Reduced reproductive success is associated with selective constraint on human genes

Genome-wide sequencing of human populations has revealed substantial variation among genes in the intensity of purifying selection acting on damaging genetic variants. Although genes under the strongest selective constraint are highly enriched for associations with 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 are associated with markedly reduced reproductive success, primarily owing to increased childlessness, with a stronger effect in males than in females. We present evidence that increased childlessness is probably mediated by genetically associated cognitive and behavioural traits, which may mean that male carriers are less likely to find reproductive partners. This reduction in reproductive success may account for 20% of purifying selection against heterozygous variants that ablate protein-coding genes. Although this genetic association may only account for a very minor fraction of the overall likelihood of being childless (less than 1%), especially when compared to more influential sociodemographic factors, it may influence how genes evolve over time.

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. The strength of selection against heterozygous PTVs has been estimated using the selection coefficient, $s_{het}$, which exhibits a continuous spectrum across human genes, although most attention has been focused on a subset of around 3,000 genes with the highest probability of loss-of-function intolerance ($pLI$).

The selection pressures acting on these most selectively constrained genes have not been fully characterized, but could include natural selection against variants increasing pre-reproductive mortality or decreasing fertility and sexual selection acting on mate choice or intra-sexual competition. Gene deletions and PTVs in these genes have been shown to be associated with lower educational attainment and general intelligence, as well as increased risk of intellectual disability and some psychiatric disorders. Moreover, these constrained genes are strongly enriched for associations with dominant early-onset Mendelian diseases (including neurodevelopmental disorders), many of which are associated with increased pre-reproductive mortality, indicating that natural selection probably makes a substantive contribution to this selective constraint.

However, the majority (65%) of highly constrained genes ($pLI \geq 0.9$) 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.

Genetic association testing

To explore the nature of selection acting on these genes we identified PTVs and gene deletions in UK Biobank data on largely post-reproductive individuals (median age at recruitment: 58 years, range: 39–73 years, birth years: 1934–1970; Supplementary Fig. 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 genetic ancestry (Supplementary Fig. 2), and identified PTVs from exome sequencing among a subset of 139,477 participants (Supplementary Fig. 3). 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 affected by these variants (under the assumption that fitness is multiplicative; 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 $\geq 0.15$ (Kolmogorov–Smirnov $P = 1.00$; Fig. 1c), and for participants with both genic deletion and PTV data available, the analogous proportions were 4.60% and 4.59%, respectively (Kolmogorov–Smirnov $P = 0.52$; Fig. 1d).

We assessed the relationship between $s_{het}$ burden and number of children using a linear regression correcting for age, genetic ancestry and birth cohort (Fig. 1a, Supplementary Table 1; Methods). We observed that an $s_{het}$ burden of 1 (the highest possible burden) is associated with...
a decrease in the overall total number of children for males (0.26 fewer children (95% confidence interval −0.02 to 0.11), \( P = 0.20 \)) when combining results from deletion and PTV-based analyses.

To determine whether the observed association of \( s_{het} \) burden with reproductive success was primarily linked to an increased likelihood of remaining childless, we performed two analyses. First, we evaluated childlessness using logistic regression and again observed a marked sex difference in participants’ probability of having any children given their \( s_{het} \) burden, for both PTVs and genic deletions (Fig. 1b). Combining the analyses of genic deletions and PTVs, we found that an \( s_{het} \) burden of 1 is associated with a lower probability of males having any children (odds ratio = 0.32 (95% confidence interval 0.24–0.41), \( P = 4.9 \times 10^{-15} \)) much more than females (odds ratio = 0.64 (95% confidence interval 0.49–0.84), \( P = 1.2 \times 10^{-11} \)). We also observed that private duplications and probably damaging private missense variants exhibit similar but weaker associations with childlessness (Extended Data Fig. 1). Second, if we remove childless individuals from the analysis, \( s_{het} \) burden ceases to have a significant correlation with the number of offspring (Extended Data Fig. 2). This is supported by the observation that when we stratify individuals with children by the number of children, the \( s_{het} \) distribution does not vary and confirms that the observed association with reproductive success relates primarily to increased childlessness (Supplementary Fig. 4). This observation is consistent with studies that have associated demographic factors with reproductive success, which have also observed that associations are weakened or disappear when childless individuals are removed from the analysis (for example, in Barthold et al.3).

