schizophrenia that is explained by its polygenic risk score (PRS) has been estimated as **0.05153** (Pardiñas *et al.* 2018). At the moment that study was carried out, this result was based on the largest available training sample (34,241 cases and 45,604 controls; Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014) and the largest available testing sample (5,220 cases and 18,823 controls; Pardiñas *et al.* 2018). The PRS used in the present study was produced by combining both samples, and thus this h^2 value could be a slight underestimation.

2- The lifetime prevalence of schizophrenia is **0.0072,** which is the median of a review of 188 independent studies (McGrath *et al*. 2008). According to the authors, the median is a more appropriate estimation of the central value than the mean due to the skewed distribution of prevalences reported by different studies.

3- The ratio of fecundities of patients with schizophrenia to unaffected individuals is **0.35** (Power *et al.* 2013). This figure might be an underestimation as it is based on a sample of people older than 40 years of age and does not account for the reduced life expectancy seen in schizophrenia (Laursen *et al.* 2014).

4- In the present study, the linear regression of the number of children of unaffected individuals on the PRS gives r^2=**0.000084** with a 95% confidence interval r^2=[**0.000023 – 0.000184**]. The direction of effect shows that higher PRS leads to more children.

5- In the present study, the mean and variance of number of children of unaffected individuals are **1.8087** and **1.3988** respectively.

RESULTS:

1- **Genetic variance for fecundity that is explained by the liability threshold model of schizophrenia**:

The liability threshold model is responsible for a fraction of the phenotypic variance for fecundity which can be calculated from the data above. The mean fecundity in the general population (i.e. affected and unaffected individuals) explained by the liability threshold model, considering only the reduction in fecundity caused by schizophrenia, is:

$$M = (1 - 0.0072) * 1 + 0.0072 * 0.35 = 0.99532$$

After scaling for average fecundity=1, the phenotypic variance for fecundity in the general population, which is explained by the liability threshold model, is

$$V = [(1 - 0.0072) * 1^2 + 0.0072 * 0.35^2 - M^2]/M^2 = 0.0030$$

These calculations implicitly assume that the phenotypic variance is generated by the onset of schizophrenia (i.e. the disorder leads to a reduction in fecundity). Therefore, the heritability of schizophrenia equals the heritability of fecundity explained by schizophrenia. Consequently, the genetic variance for fecundity in the general population explained by variation in schizophrenia PRS under the liability threshold

model is the product of the phenotypic variance for fecundity and the heritability explained by PRS:

$$V_{g\_SCZ} = 0.0030 * 0.05153 = 0.000155$$

2- **Genetic variance for fecundity in unaffected individuals explained by variation in PRS:**

The squared correlation between fecundity and PRS is the proportion of the variation in fecundity that is explained by PRS. After scaling for average fecundity=1, we get:

$$V_{g\_unaffected} = 0.000084 * \frac{1.3988}{1.8087^2} = 0.000036$$

The largest source of error in all these calculations corresponds, by far, to the estimated r^2. Using its 95% confidence interval we can get an approximation of plausible values for the genetic variance:

$$V_{g\_unaffected} = [0.000010 - \ 0.000079]$$

CONCLUSION:

If we consider that fecundity is a measure of fitness, genetic variance for fecundity is the expected change of fitness by selection in a single generation (Fisher 1930, Edwards 2014). Thus, the comparison of variances gives the relative magnitude of the two selective forces:

$$V_{g\_SCZ}/V_{g\_unaffected} = 4.31$$

A range can be approximated using the 95% confidence interval of $V_{g\_unaffected}$:

$$V_{g\_SCZ}/V_{g\_unaffected} = [1.96 \ - \ 15.50]$$

This ratio refers to the genetic variance explained by PRS in this study. Strictly, this cannot be extended to the whole genetic system affecting schizophrenia, unless we accept that the expected fertility of unaffected individuals is a linear function of the liability score of individuals for schizophrenia. Thus, although the present study detects a selective force acting in favour of the persistence of schizophrenia risk alleles in the general population, its effect is offset by the removal of these alleles caused by the reduced fecundity of affected individuals.

REFERENCES:

Edwards, A.W.F., 2014. R.A. Fisher's gene-centred view of evolution and the Fundamental Theorem of Natural Selection. Biological Reviews 89**,** 135-147.

Fisher, R.A., 1930. The genetical theory of natural selection. The Clarendon press, Oxford.

Laursen, T.M., Nordentoft, M., and Mortensen, P.B., 2014. Excess Early Mortality in Schizophrenia. Annual Review of Clinical Psychology 10**,** 425-448.

McGrath, J., Saha, S., Chant, D., and Welham, J., 2008. Schizophrenia: A Concise Overview of Incidence, Prevalence, and Mortality. Epidemiologic Reviews 30**,** 67-76.

Pardiñas, A.F., Holmans, P., Pocklington, A.J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S.E., Bishop, S., Cameron, D., et al., 2018. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nature Genetics.

Power, R.A., Kyaga, S., Uher, R., MacCabe, J.H., Långström, N., Landen, M., McGuffin, P., Lewis, C.M., Lichtenstein, P., et al., 2013. Fecundity of patients with schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, or substance abuse vs their unaffected siblings. JAMA psychiatry 70**,** 22-30.

Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014. Biological

insights from 108 schizophrenia-associated genetic loci. Nature 511**,** 421-427.