Sex-biased reduction in reproductive success drives selective constraint on human genes

Eugene J. Gardner[1], Matthew D. C. Neville[1], Kaitlin E. Samocha[1], Kieron Barclay[2,3,4], Martin Kolk[3], Mari E. K. Niemi[1], George Kirov[5], Hilary C. Martin[1], Matthew E. Hurles[1,*]

[1]Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge, Hinxton, United Kingdom
[2]Max Planck Institute for Demographic Research, Rostock, Germany
[3]Demography Unit, Department of Sociology, Stockholm University, Stockholm, Sweden
[4]Swedish Collegium for Advanced Study, Uppsala, Sweden
[5]Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, UK
*meh@sanger.ac.uk

Summary

Genome-wide sequencing of human populations has revealed substantial variation among genes in the intensity of purifying selection acting on damaging genetic variants. While genes under the strongest selective constraint are highly enriched for Mendelian disorders, most of these genes are not associated with disease and therefore the nature of the selection acting on them is not known. Here we show that genetic variants that damage these genes reduce reproductive success substantially in males but much less so in females. We present evidence that this reduction is mediated primarily by cognitive and behavioural traits, which renders male carriers of such variants less likely to find mating partners. These findings represent strong genetic evidence that sexual selection mediated through female mate choice is shaping the gene pool of contemporary human populations. Furthermore, these results suggest that sexual selection accounts for 21% of purifying selection against heterozygous variants that ablate protein-coding genes.

Main text

The most damaging genetic variants, gene deletions and protein-truncating variants (PTVs), are removed from the population by selection with strength that varies substantially from gene to gene[1,2]. The strength of selection against heterozygous PTVs has been estimated by the selection coefficient, s_{het}, which exhibits a continuous spectrum across human genes[3,4], although most attention has been focused on a subset of ~3,000 genes with the highest probability of Loss-of-function Intolerance (pLI)[1].

The selection pressures acting on these most selectively constrained genes have not been fully characterised, but, *a priori*, could include natural selection against variants increasing pre-reproductive mortality or decreasing fertility, and sexual selection acting on mate choice or intra-sexual competition[5]. Gene deletions and PTVs in these genes have been shown to be associated with lower educational attainment[6,7] and general intelligence[8], as well as increased risk of intellectual disability, and some psychiatric disorders[9]. Moreover, these

constrained genes are strongly enriched for dominant early-onset Mendelian diseases (including many neurodevelopmental disorders), many of which are associated with increased pre-reproductive mortality, indicating that natural selection likely makes a substantive contribution to this selective constraint. However, the majority (65%) of constrained genes (pLI>0.9) have not yet been associated with a Mendelian disease. This raises the fundamental question of whether natural selection represents the sole mechanism imposing this form of selective constraint on human genes, or whether other forms of selection are at work.

To explore the nature of selection acting on these genes we identified PTVs and genic deletions in the UK Biobank[10] comprising largely post-reproductive individuals (median age at recruitment: 58 years, range: 39-73 years, birth years: 1934-1970; Supplementary Figure 1), and investigated the association with reproductive success. We called large copy number variants (deletions and duplications) from SNP genotyping array data on 340,925 unrelated participants of European ancestry (Supplementary Figure 2), and identified PTVs from exome sequencing among a subset of 34,812 participants (Supplementary Figure 3)[11]. For each participant, we then calculated the cumulative burden of private (only observed in one individual), heterozygous genic deletions and PTVs by combining s_{het} selection coefficients of each autosomal gene impacted by these variants (under the assumption that fitness is multiplicative, see Methods), which we term their s_{het} burden. The distribution of s_{het} burden was statistically indistinguishable between males and females: for participants with only genic deletion data available, 0.56% and 0.55% respectively had an s_{het} burden ≥ 0.15 (Fisher's p=0.61; Figure 1B), and for participants with both genic deletion and PTV data available the analogous proportions were 6.99% and 7.06% (Fisher's p=0.80; Figure 1C).

We assessed the relationship between s_{het} burden and number of children, using a linear regression correcting for age and population structure (Methods; Supplementary Figure 4; Supplementary Table 1). We observed that an s_{het} burden of 1 is associated with a decrease in the overall total number of overall children for both males (0.53 fewer children [95% CI 0.36-0.71], $p=2.2\times10^{-9}$) and females (0.17 fewer children [95% CI 0.01-0.33], p=0.04) when combining results from deletion and PTV-based analyses.

To determine if the observed effect of s_{het} burden was due to an actual reduction in overall number of children or a result of selection against having children at all, we performed two analyses. Firstly, we evaluated childlessness using logistic regression and again observed a striking sex difference in participants' probability of having any children given their s_{het} burden, for both PTVs and genic deletions (Figure 1A). Combining the analyses of genic deletions and PTVs, we found that an s_{het} burden of 1 decreases the probability of males having any children (OR=0.29 [95 %CI 0.21-0.40], $p=9.0\times10^{-14}$) much more than females (OR=0.69 [95% CI 0.49-0.98], p=0.04). We also observed that private duplications and likely damaging private missense variants exhibit similar but weaker effects on childlessness (Supplementary Figure 5). Secondly, if we remove childless individuals from the analysis, s_{het} burden ceases to have a significant effect on the number of offspring, confirming that the observed decrease in reproductive success is determined largely by increased childlessness (Supplementary Figure 6).

We also considered whether ascertainment biases or differences in fertility between the UK Biobank population and the UK population as a whole could affect these results. As UK Biobank participants included in these analyses are biased towards females (54%), the observed sex bias is not due to having greater statistical power to detect an effect on reproductive success in males. Likewise, fertility rates between the UK Biobank population and the UK population as a whole are highly similar; the average total fertility rate in the UK from 1983-2000 (where data are available for both males and females), when UK Biobank participants would have been reproductively active, was 1.78 for males[12] and 1.76 for females[13], which is very similar to that observed among UK Biobank participants (average number of children for males = 1.77; females = 1.80).

We observed consistent sex-bias in the association of s_{het} burden with childlessness when performing this analysis in different ways, including: (i) limiting analyses to carriers of private genic deletions or PTVs in the genes under most selective constraint (following thresholds set by their authors: pLI ≥ 0.9 or s_{het} ≥ 0.15; Supplementary Figure 7), (ii) extending analysis to more frequent, but still rare genic deletions and PTVs (Supplementary Figure 8), (iii) excluding genes known to cause Mendelian disorders (male OR=0.32 [95% CI 0.21-0.47], $p=2.1 \times 10^{-8}$), and (iv) restricting analysis to individuals in specific age ranges (Supplementary Figure 9).

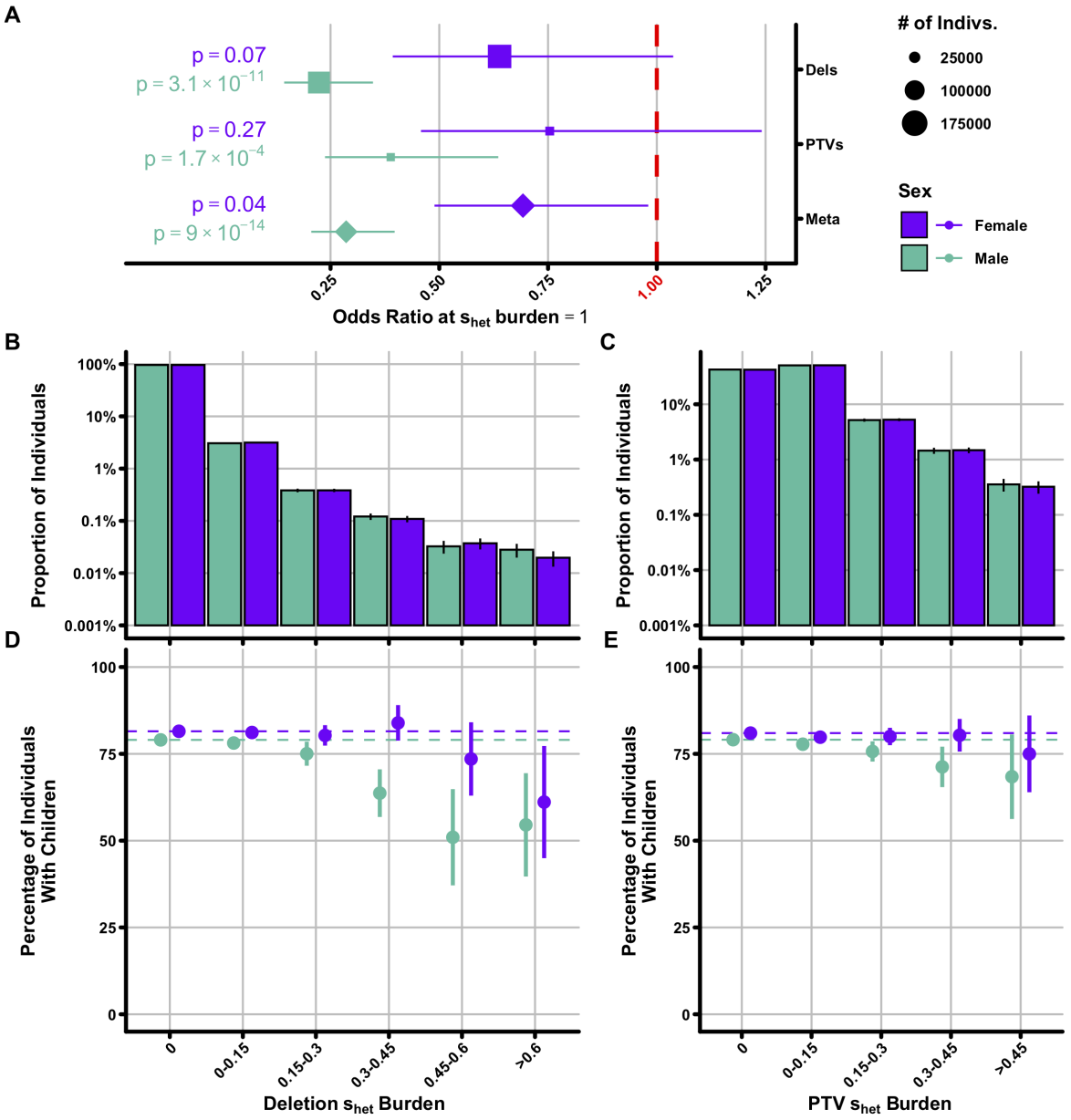


Figure 1. Differences in male and female reproductive success as a function of cumulative rare deleterious genetic variation.