We also considered whether ascertainment biases or differences in fertility between the UK Biobank sample and the UK population as a whole could affect these results. As UK Biobank participants included in these analyses are enriched for females (54%), the observed sex difference is not owing to having greater statistical power to detect an effect on reproductive success in males. Similarly, fertility rates between the UK Biobank population and the UK population as a whole are highly similar; the 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 ± 0.07 (mean ± s.e.m). for males13 and 1.76 ± 0.05 for females14, which is very similar to that observed among UK Biobank participants (average number of children for males = 1.77, for females = 1.80).

We observed a consistent sex difference in the association of \( s_{het} \) burden with childlessness when performing this analysis in different ways, including: (1) limiting analyses to carriers of private deleterious or PTVs in the genes under most selective constraint (following thresholds set by the authors of the indicated publications: pLI ≥ 0.9 (ref. 1) or \( s_{het} \) ≥ 0.15 (ref. 2) (Supplementary Fig. 5); (2) extending analysis to more frequent, but still rare genic deletions and PTVs (Supplementary Fig. 6); (3) excluding genes known to be associated with a disease (male odds ratio = 0.33 (95% confidence interval 0.24–0.46), \( P = 4.1 \times 10^{-11} \); female odds ratio = 0.68 (95% confidence interval 0.49–0.94), nominal \( P = 0.02 \); Methods); and (4) restricting analysis to individuals in specific birth cohorts (Extended Data Fig. 3).

**Evaluating different hypotheses**

We evaluated three hypotheses that could account for the association with increased childlessness: (1) impaired fertility (for example, inability to produce viable gametes), (2) adverse health conditions, and (3) cognitive and behavioural factors (which may be associated with decreased chances of finding a reproductive partner or increased voluntary childlessness). We observed that \( s_{het} \) burden is not significantly associated (after correcting for multiple testing) with an increased risk of male (odds ratio = 6.37 (95% confidence interval 1.07–37.87), nominal \( P = 0.04 \)) or female infertility (odds ratio = 0.83 (95% confidence interval 0.33–2.09), \( P = 0.70 \)) as defined on the basis of combined health outcomes data for all UK Biobank participants (combined hospital episode statistics, primary care records, self-reported conditions, and death records).
Three additional lines of evidence suggest that impaired fertility is not the predominant cause of the sex-differential association between \( s_{\text{het}} \) burden and childlessness. First, when introducing male infertility status as a covariate in the association testing, we observed minimal reduction in the strength of the association between \( s_{\text{het}} \) burden and male childlessness (odds ratio = 0.32 (95% confidence interval 0.24–0.42), \( P = 6.2 \times 10^{-12} \); Supplementary Table 2). Second, we observed minimal change in the association between \( s_{\text{het}} \) burden and male reproductive success after removing all 150 autosomal genes for which at least limited evidence exists of an association with male infertility (odds ratio = 0.32 (95% confidence interval 0.24–0.42), \( P = 1.1 \times 10^{-15} \)) or all 742 genes associated with male infertility in mice (odds ratio = 0.33 (95% confidence interval 0.25–0.43), \( P = 3.1 \times 10^{-15} \)). Finally, genes under the highest selective constraint (\( s_{\text{het}} \geq 0.15 \)) do not appear to be more highly expressed in testis, unlike the genes currently known to be associated with male infertility (Supplementary Fig. 7). 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.

Previous studies have shown that physical birth defects are associated with reduced reproductive success\(^{16,19}\). We assessed whether any adverse health conditions contributed to the association between \( s_{\text{het}} \) burden and childlessness. We independently tested 19,154 World Health Organization International Classification of Diseases (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 \( s_{\text{het}} \) burden with childlessness (Extended Data Figs. 4, 5, Supplementary Figs. 8, 9, Supplementary Table 2; Methods). We found that although many ICD-10 codes are associated with having children—in particular, associations with male-specific codes for elective sterilization and female-specific codes associated with pregnancy and childbirth (Extended Data Fig. 5)—correcting for any ICD-10 code had minimal impact on the strength of association between \( s_{\text{het}} \) burden and childlessness (Extended Data Figs. 4, Supplementary Figs. 8, 9). The ICD-10 code with the largest (but still modest) effect on the association of \( s_{\text{het}} \) burden and male childlessness was observed for ‘developmental disorders of scholastic skills’, although this only reduced the odds ratio from 0.317 to 0.325 (Extended Data Fig. 4; 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. We noted that one of these also had a modest effect on the association of \( s_{\text{het}} \) burden and male childlessness when included as a covariate in association testing (Extended Data Fig. 4, Supplementary Table 2). This code relates to social environment problems (code Z260), with the association being driven primarily by the status of living alone (subcode Z60.2). This code is also positively associated with \( s_{\text{het}} \) burden.

There is existing evidence, primarily based on examinations of western populations, that behavioural and cognitive traits are associated with reproductive success in a sex-differential manner. First, the reduced reproductive success associated with a range of psychiatric disorders is much more pronounced in males\(^{20}\). 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\(^{21,22}\). Third, although the most highly ranked characteristics in a reproductive partner are highly discordant between the sexes\(^{23}\), some mate preferences differ between the sexes, with males placing greater value on physical attractiveness and females valuing cues relating to earning potential\(^{24,25}\). Finally, low socioeconomic status and low educational attainment have been more strongly linked to increased childlessness in males than females across populations\(^{26,27}\). This has been hypothesized to be owing to males of lower socioeconomic status finding it more difficult to find a partner\(^{28,29}\).

Highly constrained genes exhibit higher expression levels and broader expression across tissues than less constrained genes. Ubiquitous expression potentially impedes insight into the underlying biology of selection. We generated \( s_{\text{het}} \) burden scores excluding genes expressed in 53 tissues characterized by GTEx\(^{30}\) and then used these scores to test the association with childlessness (Supplementary Fig. 10). As expected, owing to the large number of genes excluded for each tissue, we observed a substantial decrease in statistical significance for all tissues in males and all but one tissue in females. However, overlapping confidence intervals for effect size estimates make identification of the tissues most relevant for selection challenging. Nevertheless, observed that \( s_{\text{het}} \) was more strongly positively correlated with gene expression levels in all 13 brain regions than all 40 non-brain regions (including testis)\(^{31}\). When \( s_{\text{het}} \) values are weighted by the number of PTVs observed in UK Biobank (thus incorporating consideration of each gene’s contribution to \( s_{\text{het}} \) burden), expression in brain tissues remains more strongly correlated than any other tissue (Supplementary Fig. 11). These analyses highlight the potential greater relevance of the brain in the biological processes underlying selection against PTVs in highly constrained genes.

Some of the observations about behavioural and cognitive associations with sex-differential 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 reproductive partners. One theory posits that this is owing to differential parental investment\(^{32}\); the sex that invests more in offspring (typically female) is more discriminating in its choice of reproductive partners, especially with regard to those partners’ potential to invest in offspring. Alternative theories have been proposed regarding the causes of sexual selection, including those that focus on disparities in gamete size (the Darwin–Bateman paradigm\(^{33,34}\)).

One prediction of the hypothesis that differential mate choice underpins the observation of a stronger male association of \( s_{\text{het}} \) burden with increased childlessness is that males with a high \( s_{\text{het}} \) burden should find it harder to find reproductive partners than females. We observed that UK Biobank participants with high \( s_{\text{het}} \) burden were significantly less likely to have reported currently living with a partner—consistent with the findings from ICD-10 codes—and that, similarly to that of reproductive success, this association was significantly stronger in males than in females (Fig. 2a). UK Biobank males currently living with a partner are also much more likely to have children (odds ratio = 5.80 (95% confidence interval 5.64–5.97), \( P < 1 \times 10^{-15} \); Extended Data Fig. 6a). 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 \( s_{\text{het}} \) burden was significantly associated with a lower probability of ever having had sex for both male (odds ratio = 0.06 (95% confidence interval 0.03–0.14), \( P = 2.5 \times 10^{-10} \)) and female (odds ratio = 0.11 (95% confidence interval 0.05–0.27), \( P = 1.2 \times 10^{-14} \)) UK Biobank participants, without significant sex difference (Fig. 2b). Additionally, whereas same-sex sexual behaviour is strongly associated with increased childlessness in UK Biobank (male odds ratio = 0.14 (95% confidence interval 0.13–0.15), \( P < 1 \times 10^{-10} \); female odds ratio = 0.27 (95% confidence interval 0.25–0.29), \( P < 1 \times 10^{-10} \); Extended Data Fig. 6h), we observed no significant association after multiple-testing correction of \( s_{\text{het}} \) burden with the likelihood of having engaged in same-sex sexual behaviour among males (odds ratio = 1.98 (95% confidence interval 1.10–3.55), nominal \( P = 0.02 \); Supplementary Fig. 12) nor did we observe any change in the association between \( s_{\text{het}} \) burden and childlessness when excluding individuals who engaged in same-sex sexual behaviour from our primary model (odds ratio = 0.33 (95% confidence interval 0.25–0.43), \( P = 4.5 \times 10^{-12} \)).