(A) Odds ratio estimates for the effect of cumulative deleterious variation for deletions, PTVs, and a combined meta-analysis on childlessness separated for males (jade) and females (violet). Number of individuals included in each analysis is indicated by the size of the point. (B; C) Proportion of individuals in 0.15 s_{het} bins for deletions (B) and PTVs (C). (D; E) Percentage of individuals with children in bins based on s_{het} burden for deletions (D) and PTVs (E). All error bars are 95% confidence intervals calculated on the population proportion.

We evaluated three hypotheses that could account for increased childlessness: (i) impaired fertility (e.g. inability to produce viable gametes), (ii) adverse health conditions, and (iii) cognitive and behavioural factors (which could decrease ability to find a mate, or increase voluntary childlessness). We observed that s_{het} burden does increase the risk of male infertility, albeit with wide confidence intervals, (OR=8.93 [95% CI 1.55-51.51], p=0.01) but not female infertility (OR=0.54 [95% CI 0.17-1.75], p=0.31); however, three lines of evidence

suggest impaired fertility is not the predominant cause of the sex-biased association between s_{het} burden and childlessness. First, when introducing male infertility status available from combined health outcomes data for all UK Biobank participants (combined hospital episode statistics, primary care records, self-reported conditions and death records) as a covariate in the association testing, we observed minimal reduction in the strength of the association between s_{het} burden and male childlessness (OR=0.29 [95% CI 0.21-0.40], $p=1.5x10^{-13}$; Supplementary Table 2). Second, we observed minimal change in the association between s_{het} burden and male reproductive success after removing all 150 autosomal genes for which at least limited evidence exists of an association to male infertility (OR=0.29 [95% CI 0.20-0.40], $p=1.0x10^{-13}$)[14] or all 742 genes associated with male infertility in mice (OR=0.29 [95% CI 0.20-0.40], $p=6.8x10^{-13}$)[15]. Finally, genes under the highest selective constraint ($s_{het} \geq 0.15$) are not associated with higher expression levels in testis, unlike the genes currently known to be associated with male infertility (Supplementary Figure 10). Together, these findings are consistent with a previous study that sought but did not find a widespread role for highly-penetrant dominant deletions in spermatogenic failure[16].

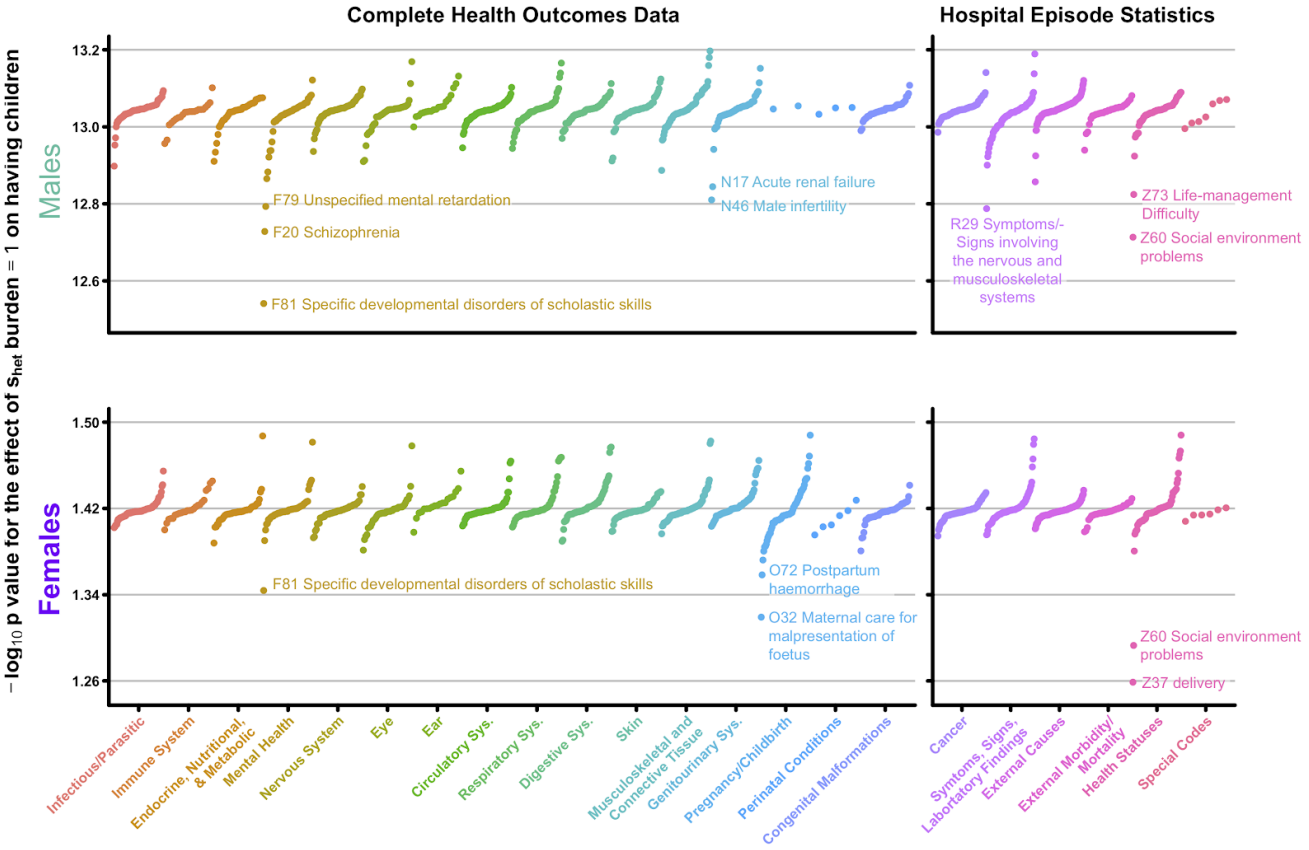


Figure 2. Modulation of the relationship between s_{het} burden and childlessness by various disorders.

Plotted is the deletion and PTV meta-analysis -log_{10} p value for the association between s_{het} burden and having children, corrected by one of 1,294 ICD-10 codes from a combination of general practitioner, hospital episode records, and self-reported conditions (left) or hospital episode records alone (right) separately for males (top) and females (bottom). Results are ordered first by ICD-10 chapter (x-axis) and then by increasing -log_{10} p value (y-axis). Visual outliers are labelled.

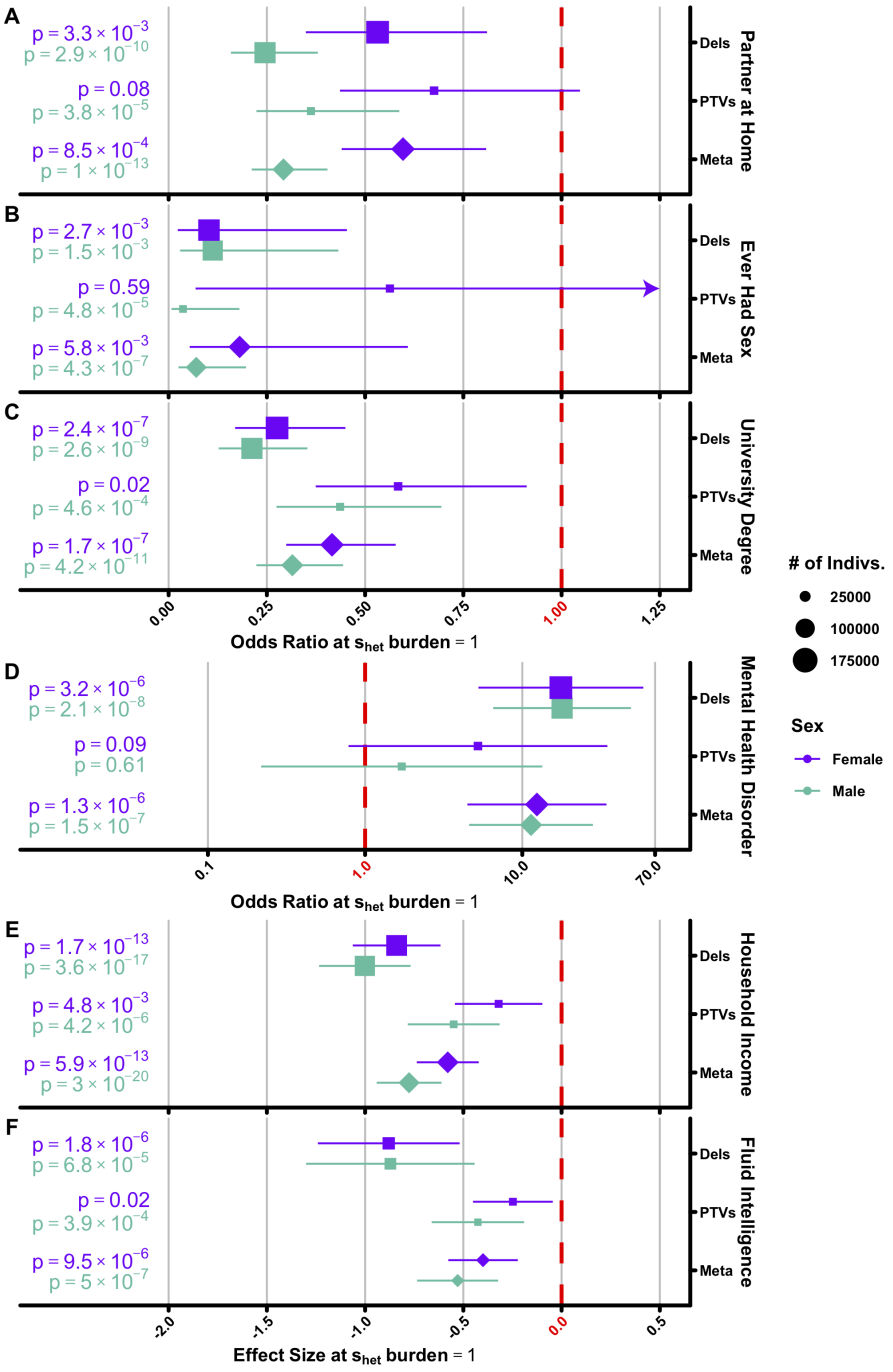
Previous studies have shown that a variety of physical birth defects are associated with reduced reproductive success[17,18]. We comprehensively assessed whether any adverse health conditions contributed to the association between s_{het} burden and childlessness. We independently tested 19,154 ICD-10 codes (from both hospital episode statistics and combined health outcomes data across four levels of the ICD-10 hierarchy; Methods) as a covariate in the association test of s_{het} burden on childlessness (Figure 2; Supplementary Figures 11-13; Supplementary Table 2; Methods). We found that while many ICD-10 codes are associated with having children, in particular positive associations with male-specific codes for elective sterilisation and female-specific codes associated with pregnancy and childbirth (Supplementary Figure 11), correcting for any ICD-10 code had minimal impact on the strength of association between s_{het} burden and childlessness (Figure 2; Supplementary Figures 12, 13). The biggest impact on the association of s_{het} burden and male childlessness was observed for 'developmental disorders of scholastic skills', although this only modestly reduced the significance of the association from $p=9.0x10^{-14}$ to $p=2.9x10^{-13}$ (Figure 2; Supplementary Table 2). In addition to diseases, clinically annotated ICD-10 codes are also available for a range of factors denoting health status and contact with health services (Chapter XXI – Health Statuses). We noted that two of these also had a modest impact on the association of s_{het} burden and male childlessness when included as covariates in association testing (Figure 2; Supplementary Table 2). These were codes relating to life management difficulty (code Z73) and social environment problems (code Z60), with the latter driven primarily by the status of living alone (code Z60.2). Both of these two health-related factors are positively associated with s_{het} burden.