Although these findings are consistent with a role for the ability to find a partner in the association between \( s_{\text{het}} \) burden and reproductive success, this relationship is not necessarily caused by sex differences in mate preferences; it could also result from sex differences in the severity of cognitive and behavioural traits that are associated with \( s_{\text{het}} \) burden, coupled to mate choice preferences that are not sex-differential. These mechanisms are not mutually exclusive; both could be contributing to an overall sex-differential reduction in reproductive success, albeit on
different traits. We explored in UK Biobank whether the association 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 summarized above. First, we investigated the association of \( s_{het} \) burden with cognition as measured by fluid intelligence in 106,299 (49,631 males, 56,668 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 (Fig. 2f). Increasing \( s_{het} \) burden is also associated with lower educational attainment (Fig. 2c), lower household income (Fig. 2e) and greater socioeconomic deprivation (Extended Data Fig. 7), again with similar effect sizes in males and females.

To evaluate the potential link between lower cognitive performance and male reproductive success in a more quantitative manner in a less biased population sample, we extended previous work relating the results of IQ tests taken by 95% of Swedish males during military conscription to their completed family size. We estimated that lower scores in IQ tests could account for 11% (95% confidence interval 8%–12%) of the reduced male reproductive success associated with high \( s_{het} \) burden (Supplementary Fig. 13; Methods). We also note that in the Swedish data, the association between reduced reproductive success and lower IQ scores was most pronounced in males with IQ below 70 (Supplementary Fig. 14); such individuals are probably proportionately fewer in the 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 that were invited by e-mail (\( n = 157,366 \)). We observed that a high \( s_{het} \) burden was very strongly associated with having an e-mail address (male odds ratio = 0.30 (95% confidence interval 0.24–0.38), female odds ratio = 0.48 (95% confidence interval 0.42–0.55); Supplementary Fig. 15), which probably explains why individuals with a high \( s_{het} \) burden were much less likely to complete the questionnaire (male odds ratio = 0.43 (95% confidence interval 0.33–0.56), female odds ratio = 0.47 (95% confidence interval 0.37–0.57); Supplementary Fig. 15). We therefore focused analyses of mental health disorders on the complete health outcomes data available for all participants. These data corroborate a previous finding that high \( s_{het} \) burden is associated with increased risk of psychiatric disorders previously linked with reduced reproductive success (intelectual disability, schizophrenia, autism, attention deficit hyperactivity disorder and bipolar disorder; Fig. 2d, Supplementary Fig. 16). In accordance with prior literature on the association of psychiatric diagnoses with reproductive success and among UK Biobank participants, psychiatric disorders are associated with decreased probability of having any children in both males (odds ratio = 0.36 (95% confidence interval 0.32–0.40), \( P = 8.9 \times 10^{-39} \)) and females (odds ratio = 0.58 (95% confidence interval 0.52–0.66), \( P = 2.1 \times 10^{-39} \)), albeit with substantially stronger association in males (Extended Data Fig. 6d). This finding is in accordance 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. Removal of carriers of well-characterized copy number variants associated with neurodevelopmental disorders (\( n = 12,608 \) individuals; Methods), which includes variants associated with schizophrenia, does not significantly alter the association of \( s_{het} \) burden with reduced male reproductive success (odds ratio = 0.31 (95% confidence interval 0.23–0.40), \( P = 6.3 \times 10^{-32} \)) and female reproductive success (odds ratio = 0.45 (95% confidence interval 0.39–0.52), \( P = 1.1 \times 10^{-5} \)).

We subsequently tested the association of \( s_{het} \) burden with childlessness, excluding individuals with any evidence of a mental health disorder linked 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 odds ratio = 0.33 (95% confidence interval 0.25–0.43), \( P = 1.6 \times 10^{-24} \); female odds ratio = 0.63 (95% confidence interval 0.58–0.69), \( P = 6.7 \times 10^{-25} \)), suggesting that the association with childlessness is not predominantly driven by this subset of mental health disorders. We investigated 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 associated with PTVs in highly constrained genes and the reduced reproductive success associated with those disorders, we estimated that these mental health disorders could account for 16% (95% confidence interval 8%–35%) of the reduced male reproductive success associated with high \( s_{het} \) burden (Methods). Thus, in UK Biobank, both lower...
performance in tests of fluid intelligence and increased risk of psychiatric disorders probably account for only modest proportions of increased male childlessness associated with $s_{het}$ 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 multiple correlated 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 with childlessness in males (68%) but a smaller proportion in females (6%), as assessed by the difference in incremental Nagelkerke’s $R^2$ of PTV $s_{het}$ burden between models (Extended Data Fig. 5a, Supplementary Fig. 17). By far the largest contribution comes from living with a partner and having had sex (Extended Data Fig. 5a). Notably, this analysis also showed that, despite the statistical significance of the association between $s_{het}$ burden and childlessness, $s_{het}$ burden accounts for less than 1% of the likelihood of being childless for both sexes.

Contribution to selective constraint

Overall, we estimate that reduced reproductive success due to $s_{het}$ burden potentially accounts for 20% (95% confidence interval 13–28%) (Extended Data Figs. 8b, 9) of the total reduction in fitness expected due to purifying selection $r$ against PTVs as predicted by $s_{het}$ (Methods), with this reduction in fitness being much stronger in males. 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 on the basis of data from aggregated research cohorts, and may thus be biased upwards due to individuals with high $s_{het}$ burden probably being under-represented; participation in research has previously been shown to be biased with respect to gender, socioeconomic status and genetic variation. Therefore, PTVs within highly constrained genes may be present 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 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 genetic purifying selection is likely to change over time. Medical advances have altered the landscape of infertility and pre-reproductive mortality substantially. Moreover, childlessness is highly dynamic over time. Demographic data demonstrate that population-wide childlessness can double in just two decades, a trend that is readily apparent in UK Biobank (Supplementary Fig. 1). We cannot discount that sex-differential sociodemographic factors, in addition to sex differences in trait severity and mate choices, could also be contributing to 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 females. Indeed, when we correct the association between $s_{het}$ burden and childlessness for having a university degree, we see an increased association of $s_{het}$ burden with childlessness in females (odds ratio $= 0.56$ (95% confidence interval 0.42–0.73), $P = 2.8 	imes 10^{-5}$; Extended Data Fig. 5a, Supplementary Fig. 17); but the association with male childlessness remains considerably stronger in females.

Discussion

We find that in individuals of European ancestry within the UK Biobank, reduced reproductive success—especially in males—is associated with purifying selection acting on human genes, and that this may be mediated primarily by mate choice on the basis of preferences relating to cognitive and behavioural traits. The mechanisms underlying this association have not been fully resolved here. Mate preferences are multi-dimensional and vary across cultures and time, and studies of mate preferences are often biased to Western contexts. It is likely that reduced reproductive success associated with increased $s_{het}$ burden involves multiple cognitive and behavioural traits, not all of which will have been studied here. The negative association of $s_{het}$ burden with measures of fluid intelligence, household income and educational attainment, together with the previously documented stronger female preference for reproductive partners with good financial prospects, suggests that sex differences in mate preferences might contribute in part to the sex difference in reproductive success with increased $s_{het}$ burden. However, as we are not able to assess the association of $s_{het}$ burden with all characteristics that are valued in a reproductive partner, especially those that are ranked most highly by both sexes (for example, emotional stability and maturity), we cannot exclude the possibility that sex differences in the association of $s_{het}$ burden with these traits also contribute to the sex difference in reproductive success. Nonetheless, this study is consistent with the relevance of Darwin’s theory of sexual selection to contemporary human populations. Further discussion of the limitations and implications of these findings is presented in Supplementary Note 1, “Frequently asked questions”.

We note that whereas this study demonstrates that rare heterozygous genetic variation has a much stronger association with reproductive success in males than females, genome-wide association studies (predominantly in individuals with European ancestry) have suggested a greater contribution of common genetic variation to variance in reproductive success in females than males, although the genetic correlation in total fertility between men and women is highly correlated. As narrow-sense heritability of childlessness does not differ between sexes, 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 studies have demonstrated that the proportion of the genome that is homozygous is also associated with decreased reproductive success as measured through both increased childlessness and number of children ever born, although without an apparent sex difference in the latter, and proposed that these associations are largely driven by rare homozygous variation, which we did not assess here. Our study is focused on understanding why some genes might be under much stronger selective constraint than others, and not on a comprehensive analysis of the genetic basis of reproductive success. These findings may help to explain why only a minority of highly constrained genes have been associated with genetic disorders that increase pre-reproductive mortality or cause infertility. Although there are probably more genetic disorders to be discovered among these genes, we anticipate that these highly constrained genes will not neatly divide into those that cause genetic disorders and those that reduce reproductive success. Rather, we predict that damaging variants in many of these genes will perturb neurodevelopment, resulting in a spectrum of cognitive and behavioural outcomes, which may increase an individual’s risk of childlessness, but only result in a clinically ascertainable condition in some individuals.