There is substantial existing evidence that behavioural and cognitive traits influence reproductive success in a sex-biased manner. First, the reduced reproductive success associated with a range of psychiatric disorders is much more pronounced in males than in females[19]. Second, personality traits associated with increased reproductive success differ between males and females, with increased extraversion in males but greater neuroticism in females being linked to increased reproductive success[20]. Third, although the most highly ranked mate characteristics are highly concordant between the sexes[21], some mate preferences differ between the sexes, with males placing greater value on physical attractiveness and females valuing cues relating to earning potential[20–23]. Finally, low socioeconomic status and low educational attainment have been more strongly linked to increased childlessness in males than females across populations[24–27]. This has typically been ascribed to males of lower socioeconomic status finding it harder to attract a partner[28,29].

Some of these observations about sex-biased reproductive success have been related to sexual selection by mate choice, in which one sex (typically female) tends to be more discriminating in their choice of mates. Alternative theories have been proposed regarding the causes of sexual selection, including those that focus on disparities in gamete size (Darwin-Bateman paradigm[30]) and others that focus on differential parental investment[31]. This latter hypothesis posits that sexual selection by mate choice is driven, in large part, by the sex that invests more in offspring (typically female) being more discriminating in their choice of mates, especially with regard to their potential to invest in offspring. However, a sex-biased reduction in reproductive success need not be caused by sex differences in mate

preferences. Sex-biased reproductive success could also result from a sex bias in trait severity coupled to mate choice preferences that are not sex-biased. These mechanisms are not mutually exclusive; both could be contributing to an overall sex-biased reduction in reproductive success, albeit on different traits.

A key prediction of the hypothesis that differential mate choice underpins the observation of a male-biased association of s_{het} burden with increased childlessness is that males with a high s_{het} burden should find it harder to find mates than females. We observed that UK Biobank participants with high s_{het} burden were significantly less likely to have reported currently living with a partner (at the time of assessment), consistent with the findings from the ICD10 codes, and that, like reproductive success, this effect was significantly stronger in males than in females (Figure 3A). UK Biobank males currently living with a partner are also much more likely to have children (OR=5.80 [95% CI 5.65-5.96] , $p < 1x10^{-100}$; Supplementary Figure 14). We note that the status of currently living with a partner is an imperfect proxy for partner status during peak reproductive years, but the latter information is not currently available in UK Biobank. We also found that s_{het} burden was significantly positively associated with reporting never having had sex for both male (OR=0.07 [95% CI 0.03-0.20], $p=4.3x10^{-07}$) and female (OR=0.18 [95% CI 0.05-0.61], $p=5.8x10^{-03}$) UK Biobank participants, without significant sex-bias (Figure 3B). Additionally, while same sex sexual behaviour is strongly associated with increased childlessness in UK Biobank (male OR=0.14 [95% CI 0.13-0.15], $p<1x10^{-100}$; female OR=0.27 [95% CI 0.25-0.29] $p<1x10^{-100}$; Supplementary Figure 14), we observed no significant impact of s_{het} burden on the likelihood of having engaged in same sex sexual behaviour among males (OR=1.27 [95% CI 0.59-2.69], p=0.54; Supplementary Figure 15) nor did we observe any change in the relationship between s_{het} burden and childlessness when excluding individuals who engaged in same-sex sexual behaviour from our primary model (OR=0.27 [95% CI 0.19-0.39], $p=1.2x10^{-13}$).



221

222 ***Figure 3. Effect of s_{het} burden on traits known to be associated with reproductive success.***
223 *Shown are similar plots to Figure 1A, except for six phenotypes which have been previously*
224 *associated with reproductive success: (A) having a partner at home, (B) ever having engaged in*
225 *sexual intercourse, (C) educational attainment as measured by having a university degree, (D)*
226 *household income (as measured by income bracket and corrected for having a partner at home; see*
227 *methods), (E) fluid intelligence (in standard deviations), and (F) having a mental health disorder. For*
228 *each trait, we tested using a logistic (A,B,C,D) or linear (E,F) model the effect of s_{het} burden on each*
229 *phenotype shown above, corrected for age, age^2, and the first thirty ancestry principal components.*
230 *The arrow in plot (B) indicates the upper confidence interval for female PTVs is outside the range of*
231 *the X-axis. Note that plot (D) is in log rather than linear scale.*
232

We explored in UK Biobank whether the impact of s_{het} burden on reproductive success might plausibly be mediated through some of the specific factors highlighted by the previous psychiatric, demographic and psychosocial research summarised above. Firstly we investigated the impact of s_{het} burden on cognition as measured by fluid intelligence in 110,190 (51,378 males, 58,812 females) UK Biobank participants. We found that s_{het} burden was associated with significantly reduced fluid intelligence scores of males and females with similar effect sizes (Figure 3F). Increasing s_{het} burden is also associated with lower educational attainment and household income (Figure 3C,E), again with similar effect sizes in males and females. To evaluate the potential impact of this reduced cognition on male reproductive success, we extended previously published work relating the results of IQ tests taken by 95% of Swedish males (during military conscription) to their completed family size[32]. We estimated that the decrement in cognition observed in UK Biobank males could account for 6% [95% CI 5%-9%] of the reduced male reproductive success associated with high s_{het} burden (Supplementary Figure 16, Methods). We also note that the decrease of reproductive success with decreasing IQ was most pronounced in males with IQ<70 (Supplementary Figure 17)[32], who are likely depleted in UK Biobank relative to the general population.

Analysis of psychiatric disorders in UK Biobank is complicated by both recruitment bias away from more severe psychiatric disorders and incomplete data on participants[10,33–35]. The most comprehensive data are available on a subset of UK Biobank individuals from a mental health questionnaire for which participants were invited by email (n = 157,366)[34]. We observed that a high s_{het} burden was very strongly associated with not having an email address (male OR=0.30 [95% CI 0.21-0.41], $p=8.4x10^{-14}$; female OR=0.48 [95% CI 0.35-0.65], $p=3.1x10^{-6}$; Supplementary Figure 18), which likely explains why individuals with a high s_{het} burden were much less likely to complete the questionnaire (male OR=0.43 [95% CI 0.31-0.60], $p=7.6x10^{-7}$; female OR=0.41 [95% CI 0.30-0.56], $p=1.7x10^{-8}$; Supplementary Figure 18). Therefore we focused analyses of mental health disorders on the complete health outcomes data available for all participants. These data corroborate a previous finding[9] that high s_{het} burden increases the risk of psychiatric disorders previously associated with reduced reproductive success (intellectual disability, schizophrenia, autism, attention deficit hyperactive disorder, and bipolar disorder; Figure 3D, Supplementary Figure 19)[19], and that these psychiatric disorders are associated with increased childlessness in both male (OR=0.35 [95% CI 0.32-0.39], $p=8.3x10^{-85}$) and female (OR=0.58 [95% CI 0.52-0.65], $p=2.7x10^{-19}$) UK Biobank participants, albeit with substantial sex-bias (Supplementary Figure 14). This finding accords with a previous study showing that copy number variants associated with increased risk of schizophrenia are also associated with disproportionately reduced reproductive success in males[36]. Carriers of well-characterised neurodevelopmental disorder-associated copy number variants, which include those with a strong association to schizophrenia (Methods), only account for 3.7% (n = 12,608) of individuals in UK Biobank. Removal of these individuals from the dataset does not significantly alter the association of s_{het} burden with reduced male reproductive success (OR=0.29 [95% CI 0.21-0.41], $p=3.1x10^{-12}$).