Sex-differential reproductive success potentially contributes to some other sex-differential patterns of genetic associations for cognitive and behavioural traits. For example, preferential transmission from mothers of alleles increasing risk of neurodevelopmental disorders has been related to the greater ‘resilience’ of females to such alleles. However, our findings that the association of cognitive and behavioural traits with the damaging genetic variation studied here is similar between the sexes suggests that sex-differential reproductive success may be a more plausible explanation for some of these observations, as seen for the 22q11.2 deletion, although it is unlikely to explain sex-differential genetic associations in autism spectrum disorder.
Childlessness can be voluntary or involuntary. A survey of childless individuals (at age 42) in the 1970 British Birth Cohort found that 28% of men and 31% of women reported not wanting to have children49. Involuntary childlessness can have serious consequences for mental health, and further studies of the factors influencing involuntary childlessness are warranted (further discussed in Supplementary Note 1).

These analyses have several limitations. First, longitudinal relationship data for UK Biobank participants that might clarify the potential association of \( s_{het} \) burden with the ability to initiate and/or sustain a reproductive partnership during peak reproductive years is not available. Second, we have not investigated the association of \( s_{het} \) burden with the full range of cognitive and behavioural traits relating to mate preferences and 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 status47, and estimates of the negative association of \( s_{het} \) burden with reproductive fitness in UK Biobank possibly underestimate the true effects in the general population. Finally, we cannot account for yet-undiscovered male infertility genes in these analyses; nonetheless, our results based on clinically and self-reported infertility (Extended Data Figs. 4, 5) suggest a minor contribution of male infertility to the sex-differential 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 cultures46, including higher levels of childlessness than females27, highlighting the potential for sexual selection acting on male reproductive success to act across populations. We also note that some trends relating to mate preferences and male childlessness have some evidence of being cross-cultural in nature22,26,50. We anticipate future studies that integrate genome-wide sequencing data on large population samples from a range of ancestries to more fully characterize the nature of selection pressures acting on damaging genetic variation in our species.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi-org.ezp.lib.cam.ac.uk/10.1038/s41586-022-04549-9.

1. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 536, 285–291 (2016).
2. Cassa, C. A. et al. Estimating the selective effects of heterozygous protein-truncating variants from human exome data. Nat. Genet. 49, 806–810 (2017).
3. Collins, R. L. et al. A structural variation reference for medical and population genetics. Nature 581, 444–451 (2020).
4. 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).
5. Darvin, C. The Descent of Man, and Selection in Relation to Sex (A. L. Burt, 1874); https://doi-org.ezp.lib.cam.ac.uk/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 individuals (at age 42) in the 1970 British Birth Cohort found that 28% of men and 31% of women reported not wanting to have children. Nature 536, 285–291 (2016).
8. Tajima, F. Estimation of mutation parameters from segregating sites. Mol. Biol. Evol. 26, 1530–1538 (2009).
9. Ganna, A. et al. Quantifying the impact of rare and ultra-rare coding variation across the human genome. Nat. Genet. 51, 284–286 (2019).
10. Ganna, A. et al. Ultra-rare disruptive and damaging mutations influence educational attainment in the general population. Nature 581, 444–451 (2020).
11. Szostakowski, J. D. et al. Advancing human genetics research and drug discovery through exome sequencing of the UK Biobank. Nat. Genet. 53, 942–948 (2021).
12. Barthold, J. A., Myrskylä, M. & Jones, O. P. Childlessness drives the sex difference in the association between income and reproductive success of modern Europeans. Evol. Hum. Behav. 33, 628–638 (2012).
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://biobank.ndph.ox.ac.uk/showcase/; data acquired 22 Jan 2020). Owing to ascertainment biases with post-recruitment data (Supplementary Fig. 15), we only retained data which were ascertained at time of recruitment as opposed to those ascertained via followup (that is, instance 0 in the UK Biobank data showcase). See Supplementary Table 1 for detailed descriptions of all phenotypes assessed in this manuscript, which includes how each was 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.51, 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 ukbgenet tool, which contains 107,124 relatedness pairs among UK Biobank participants51. 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, leaving 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 (CNVs) from 488,377 UK Biobank participants with available genetic data51, we used the PennCNV CNV-ascertainment pipeline52. 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-genome_cluster.pl’ and ‘normalize_affy-genome_cluster.pl’ scripts provided as part of PennCNV, genotyped samples within each batch were clustered and normalized, respectively. Normalized 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 Y or X 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.53 but instead used the intersection of array- and whole-exome sequencing (WES)-ascertained CNVs as the dependent variable in a random forest model54, with various per-individual and per-CNV metrics as predictors. To train this model, we used an additional set of 46,856 individuals collected as part of the INTERVAL study39 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: XHMM56, CANOES57, and CLAMMS58. 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 control59-61, 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 tenfold 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 1,612,831 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 328,899 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.62 and annotated using Variant Effect Predictor (VEP) v9761. Only loci with an annotation of ‘transcript_ablation’ or ‘feature_truncation’ and ‘coding_sequence_variant’ for deletions, 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 Fig. 2).

Processing SNV and indel data from WES
To collate protein truncating, missense, and synonymous variants for all 200,629 individuals whole-exome sequenced by UK Biobank51, we downloaded the GRCh38-aligned population-level variant call format files from the UK Biobank (UK Biobank field 23156). All autosomal variants were then annotated with VEP v10260, CADDV1.61, allele frequency from gnomAD62, and, where relevant, MPC63 and LOFTEE62. MPC scores were converted from build37 to build38 using the CrossMap tool64. 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 change. We then retained only variants with a gnomAD-specific allele frequency < 1 × 10−4. Missense variants were only retained if they had MHC > 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 ENSEMBL62. We next applied genotype-level filtering where a given genotype was set to null (that is, ./.) if the genotype had a depth < 7, genotype quality < 20, or a binomial test P value for alternate versus reference reads for only heterozygous genotypes ≤ 0.001. If more than 50% of genotypes were missing for a given variant, that variant was filtered. These filtering approaches left a total of 11,455,900 redundant autosomal SNVs and indels across all 139,480 unrelated individuals of broadly European ancestry included in this study (Supplementary Fig. 3).
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 used the following formula:

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

where $s_{het}(i,v,g)$ 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 $i$. As indicated by the formula above, $s_{het}$ values were calculated independently for each variant type, with possible values for $v$ of TTV, missense, synonymous, deletion or duplication. Per-gene $s_{het}$ values were obtained from Weghorn et al. 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 association of $s_{het}$ burden with childlessness is unaffected by the version of $s_{het}$ we use to calculate our burden scores, we also used an earlier derivation of $s_{het}$ from Cassa et al. which does not take into account a demographic model. Use of $s_{het}$ scores from Cassa et al. did not significantly change our primary result (Supplementary Fig. 18).

To explore if genes known to be associated with male infertility were responsible for our observed association with 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 S from Oud et al.). Genes with an annotation of limited, moderate, strong, or definitive evidence were excluded from calculated $s_{het}$ scores. Similarly, and to test if a greater number of 742 genes associated with male infertility in mice were responsible for our observed association with male reproductive success, we queried all genes from Mouse Genome Informatics 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://www.decypher genomics.org/ddd/ddgenes), 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 with a pathogenic or likely pathogenic variant linked to a phenotype.

Logistic and linear modelling of phenotypes

To test the association of each $s_{het}$ burden (that is, $s_{het}(i)$) per variant class with a given phenotype (for example, those in Supplementary Table 1), we used a general linear model via the ‘glm’ function in R of the form

$$\text{phenotype} \sim s_{het}(i,v) + \text{age} + \text{age}^2 + \text{birth . cohort} + \text{wes . cohort} + \text{PC1} . \text{PC40} + \text{rare . PC1} . \text{PC100}$$

As with estimating the contribution of $s_{het}$ burden to phenotypes, all analyses were run separately for both males and females and with a WES batch covariate as necessary. 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.

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 association of high s\textsubscript{het} burden with childlessness, we queried two relevant datasets provided by the UK Biobank\textsuperscript{10}: hospital episode statistics (HES) and combined health outcomes data (CHOD; Supplementary Table I). In brief, 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 000-008 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. For ease of display, tests involving cancer codes are shown with HES data (for example, in Extended Data Figs. 4, 5, Supplementary Figs. 8, 9).

We used 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 (for example, 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\textsubscript{het} burden and whether a participant had general practitioner records or not, we did not observe a significant association for either males (odds ratio = 1.03 (95% confidence interval 0.80–1.31), P = 0.84) or females\textsuperscript{35} (odds ratio = 1.14 (95% confidence interval 0.91–1.43), P = 0.25; Supplementary Fig. 15). This indicated that we were unlikely to see biases due to including CHOD from individuals who were missing general practitioner records. As such, we prioritized the use of CHOD in most analyses presented in the text, figures, and Supplementary Information—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 variety of conditions as they incorporate aforementioned HES data with both general practitioner records and self-reported outcomes.

To determine the role of 19,154 and 2,361 diseases, disorders, and special codes collated from HES and CHOD, respectively, in the relationship between s\textsubscript{het} burden and childlessness (Extended Data Figs. 4, 5, Supplementary Figs. 8, 9), we used a modified version of our primary logistic model of the form

\[
\text{logit}(p) = \beta_0 + \beta_1\text{c}\text{ode} + \beta_2\text{age} + \beta_3\text{birth cohort} + \beta_4\text{wes cohort} + \beta_5\text{PCL}\text{PC40} + \beta_6\text{rare}\text{PCL}\text{PC100}
\]

where \(\text{c}\text{ode}\) represents a binarized presence or 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 two individuals in both genetic (that is, CNVs) and WES (that is, PTVs) data. When considering individuals who have a particular code c, we used 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 association with childlessness, we also considered individuals with any sub-code (that is, 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 (for example, E11.0). For each model, we retained both an odds ratio for the association of individual s\textsubscript{het} burden and presence of a given code with childlessness (Supplementary Table 2).