We subsequently tested the impact of s_{het} burden on childlessness only for individuals without any evidence of a mental health disorder associated with reduced reproductive

success (from hospital episode statistics, combined health outcomes data, or the mental health questionnaire). We observed very similar effect sizes to when analysing all individuals (male OR=0.31 [95% CI 0.22-0.43], $p=1.0x10^{-11}$; female OR=0.75 [95% CI 0.53-1.07], p=0.12), suggesting that the effect on childlessness is not predominantly driven by this subset of mental health disorders. We explored this further, using data external to UK Biobank that are less affected by the limitations described above. Using previous estimates of the increased risk of mental health disorders caused by PTVs in highly constrained genes[9], and the reduced reproductive success associated with those disorders[19], we estimated that these mental health disorders could account for 15% [7 - 33%] of the reduced male reproductive success associated with high s_{het} burden (Methods). Thus, in UK Biobank, both reduced fluid intelligence and increased risk of psychiatric disorders account for only modest proportions of increased male childlessness due to s_{het} burden.

We next used multiple regression to explore how much of the association between s_{het} burden and childlessness can be accounted for by the factors described above (where available for the entire cohort), namely: living with a partner, having had sex, having a mental health disorder associated with reduced reproductive success, having a university degree and having an infertility code in health records. Collectively, these factors can account for most of the association of s_{het} burden and childlessness in both males (75%) and females (65%), as assessed by the difference in incremental Nagelkerke's r^2 of deletion s_{het} burden between models (Methods; Figure 4A; Supplementary Figure 20). By far the biggest contribution comes from living with a partner and having had sex, which together can account for 73% of the association in males.

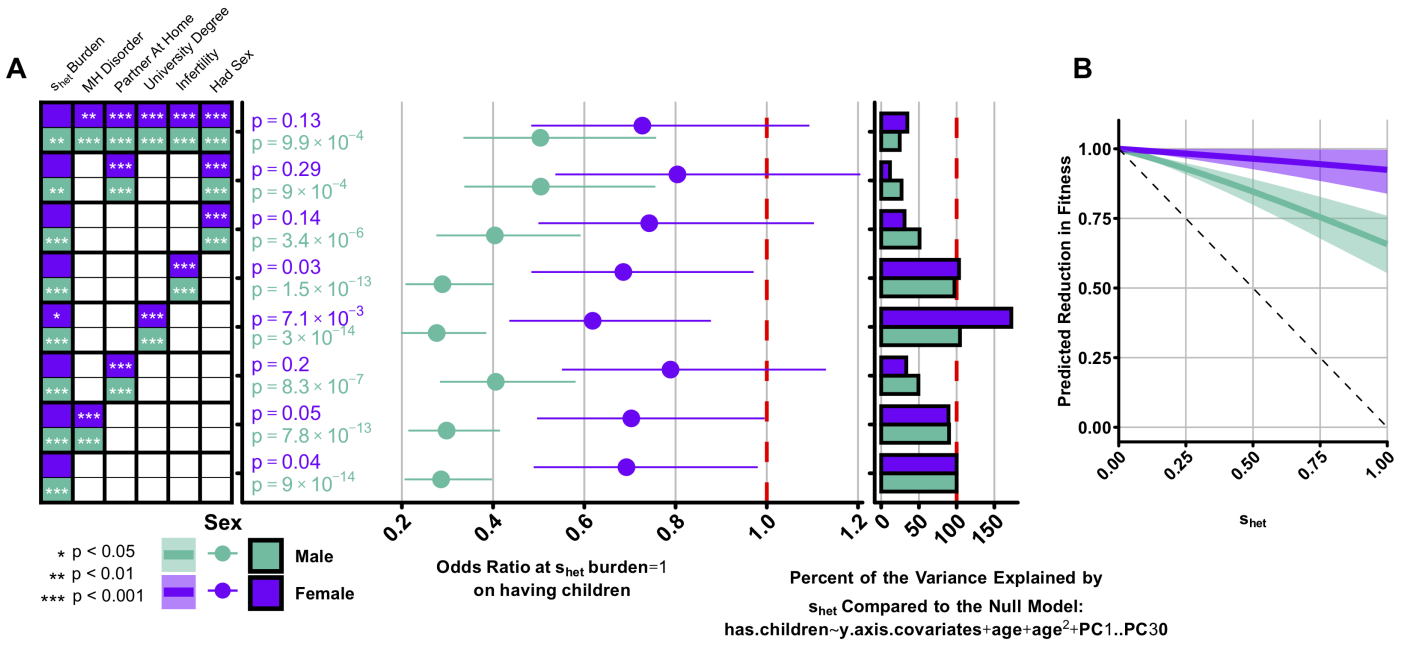


Figure 4. The role of individual phenotypes in the relationship between s_{het} burden and fitness. *(A) Odds ratio estimates for the effect of cumulative deleterious variation for a combined meta-analysis (deletions + PTVs) on childlessness (middle), corrected for a combination of whether or not a study participant has a mental health disorder, a partner at home, a university degree, infertility, or ever had sex; traits included in each model are indicated as coloured boxes (males – jade, females – violet) on the y-axis. Stars within boxes indicate significance level with childlessness for each covariate independently when correcting for deletion s_{het} burden. For all possible combinations of these traits, see Supplementary Figure 20. As indicated by coloured boxes, all models include s_{het} burden and were run separately for males (jade) and females (violet). The marginal bar plot to the right gives the proportion of the variance in childlessness explained by s_{het} burden as calculated for deletions only, scaled to the model which only includes s_{het} burden (i.e. the model on the bottom of the plot). (B) Predicted reduction in overall fitness as a factor of individual s_{het} burden. Displayed is the expected reduction in fitness as a factor of increasing s_{het} burden, independently for each sex. Error is shown as the lighter shaded area surrounding the trend line, and is based on the confidence intervals on the odds ratio as determined by our logistic regression model (Figure 1A; Methods). The dashed line represents the theoretical reduction in fertility as predicted by s_{het}^3.*

Overall, we estimate that reduced reproductive success due to s_{het} burden explains 21% [12-30%] (Figure 4B; Supplementary Figure 21) of the total reduction in fitness expected due to purifying selection against PTVs as predicted by s_{het} (Methods)[3], with this reduction in fitness being much stronger in males. This suggests that such selection may not be borne equally by males and females. We note that the total reduction in fitness predicted by s_{het} will include a substantive contribution from pre-reproductive mortality, which is not quantified here. We also note that current estimates of s_{het} are based on data from aggregated research cohorts, and may thus be biased upwards. This is because individuals with high s_{het} burden are likely to be under-represented within research cohorts (since participation in research has been shown to be biased with respect to gender, socioeconomic status and genetic variation[37]), so PTVs within genes under strong selective constraint may well be segregating at higher frequencies in the general population than in research cohorts. This bias could result in the true value of s_{het} being lower than currently estimated, and,

consequently, the contribution of reduced reproductive success to the overall reduction in fitness due to purifying selection being greater than estimated here.

These estimates of reproductive success and selection coefficients are inevitably reflective of a population at a particular point in time. The proportionate contribution of reduced reproductive success to the overall reduction in fitness associated with genic purifying selection is likely to change over time. Medical advances over recent decades have altered the landscape of infertility and pre-reproductive mortality substantially. Moreover, overall childlessness is highly dynamic over time. Demographic data demonstrate that population-wide childlessness can double in just two decades, a nationwide trend that is readily apparent in UK Biobank (Supplementary Figure 1). We cannot discount that sex-biased sociodemographic factors, in addition to sexual selection, could also be contributing to dampening the apparent association between s_{het} burden and childlessness in women. Higher educational attainment has been shown to be one of the factors most strongly positively associated with childlessness in a female-biased manner[26]. Indeed, when we correct the association between s_{het} burden and childlessness for having a university degree, we see a more significant effect of s_{het} burden on childlessness in females (OR=0.62 [95% CI 0.44-0.88], $p=7.1 \times 10^{-3}$; Figure 4A; Supplementary Figure 20); but the effect on male childlessness remains considerably stronger than in females.

In summary, we find that reduced reproductive success, especially in males, makes a substantial contribution to purifying selection acting on human genes, and that this is likely mediated primarily by mate choice on cognitive and behavioural traits. Mate preferences are multi-dimensional, and vary across cultures and time[23]. It is likely that male-biased reduced reproductive success associated with increased s_{het} burden involves multiple cognitive and behavioural traits. The negative impact of s_{het} burden on fluid intelligence, household income and educational attainment, together with the previously documented female-biased preference for mates with good financial prospects[21] suggest that sex-biased mate preferences contribute in part to the sex-bias in reproductive success with increased s_{het} burden. However, as we are not able to assess the effect of s_{het} burden on all characteristics that are valued in a mate, especially those that are ranked most highly by both sexes (e.g. emotional stability and maturity)[21], we cannot exclude the possibility that sex biases in the impact of s_{het} burden on these traits also contribute to the sex bias in reproductive success. Nonetheless, this study represents an important validation of the relevance of Darwin's theory of sexual selection[5] to contemporary human populations.

We note that while this study demonstrates that rare heterozygous genetic variation has a bigger impact on reproductive success in males than females, heritability analyses have suggested a greater contribution of common genetic variation to variance in reproductive success in females than males[38]. These two observations are potentially complementary: the larger contribution of heterozygous rare genetic variation to male reproductive success could be lowering the proportionate contribution of common genetic variation. Previous work demonstrated that the proportion of the genome that is homozygous is also associated with decreased reproductive success through increased childlessness, although without an apparent sex-bias, and proposed that this association is largely driven by rare homozygous variation, which we did not assess here[39]. Involuntary childlessness can have serious

consequences for mental health, and further studies of the genetic contributions to involuntary childlessness are warranted.

These findings may help to explain, at least in part, why only a minority of genes under the highest selective constraint have been associated with single gene disorders that increase pre-reproductive mortality or cause infertility. While there are clearly many more single gene disorders to be discovered among these genes[40,41], we anticipate that these highly constrained genes will not be neatly divided into those that cause single gene disorders and those that impact on reproductive success without causing a clinical condition. Rather, we predict that damaging variants in many of these genes will perturb neurodevelopment resulting in a broad spectrum of cognitive and behavioural outcomes, which will increase an individual's risk of childlessness, but only in some cases result in a clinically-ascertainable condition.

When investigating sex-biased patterns of genetic associations for cognitive and behavioural traits, the potential contribution of reproductive success and mate choice ought to be considered. For example, it has been posited that the preferential transmission from mothers of inherited alleles increasing risk of neurodevelopmental disorders potentially relates to the greater 'resilience' of females to such alleles[42]. However, our findings that the impact on cognition of the damaging genetic variation studied here is similar between the sexes suggests that, other than for autism spectrum disorder[43], mate choice may be a more plausible explanation for such observations, as seen for the 22q11.2 deletion[44].

These analyses have several limitations. First, we do not have longitudinal relationship data for UK Biobank participants that might shed more light on the impact of s_{het} burden on the ability to attract a partner during peak reproductive years. Second, we have not been able to explore the impact of s_{het} burden on the full range of cognitive and behavioural traits that relate to mate preferences and influence reproductive success. We anticipate that teasing out the relative contributions of correlated cognitive and behavioural traits will be challenging. Third, UK Biobank participants are biased towards higher health, educational attainment and socioeconomic status[37], and as such the estimates of the negative effect of s_{het} burden on reproductive fitness possibly underestimate the true effects in the general population. Finally, we cannot completely account for as-yet-undiscovered male infertility genes in these analyses; nonetheless, these results – in particular those based on clinically- and self-reported health outcomes (Figure 2; Supplementary Figure 11) – suggest a minor contribution of male infertility to the relationship between s_{het} burden and childlessness.

Our study focused on individuals of European ancestry and analogous studies across different populations and cultures are needed. Males have considerably greater variance in reproductive success than females across cultures[45], including higher levels of childlessness than females[26], highlighting the potential for sexual selection acting on male reproductive success to act across populations. We also note that many of the fundamental trends relating to mate preferences and male childlessness have been shown to be cross-cultural in nature[21,25,45]. We anticipate future studies that integrate genome-wide sequencing data on large population samples from a range of ancestries to more fully characterise the impact of sexual selection on our species.

Methods

Sample Selection and Phenotype Collation

To collate phenotypes for all individuals in UK Biobank, we downloaded bulk phenotype files from the UK Biobank data showcase (https://www.ukbiobank.ac.uk/data-showcase/; data acquired 22 Jan 2020). Due to ascertainment biases with post-recruitment data (Supplementary Figure 18), we only retained data which were ascertained at time of recruitment as opposed to those ascertained via followup (i.e. instance 0 in the UK Biobank data showcase). Please see Supplementary Table 1 for detailed descriptions of all phenotypes assessed in this manuscript, including how they were processed, if applicable. Individuals with missing data for a relevant phenotype were excluded from analysis when testing that phenotype.

Following phenotype collation, we next selected for final analysis individuals of broadly European ancestry as determined by Bycroft et al.[46], which left a total of 409,617 individuals. To identify and remove related individuals, we first downloaded the relatedness file from the UK Biobank data showcase using the ukbbgene tool, which contains 107,124 relatedness pairs among UK Biobank participants[46]. Next, we sorted individuals by the total number of related pairs within this file, and removed the individual with the most related pairs and recalculated the total number of relationships for all other individuals. We repeated this process until no related pairs remained, which left a total of 342,717 individuals for downstream analysis.

Calling, Quality Control, and Annotation of Copy Number Variants from SNP Microarrays

To ascertain copy number variants from 488,377 UK Biobank participants with available genetic data[46], we utilised the PennCNV CNV-ascertainment pipeline[47]. Raw CEL files were downloaded in 107 independent batches, of which 95 batches were genotyped with the standard UK Biobank array platform and 12 batches were genotyped with the UKBiLEVE array platform. Each batch was then processed independently through the following calling pipeline: first, raw CEL files were genotyped with Affymetrix power tools (http://media.affymetrix.com/support/developer/powertools/changelog/index.html) 'genotype' with default settings. Next, using the 'generate_affy-geno_cluster.pl' and 'normalize_affy_geno_cluster.pl' scripts provided as part of PennCNV, genotyped samples within each batch were clustered and normalised, respectively. Normalised clustering output was then split into one file per individual and provided as input to 'detect_cnv.pl' to generate an initial call set of CNVs. Finally, initial CNVs were then passed to the 'clean_cnv.pl' script with "-fraction" set to 0.25 in order to merge nearby CNV calls in each individual. Following CNV calling, we excluded all individuals with ≥ 20 CNVs and absolute waviness factor > 0.3, and all variants on either the X or Y chromosome, which left 485,593 individuals and 3,101,974 raw, redundant CNVs.

To perform quality control of ascertained CNVs, we developed a novel approach which uses individuals for which CNVs have been ascertained with both array and exome-based approaches. In short, we started with the basic logistic regression concept outlined in Mace et al.[48] but instead used the intersect of array- and WES-ascertained CNVs as the dependent variable in a random forest model[49], with various per-individual and per-CNV metrics as predictors. To train this model, we utilised an additional set of 46,856 individuals collected as part of the INTERVAL study[50] genotyped on the same array as participants in UK Biobank, of which 4,465 also had matched WES data. For INTERVAL individuals, we performed array-based CNV calling identically to the method as described above and ascertained exome-based CNVs using three different algorithms with default settings: XHMM[51], CANOES[52], and CLAMMS[53]. For each INTERVAL participant for which we had both array and exome-based CNVs, we then determined a "WES overlap score" as a product of the overlap of each array-based CNV with the three WES-based callers, corrected for whether or not any overlap was possible due to probe/exon bias. Scoring results in a roughly continuous metric for each array-ascertained CNV of between zero and three, where zero represents a lack of overlap with any WES CNV call and three represents a perfect overlap with all three algorithms. For predictor covariates, we used several metrics already shown to be of high quality for CNV quality control[48,54], per-CNV metrics based on these (e.g. mean log R ratio for each probe within a CNV rather than for all probes across an entire individual), and a novel metric which uses specific probes on the array known to be biased for CNV calls on bad arrays (Supplementary Table 3; see code availability). To determine estimated sensitivity/specificity of our model we performed 10-fold cross-validation, where all array CNVs which overlapped at least two exons were split into equal test and training sets and provided, separately for deletions and duplications, as input into the randomForest implementation in R as a linear predictor with nTrees set to 500. To generate a call set of final quality controlled CNVs for downstream analyses, we then trained a final random forest using all INTERVAL individuals with matched array and WES data and generated predicted WES overlap scores for all 3,101,974 raw UK Biobank CNVs identified with PennCNV as described above. CNVs were then filtered based on a predicted sensitivity of 95% based on cross-validation, leaving a remaining 1,612,931 CNVs (1,043,717 deletions, 569,114 duplications).

CNVs passing quality control were then provided as input to a custom java pipeline which merged all CNVs, regardless of whether they were deletions or duplications, based on 75% reciprocal overlap to generate a set of 173,871 nonredundant loci. Following filtering to 342,717 unrelated individuals of broadly European ancestry for which CNV data was available, each locus was quantified for allele frequency. Loci were then assessed for overlap with a set of known pathogenic CNVs identically to Crawford, et al.[54] and annotated using Variant Effect Predictor (VEP) v97[55]. Only loci with an annotation of 'transcript_ablation' or 'feature_trunctation' and ‘coding_sequence_variant’ for deletions, and 'transcript_amplification' or 'feature_elongation' and ‘coding_sequence_variant’ for duplications were considered to be affecting a target gene. A total of 1,118,859 redundant CNVs remained for downstream analysis following all filtering and annotation (721,536 deletions, 397,323 duplications; Supplementary Figure 2).

Processing SNV/InDel Data from WES

To collate protein truncating, missense, and synonymous variants for all 49,960 individuals whole exome sequenced by UK Biobank[11], we downloaded the GRCh38-aligned population-level variant files from the UK Biobank (UK Biobank field 23160) and converted them to variant call format. All autosomal variants were then annotated with VEP v97[55], CADDv1.5[56], allele frequency from gnomAD[57], PEXT[58] and, where relevant, MPC[59] and LOFTEE[57]. PEXT and MPC scores were converted from build37 to build38 using the CrossMap tool[60]. Variants were assigned to a gene based on the primary ENSEMBL transcript with the most severe consequence. Variants were considered to be PTVs if they were annotated by VEP as having a splice acceptor/donor, stop gained, or frameshift consequence. We then retained only variants with a gnomAD or UK Biobank-specific allele frequency ≤ $1x10^{-3}$ and with a PEXT mean cerebellum score > 0.1. Missense variants were only retained if they had MPC > 2 and CADD > 25. PTVs were only retained if they were annotated by LOFTEE as high confidence, had CADD > 25, and were not located in the last exon or intron of the canonical transcript as annotated by ENSEMBL[61]. This filtering approach left a total of 2,658,431 redundant autosomal SNVs and InDels across all 34,812 unrelated individuals of broadly European ancestry included in this study (Supplementary Figure 3).

It has recently been reported that the UK Biobank exome sequencing data is missing variant calls in regions where all reads were assigned MAPQ=0 (for more details, see Jia et al.[62]). While this issue affects 703 genes with an s_{het} value assessed in this study, genes with the highest constraint scores (i.e. $s_{het} \geq 0.15$) are less likely to be affected by this problem (3.3% of genes with $s_{het} \geq 0.15$, 4.5% of genes with $s_{het} < 0.15$; Fisher's p=0.02). Secondly, this issue is consistent across all individuals with WES within the UK Biobank and thus results in a simple loss of power equivalent to having insufficient coverage to call variants across ~4% of the exome. Finally, as CNV calling was performed using genotyping arrays, and thus unaffected by issues with sequence alignment, our findings are independently robust. Information on exome capture baits and genes affected by alignment issues for producing this statement were acquired from the UK Biobank data showcase (http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=1911).

Calculating s_{het} Burden for UK Biobank Participants

To calculate an individual's s_{het} burden, assuming that fitness is multiplicative and that there is no epistasis between genes which are lost, we utilised the following formula:

$$s_{het[i,v]} = 1 - \prod_{g}(1 - s_{het[i,v,g]})$$

where $s_{het[i,v]}$ indicates individual i's s_{het} burden for variant class v and $s_{het[i,v,g]}$ indicates the s_{het} score for gene g with a qualifying annotation for variant class v in individual i. Possible values for v are PTV, missense, synonymous, deletion, or duplication. As indicated by the formula above, s_{het} values were calculated independently for each variant type. Per-gene s_{het} values were obtained from Weghorn et al.[3], under their demographic model which includes

drift and scores for 16,189 protein coding genes which we were able to definitively map to an ENSEMBL gene ID. To ensure that our primary result of the effect of s_{het} burden on childlessness is unaffected by the version of s_{het} we use to calculate our burden scores, we also utilised an earlier derivation of s_{het} from Cassa et al.[4] which does not take into account a demographic model. This did not significantly change our primary result (Supplementary Figure 22).

To explore if genes known to be associated with male infertility were responsible for our observed effect on male reproductive success, we also generated individual s_{het} scores for each variant class excluding a set of 150 autosomal genes known to be associated with male infertility (Supplementary Table 5 from Oud et al.[14]). Genes with an annotation of limited, moderate, strong, or definitive evidence were excluded from calculated s_{het} scores[15]. Similarly, and to test if a greater number of 742 genes associated with male infertility in mice were responsible for our observed effect on male reproductive success, we queried all genes from Mouse Genome Informatics[15] with a phenotype code of MP:0001925. Gene IDs were then translated to their human homologues and s_{het} burden scores excluding these genes were then generated and provided as input to logistic regression as described above.

To test if our observed relationship was robust when excluding genes with a known disease annotation, we also generated individual s_{het} scores where we removed 4,414 disease-associated genes. We considered a gene to be disease-associated based on being a confirmed or probable developmental disorder gene in the Developmental Disorders Genotype-Phenotype Database (DDG2P; https://decipher.sanger.ac.uk/info/ddg2p), in the Online Mendelian Inheritance in Man (OMIM; https://omim.org/) Morbid Map after excluding 'non disease' and 'susceptibility to multifactorial disorder' entries, or in ClinVar[63] with a pathogenic/likely pathogenic variant linked to a phenotype.

Logistic and Linear Modelling of Phenotypes

To test the association of each s_{het} burden (i.e. $s_{het[i,v]}$) per variant class with a given phenotype (e.g. those in Supplementary Table 1), we used a general linear model via the 'glm' function in R of the form:

$$phenotype \sim s_{het[i,v]} + age + age^2 + PC1..PC30$$

All models were run separately for males and females. For binary phenotypes, 'family' was set to 'binomial' and for continuous phenotypes 'family' was set to 'gaussian'. To combine the effect sizes or log odds ratios for CNVs and PTVs (e.g. for Figure 1A), we used the 'metagen' function from the 'meta' package[64] in R to perform a fixed-effects meta analysis. For logistic regression, we set parameters 'method.tau' to 'SJ' and 'sm' to 'OR'. For linear regression, we set the parameter 'sm' to "SMD". To avoid including an individual twice in our meta analysis, for samples with both CNV and PTV data available, we prioritised PTV-derived s_{het} scores.

When using raw variant counts as in Supplementary Figure 7, the s_{het} term in the above formula was changed to the total number of qualifying genes affected per individual, where qualifying genes were either those with pLI ≥ 0.9[57] or those with s_{het} ≥ 0.15[3]. Individuals with > 3 genes lost for deletions (pLI ≥ 0.9 n = 43; s_{het} ≥ 0.15 n = 16) and PTVs (pLI ≥ 0.9 n = 2; s_{het} ≥ 0.15 n = 1) were removed prior to regression analyses. To provide a negative control for our association tests, we also performed associations for several neutral phenotypes we hypothesised to not be under negative selection: fresh fruit intake, handedness, and blonde hair colour (Supplementary Table 1). None of these associations were significant after correcting for multiple testing (Supplementary Figure 23).

To test the effect of individual phenotypes on likelihood of having children (Supplementary Figure 14), we used a logistic model (with the 'family' parameter of the 'glm' function set to 'binomial' in R) of the form:

$$has.children \sim phenotype + age + age^2 + PC1..PC30$$

As with estimating the contribution of s_{het} burden to phenotypes, all analyses were run separately for both males and females. For all models involving household income, we additionally included partner at home status as a covariate, as household income was recorded per household, not per recruited individual.

Pre-computed ancestry principal components for each UK Biobank participant were taken from Bycroft et al.[46]. To alleviate concerns about a potentially arbitrary selection of the number of ancestry principal components used in our models, we repeated our primary analysis of the association between having children and s_{het} burden in males with between 10 and 30 ancestry principal components and did not observe any significant differences in our result (Supplementary Figure 24).

All odds ratios, effect sizes, standard errors, p values, and total individuals per association test reported in this manuscript can be found in Supplementary Table 4.

Collation and Testing of Participant Medical Data

To assess if a broad range of medical conditions play a role in mediating the effect of high s_{het} burden on childlessness, we queried two relevant datasets provided by the UK Biobank[10]: hospital episode statistics (HES) and combined health outcomes data (CHOD; Supplementary Table 1). Briefly, for each UK Biobank participant, HES data incorporates electronic inpatient data provided directly from NHS hospitals and CHOD aggregates HES, general practitioner records, self-reported conditions, and death records. All data sources are coded according to the International Classification of Disorders v10 (ICD-10). For the purposes of this work, we ignored all cancer codings from HES and CHOD data (ICD chapter II and codes O00-O08 of chapter XV), and instead used independent cancer registry data; the UK Biobank acquires information on cancer diagnoses from the UK cancer registry which aggregates a wide range of data sources including general practitioners, nursing homes, and hospitals and is considered the more accurate data source for UK Biobank

participant cancer diagnoses. Cancer codes were retained only when testing HES data (Figure 2, Supplementary Figures 11-13).

We utilised both HES and CHOD sources to examine a broad set of medical conditions in UK Biobank participants. Complete HES are available for all UK Biobank participants but are probably depleted of conditions that are unlikely to be seen in a hospital setting (e.g. male infertility). CHOD are also likely to be more sensitive to a wide variety of conditions as they incorporate aforementioned HES data with both general practitioner records and self-reported outcomes. Importantly, while general practitioner records are only available for 46% (n = 230,090) of UK Biobank participants, when we tested for an association between s_{het} burden and whether a participant had general practitioner records or not, we did not observe a significant association for either males (OR=1.02 [95% CI 0.75-1.37], p=0.92) or females (OR=1.26 [95% CI 0.94-1.68], p=0.12; Supplementary Figure 18). This indicated that we were unlikely to see biases due to including CHOD from individuals who were missing general practitioner records. As such, we prioritised the use of CHOD in most analyses presented in the text, figures, and supplementary information of this manuscript – complete results for all codes in both HES and CHOD data are available as Supplementary Table 2. Exceptions include when testing codings from chapters XVII to XXII which are beyond the diagnostic scope of CHOD, and cancer codes better ascertained from the UK Biobank cancer registry as noted above (Figure 2; Supplementary Figures 12, 13).

To determine the role of 19,154 diseases, disorders, and special codes collated from HES and CHOD in the relationship between s_{het} burden and childlessness (Figure 2), we used a modified version of our primary logistic model of the form:

$$has.children \sim s_{het[i,v]} + icd.code_c + age + age^2 + PC1..PC30$$

Where $icd.code$ represents a binarised presence/absence of one of 19,154 different ICD-10 diseases, disease groups, and chapters, c. Tests were only performed when a given code was represented by at least 2 individuals in both genetic (i.e. CNVs) and WES (i.e. PTVs) data. When considering individuals who have a particular code, c, we utilised the hierarchical information present within the ICD-10 coding system. For example, when we tested if inclusion of a term for individuals with non-insulin-dependent diabetes mellitus (ICD-10 code E11) has an effect on childlessness, we also considered individuals with any sub-code (i.e. E11.0-E11.9). This same principle was also used for disease groupings – when testing the more general diabetes mellitus group (ICD-10 block E10-E14) we included all individuals with any code between E10 and E14, including disease subtypes (e.g. E11.0). For each model, we retained both an odds ratio for the effect of individual s_{het} burden and presence of a given code on childlessness (Supplementary Table 2).

Evaluation of Gene Expression in Testis

To determine the expression in testis of all genes assessed in this study, we downloaded processed median transcripts per million values for all genes provided by v7 of the GTEx study

(https://storage.googleapis.com/gtex_analysis_v8/rna_seq_data/GTEx_Analysis_2017-06-05_v8_RNASeQCv1.1.9_gene_median_tpm.gct.gz). Only genes for which an s_{het} score was available[3] were retained from this file. We then determined if each gene was affected by either a private deletion or PTV in a UK Biobank individual. We then plotted $ln\,(testis\ expression)$ in two ways: (i) as a factor of being a male infertility gene or not or (ii) having or not having a qualifying variant (Supplementary Figure 10). To determine significance we used a one-sided Wilcoxon test, with the alternate hypothesis that expression in testis of male infertility genes or genes with private variants is greater than the alternative set.

Modelling the Contribution of Phenotypes to Observed Reduction in Fitness

Variant s_{het} Burden

To estimate the contribution of s_{het} to overall fitness (Figure 4B), we extracted log odds ratio estimates for the effect of s_{het} on having children from our logistic model and estimated the proportion of childless individuals at various s_{het} scores (0 to 1 at 0.1 intervals; Supplementary Figure 21). To calculate the error in our estimates (i.e. the shaded areas in Figure 4B), we used the 95% confidence intervals for the s_{het} burden log odds ratio from our original logistic regression. Please see Supplementary Note 1 for a more detailed description of how the contribution of s_{het} to overall fitness was calculated.

General Cognition

When possible, we used independent estimates from population level or external data to alleviate biases in UK Biobank phenotype ascertainment (Supplementary Figure 18)[35]. As such, data on cognitive ability and fertility are collected from Swedish population-level government administrative registers that have been linked to Swedish conscription registers[65]. To assess assignment into different branches of a universal conscription for Swedish men, the Swedish government included an extensive cognitive ability test which all men in Sweden had to take part in. Information on childbearing is based on birth records, and linkage to both men and women is nearly universal, partly due to universal government identity numbers, combined with serious paternity investigations in case of missing information of the biological father. This information was used to calculate reproductive fertility histories in 2012 for all men included in this study. We include data on all Swedish born men who participated in the military conscription test at age 18-20 who were born 1965-1967. The conscription registers are described in more detail elsewhere[32,66].

For the current study, we did not rely on the official cognitive ability scores assigned for each man following their cognitive ability test as in Kolk and Barclay[67], but instead made manual calculations to create a more finely grained measure from raw test scores based on a battery of cognitive ability tests that are available for 3 years in our conscription registers. The Swedish military created an official IQ-measure based on a 9-score stanine scale that has been used in a large number of scientific studies[67,68]. In the current study we developed a more detailed score using information on the actual test scores of men participating in the test. The conscription test consisted of four large subtests measuring different dimensions of IQ with logical, spatial, verbal, and technical subtest[66,69,70]. To get a more finely tuned IQ

measure than the official stanine measure we used the raw test scores of each of these four tests and summed the total number of correct questions for these four sub-tests. Within each stanine IQ score, we then examined the distribution of test scores and after standardising the test scores using only variation within each stanine score, calculated a new detailed IQ score. This procedure is done to anchor our new IQ measure in the official stanine IQ score. As our test scores have some missing values for men with very high and very low stanine scores, this procedure results in a slightly underdispersed distribution and our new calibrated IQ score has μ = 100 & σ = 12, as compared to the official stanine measure with μ = 100 & σ = 15.

This score allows us to calculate cognitive ability by single digit IQ scores (Supplementary Table 5); however, as we had to rely on only observations with complete test scores for all test batteries, our data has a higher share of excluded men than the official cognitive ability scores (used by Kolk and Barclay[32] and others). In addition to the ~5% of men that did not take the test (e.g. they were ineligible for military service due to handicap such as visual impairments, that they were abroad, or were conscripted at an atypical age), we additionally excluded a number of men for which scores of all test batteries were not available. Our manually computed fine-grained measure was later standardised against the official cognitive ability test score to maintain comparability and to assure our slightly smaller population is still representative of the complete cohort. Compared to most other measures of cognitive ability in the scientific literature, we argue that our population is unusually representative as little (indirect) pre-selection due to cognitive ability took place.

We first estimated the effect of overall s_{het} burden on fluid intelligence (Figure 3F) and, because fluid intelligence is normalised and IQ is normally distributed, converted this effect to a predicted change of IQ. To then estimate childlessness and fertility for low IQ values not actually observed in the general population, we fit actual observations to a sigmoidal model using the function 'nls' in R (Supplementary Figure 17; Supplementary Table 5). As our empirical distribution did not conform to a standard test distribution, we then simulated 100,000 individuals, with IQ values for each individual randomly selected from our original Swedish IQ distribution with the mean shifted by the expected reduction in IQ as explained by our s_{het} model. We then assigned each simulated individual an expected number of children and predicted probability of childlessness based on their simulated IQ value as given in Supplementary Table 5. Number of children across all 100,000 individuals was then averaged to generate an expected mean fertility for a given s_{het} score (Supplementary Figure 16). This value was then compared to the mean number of children for the unburdened population via the following formula:

$$fertility.ratio = \frac{fertility_{shet(1)}}{fertility_{shet(0)}}$$

To then calculate the proportion of reduced reproductive success explained by IQ (and by extrapolation, other traits) we then divide this fertility ratio by the overall reduction in fitness given by s_{het} as described above. Please see Supplementary Note 2 for a more detailed example of how we performed this calculation.

Mental Health Disorders

As with general cognition, we used estimates from external studies to alleviate biases in UK Biobank phenotype ascertainment (Supplementary Figure 18)[35]. In this case, as we were

unable to accurately estimate the increased risk of developing individual mental health disorders as a factor of individual s_{het} burden, we instead utilised odds ratios from Ganna et al.[9]. Only odds ratios for schizophrenia, autism spectrum disorder, and bipolar disorder were retained. As Ganna et al.[9] estimated the risk based on total count of high pLI (≥0.9)[57] genes with PTVs per individual instead of with s_{het}, we assumed that an individual carrying one such variant had an s_{het} burden of 0.162, or the mean s_{het} value of all high pLI (≥0.9) genes. We then converted this into a proportion of individuals with a given mental health disorder, *t*, at s_{het} burden, *x*, by scaling the odds ratio with the following formula:

$$log(OR_{s_{het}[x,t]}) = \frac{log(OR_{ganna}) * s_{het}[x]}{0.162}$$

To establish a baseline expectation for the prevalence of each mental health disorder at s_{het} 0 we utilised population-level data from Power et al.[19] and extrapolated for each trait at increasing s_{het} values (Supplementary Figure 25). To generate an expected mean number of children for simulated individuals with mental health disorders, we used fertility statistics generated by Power et al.[19]. As Power et al.[19] did not provide childlessness data, we were unable to generate expected childlessness as we did for other traits. Overall predicted reduced fitness attributable to mental health and all values used for performing the above analyses are provided in Supplementary Table 6.

Having a Partner at Home, Having had Sex, Educational Attainment, and Infertility

To determine the contribution of having a partner at home, ever having had sex, educational attainment, and a medical diagnosis of infertility to the relationship between s_{het} burden and childlessness, we utilised a multiple regression model incorporating various combinations of these traits (Figure 4A; Supplementary Figure 20). First, we calculated the variance in childlessness explained by a null model consisting of age, age^2, and the first 30 ancestry PCs using Nagelkerke's pseudo-r^2 as calculated using the "nagelkerke" function from the R package "rcompanion" (https://rcompanion.org/handbook/). Next, to determine the proportion of variance explained in childlessness by s_{het} alone, we calculated incremental pseudo-r^2 between this null model and a model additionally incorporating a term for deletion s_{het} burden. We then repeated this analysis, except now including an additional covariate (e.g. having a partner at home) to determine the reduction in variance explained by deletion s_{het} when correcting for the additional covariate. This reduction in variance was then converted to a percent change via the following formula:

$$\%\ reduction\ in\ variance\ explained\ by\ deletion\ s_{het} = \frac{s_{het}\ incremental\ r^2 with\ covariate}{s_{het}\ incremental\ r^2\ without\ covariate}$$

This basic analysis was then repeated for all possible combinations of having a partner at home, ever having had sex, having a university degree, and having a medical diagnosis of infertility (Figure 4A; Suppelementary Figure 20). Percent reduction in variance explained values plotted in Figure 4A and Supplementary Figure 20 are displayed for s_{het} calculated using deletions only.

Relatively complete mental health disorder data are available for all individuals via the complete health outcomes data; therefore, we also included a covariate for having a mental health disorder as a covariate in our multiple regression model (Figure 4A). As income is provided by UK Biobank on a per-household basis (Figure 3E) and the number of individuals with fluid intelligence data recorded at recruitment is significantly smaller than for other covariates (Figure 3F), we did not include these as part of our multiple regression model.

Acknowledgements

We thank Leopold Parts, Molly Przeworski and George Davey-Smith for useful discussions and comments during manuscript preparation. We thank the INTERVAL study for sharing genotyping and exome data that allowed us to refine our CNV filtering methodology. We thank Manon Oud and Joris Veltman for helpful discussions regarding male infertility in mice. We thank the reviewers for constructive comments and criticism. This work has been funded by core Wellcome funding to the Wellcome Sanger Institute (grant WT098051). This work has been conducted using the UK Biobank Resource under application numbers 14421 (to G.K.) and 44165 (to H.C.M.).

Author Contributions

E.J.G, M.D.C.N, and K.E.S. assessed the contribution of rare genetic variation to the phenotypes and vital statistics presented in this manuscript. E.J.G and G.K. performed CNV calling. E.J.G. and M.E.K.N. annotated and assessed SNV and InDel variants from provided WES data. K.B., M.K., E.J.G, and M.E.H. curated and analysed Swedish IQ data. E.J.G., K.E.S., H.C.M., and M.E.H. designed experiments, oversaw the study and wrote the manuscript.

Data Availability

CNVs, SNVs and InDels included in this study will be returned to the UK Biobank following study publication, as per UK Biobank guidelines.

Code Availability

Code used as part of this project to perform phenotype testing, CNV calling, variant quality control, and generate all main text figures, supplementary figures and supplementary tables is available on github: https://github.com/eugenegardner/UKBBFertility.

References

1. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).

2. Collins, R. L. *et al.* A structural variation reference for medical and population genetics. *Nature* **581**, 444–451 (2020).

3. Weghorn, D. *et al.* Applicability of the Mutation-Selection Balance Model to Population

Genetics of Heterozygous Protein-Truncating Variants in Humans. *Mol. Biol. Evol.* **36**, 1701–1710 (2019).

4. Cassa, C. A. *et al.* Estimating the selective effects of heterozygous protein-truncating variants from human exome data. *Nat. Genet.* **49**, 806–810 (2017).
5. Darwin, C. The descent of man, and selection in relation to sex. By Charles Darwin. (1874) doi:10.5962/bhl.title.16749.
6. Ganna, A. *et al.* Ultra-rare disruptive and damaging mutations influence educational attainment in the general population. *Nat. Neurosci.* **19**, 1563–1565 (2016).
7. Männik, K. *et al.* Copy number variations and cognitive phenotypes in unselected populations. *JAMA* **313**, 2044–2054 (2015).
8. Huguet, G. *et al.* Measuring and Estimating the Effect Sizes of Copy Number Variants on General Intelligence in Community-Based Samples. *JAMA Psychiatry* **75**, 447–457 (2018).
9. Ganna, A. *et al.* Quantifying the Impact of Rare and Ultra-rare Coding Variation across the Phenotypic Spectrum. *Am. J. Hum. Genet.* **102**, 1204–1211 (2018).
10. Sudlow, C. *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.* **12**, e1001779 (2015).
11. Van Hout, C. V. *et al.* Exome sequencing and characterization of 49,960 individuals in the UK Biobank. *Nature* **586**, 749–756 (2020).
12. Dudel, C. & Klüsener, S. Estimating men's fertility from vital registration data with missing values. *Population Studies* vol. 73 439–449 (2019).
13. Office of National Statistics. Birth Summary Tables, England and Wales 2019. (2020).
14. Oud, M. S. *et al.* A systematic review and standardized clinical validity assessment of male infertility genes. *Hum. Reprod.* **34**, 932–941 (2019).
15. Bult, C. J. *et al.* Mouse Genome Database (MGD) 2019. *Nucleic Acids Res.* **47**,

868 D801–D806 (2019).

869 16. Huang, N. *et al.* A Screen for Genomic Disorders of Infertility Identifies MAST2
870 Duplications Associated with Nonobstructive Azoospermia in Humans. *Biol. Reprod.* **93**,
871 61 (2015).

872 17. Skjaerven, R., Wilcox, A. J. & Lie, R. T. A population-based study of survival and
873 childbearing among female subjects with birth defects and the risk of recurrence in their
874 children. *N. Engl. J. Med.* **340**, 1057–1062 (1999).

875 18. Lie, R. T., Wilcox, A. J. & Skjaerven, R. Survival and reproduction among males with
876 birth defects and risk of recurrence in their children. *JAMA* **285**, 755–760 (2001).

877 19. Power, R. A. *et al.* Fecundity of patients with schizophrenia, autism, bipolar disorder,
878 depression, anorexia nervosa, or substance abuse vs their unaffected siblings. *JAMA*
879 *Psychiatry* **70**, 22–30 (2013).

880 20. Allen, M. S. The Role of Personality in Sexual and Reproductive Health. *Current*
881 *Directions in Psychological Science* vol. 28 581–586 (2019).

882 21. Buss, D. M. *et al.* International Preferences in Selecting Mates: A Study of 37 Cultures.
883 *J. Cross. Cult. Psychol.* **21**, 5–47 (1990).

884 22. Pawłowski, B. & Dunbar, R. I. Impact of market value on human mate choice decisions.
885 *Proc. Biol. Sci.* **266**, 281–285 (1999).

886 23. Buss, D. M. & Schmitt, D. P. Mate Preferences and Their Behavioral Manifestations.
887 *Annu. Rev. Psychol.* **70**, 77–110 (2019).

888 24. Fieder, M., Huber, S. & Bookstein, F. L. Socioeconomic status, marital status and
889 childlessness in men and women: an analysis of census data from six countries. *J.*
890 *Biosoc. Sci.* **43**, 619–635 (2011).

891 25. Nettle, D. & Pollet, T. V. Natural selection on male wealth in humans. *Am. Nat.* **172**,
892 658–666 (2008).

893 26. Miettinen, A., Rotkirch, A., Szalma, I., Donno, A. & Tanturri, M.-L. *Increasing*

childlessness in Europe: time trends and country differences. (Family and Societies Working Paper 33, 2015).

27. Jalovaara, M. *et al.* Education, Gender, and Cohort Fertility in the Nordic Countries. *Eur. J. Popul.* **35**, 563–586 (2019).
28. Fieder, M. & Huber, S. The effects of sex and childlessness on the association between status and reproductive output in modern society. *Evolution and Human Behavior* vol. 28 392–398 (2007).
29. Barthold, J. A., Myrskylä, M. & Jones, O. R. Childlessness drives the sex difference in the association between income and reproductive success of modern Europeans. *Evolution and Human Behavior* vol. 33 628–638 (2012).
30. Bateman, A. J. Intra-sexual selection in Drosophila. *Heredity* **2**, 349–368 (1948).
31. Trivers, R. Parental Investment and Sexual Selection. in *Sexual selection and the descent of man* (ed. Campbell, B.) (Aldine, 1972).
32. Kolk, M. & Barclay, K. Cognitive ability and fertility among Swedish men born 1951-1967: evidence from military conscription registers. *Proc. Biol. Sci.* **286**, 20190359 (2019).
33. Kendall, K. M. *et al.* Cognitive Performance Among Carriers of Pathogenic Copy Number Variants: Analysis of 152,000 UK Biobank Subjects. *Biol. Psychiatry* **82**, 103–110 (2017).
34. Davis, K. A. S. *et al.* Mental health in UK Biobank - development, implementation and results from an online questionnaire completed by 157 366 participants: a reanalysis. *BJPsych Open* **6**, e18 (2020).
35. Tyrrell, J. *et al.* Genetic predictors of participation in optional components of UK Biobank. *Cold Spring Harbor Laboratory* 2020.02.10.941328 (2020) doi:10.1101/2020.02.10.941328.
36. Stefansson, H. *et al.* CNVs conferring risk of autism or schizophrenia affect cognition in

controls. *Nature* vol. 505 361–366 (2014).

37. Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am. J. Epidemiol.* **186**, 1026–1034 (2017).
38. Watanabe, K. *et al.* A global overview of pleiotropy and genetic architecture in complex traits. *Nat. Genet.* **51**, 1339–1348 (2019).
39. Clark, D. W. *et al.* Associations of autozygosity with a broad range of human phenotypes. *Nat. Commun.* **10**, 4957 (2019).
40. Stanley, K. E. *et al.* Causal Genetic Variants in Stillbirth. *N. Engl. J. Med.* (2020) doi:10.1056/NEJMoa1908753.
41. Kaplanis, J. *et al.* Evidence for 28 genetic disorders discovered by combining healthcare and research data. *Nature* **586**, 757–762 (2020).
42. Girirajan, S. *et al.* Phenotypic heterogeneity of genomic disorders and rare copy-number variants. *N. Engl. J. Med.* **367**, 1321–1331 (2012).
43. De Rubeis, S. *et al.* Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* **515**, 209–215 (2014).
44. Costain, G., Chow, E. W. C., Silversides, C. K. & Bassett, A. S. Sex differences in reproductive fitness contribute to preferential maternal transmission of 22q11.2 deletions. *J. Med. Genet.* **48**, 819–824 (2011).
45. Betzig, L. Means, variances, and ranges in reproductive success: comparative evidence. *Evolution and Human Behavior* vol. 33 309–317 (2012).
46. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
47. Wang, K. *et al.* PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.* **17**, 1665–1674 (2007).

946 48. Macé, A. *et al.* New quality measure for SNP array based CNV detection. *Bioinformatics*
947 **32**, 3298–3305 (2016).
948 49. Liaw, A. & Wiener, M. Classification and Regression by Randomforest. *R news* **2**, 285
949 (2002).
950 50. Di Angelantonio, E. *et al.* Efficiency and safety of varying the frequency of whole blood
951 donation (INTERVAL): a randomised trial of 45 000 donors. *Lancet* **390**, 2360–2371
952 (2017).
953 51. Fromer, M. *et al.* Discovery and statistical genotyping of copy-number variation from
954 whole-exome sequencing depth. *Am. J. Hum. Genet.* **91**, 597–607 (2012).
955 52. Backenroth, D. *et al.* CANOES: detecting rare copy number variants from whole exome
956 sequencing data. *Nucleic Acids Res.* **42**, e97 (2014).
957 53. Packer, J. S. *et al.* CLAMMS: a scalable algorithm for calling common and rare copy
958 number variants from exome sequencing data. *Bioinformatics* **32**, 133–135 (2016).
959 54. Crawford, K. *et al.* Medical consequences of pathogenic CNVs in adults: analysis of the
960 UK Biobank. *J. Med. Genet.* **56**, 131–138 (2019).
961 55. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biology* vol. 17
962 (2016).
963 56. Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J. & Kircher, M. CADD: predicting
964 the deleteriousness of variants throughout the human genome. *Nucleic Acids Research*
965 vol. 47 D886–D894 (2019).
966 57. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in
967 141,456 humans. *Nature* **581**, 434–443 (2020).
968 58. Cummings, B. B. *et al.* Transcript expression-aware annotation improves rare variant
969 interpretation. *Nature* **581**, 452–458 (2020).
970 59. Samocha, K. E. *et al.* Regional missense constraint improves variant deleteriousness
971 prediction. (2017) doi:10.1101/148353.

60. Zhao, H. *et al.* CrossMap: a versatile tool for coordinate conversion between genome assemblies. *Bioinformatics* **30**, 1006–1007 (2014).

61. Kersey, P. J. *et al.* Ensembl Genomes 2016: more genomes, more complexity. *Nucleic Acids Res.* **44**, D574–80 (2016).

62. Jia, T., Munson, B., Lango Allen, H., Ideker, T. & Majithia, A. R. Thousands of missing variants in the UK Biobank are recoverable by genome realignment. *Ann. Hum. Genet.* **84**, 214–220 (2020).

63. Landrum, M. J. *et al.* ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* **46**, D1062–D1067 (2018).

64. Balduzzi, S., Rücker, G. & Schwarzer, G. How to perform a meta-analysis with R: a practical tutorial. *Evid. Based. Ment. Health* **22**, 153–160 (2019).

65. Sweden, S. *Multi-generation register 2016: A description of contents and quality*. (Statistics Sweden, Population and Welfare Department, 2017).

66. Carlstedt, B. Cognitive abilities - aspects of structure, process and measurement. (University of Gothenburg, 2000).

67. Kolk, M. & Barclay, K. Cognitive ability and fertility among Swedish men born 1951-1967: evidence from military conscription registers. *Proc. Biol. Sci.* **286**, 20190359 (2019).

68. Hällsten, M. Inequality across three and four generations in Egalitarian Sweden: 1st and 2nd cousin correlations in socio-economic outcomes. *Research in Social Stratification and Mobility* vol. 35 19–33 (2014).

69. Mårdberg, B. & Carlstedt, B. Swedish Enlistment Battery (SEB): Construct Validity and Latent Variable Estimation of Cognitive Abilities by the CAT‐SEB. *International Journal of Selection and Assessment* vol. 6 107–114 (1998).

70. Rönnlund, M., Carlstedt, B., Blomstedt, Y., Nilsson, L.-G. & Weinehall, L. Secular trends in cognitive test performance: Swedish conscript data 1970–1993. *Intelligence* vol. 41

19–24 (2013).