**Evaluation of gene expression**
To determine the expression in various tissues of all genes assessed in this study, we downloaded processed median transcripts per million (TPM) values for all genes provided by v7 of the GTEx study\textsuperscript{10} (https://storage.googleapis.com/gtex_analysis_v8/ma_qseq_data/GTEx_Analysis_2017-06-05_v8_RNASeQCv1.1.9_gene_median_tpm.gct.gz). Only genes for which an s\textsubscript{het} score was available were retained from this file.

To evaluate genes known to be involved in male infertility, we assessed 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: (1) as a factor of being a male infertility gene or not or (2) having or not having a qualifying variant (Supplementary Fig. 7).

To assess if genes expressed in a particular tissue modulated the relationship between s\textsubscript{het} burden and childlessness, we also generated tissue-specific s\textsubscript{het} scores for all individuals where we excluded all genes with a tissue-specific TPM value greater than 0.5. s\textsubscript{het} scores for all 53 tissues were calculated as described above and used to test the association between s\textsubscript{het} burden and childlessness (Supplementary Fig. 10).

To examine the correlation between expression and s\textsubscript{het}, we performed two linear regressions in all 53 tissues queried by the GTEx study: (1) the association between s\textsubscript{het} and ln(expression\textsubscript{s\textsubscript{het}}) corrected for gene coding sequence length and (2) the association between s\textsubscript{het} times the number of singleton loss-of-function variants and ln(expression\textsubscript{s\textsubscript{het}}) corrected for gene coding sequence length. For each tissue and model, we extracted the overall variance explained for each model (\(R^2\)) and plotted it as part of Supplementary Fig. 11.

**Modelling the contribution of phenotypes to observed reduction in fitness**

**Variant s\textsubscript{het} burden.** To estimate the contribution of s\textsubscript{het} to overall fitness (Extended Data Fig. 8b), we extracted log odds ratio estimates for the association of s\textsubscript{het} with having children from our logistic model and estimated the proportion of childless individuals at various s\textsubscript{het} scores (0 to 1 at 0.1 intervals; Extended Data Fig. 9). To calculate the error in our estimates (that is, the shaded areas in Extended Data Fig. 8b), we used the 95% confidence intervals for the s\textsubscript{het} burden log odds ratio from our original logistic regression. We note that we assume that input parameters calculated with UK Biobank data are generalizable to the UK population as a whole, but may slightly underestimate the error in our final result. See Supplementary Note 2 for a more detailed description of how the contribution of s\textsubscript{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 Fig. 15). 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\textsuperscript{15}. To assess assignment into different branches of a universal conscription for Swedish men, the Swedish government includes an extensive cognitive ability test which all men in Sweden had to take part in. Information on childbirth 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 1963–1967. The conscription registers are described in more detail elsewhere\textsuperscript{32,37}.

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\textsuperscript{34}, 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 three years in our conscription registers. The Swedish military created an official IQ-measure based on a nine-score stanine scale that has been used in a large number of scientific studies\(^{4,23}\). 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\(^{25,47,52}\). 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 standardizing 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 \(\mu = 100\) and \(\sigma = 12\), as compared to the official stanine measure with \(\mu = 100\) and \(\sigma = 15\).

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\(^{34}\) and others). In addition to the -5\% of men who did not take the test (for example, they were ineligible for military service owing to disability (such as visual impairments), that they were abroad or were conscripted at an atypical age), we additionally excluded a number of men for whom scores of all test batteries were not available. Our manually computed fine-grained measure was later standardized 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 association of overall \(s_{\text{het}}\) burden with fluid intelligence (Fig. 2f) and, because fluid intelligence is normalized and IQ is normally distributed, converted this effect size 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 Fig. 14; 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 \(s_{\text{het}}\) 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 \(s_{\text{het}}\) score (Supplementary Fig. 13). This value was then compared to the mean number of children for the unburdened population via the following formula:

\[
\text{fertility ratio} = \frac{\text{fertility}_{\text{het}(i)}}{\text{fertility}_{\text{het}(0)}}
\]

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 \(s_{\text{het}}\) as described above. See Supplementary Note 3 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\(^{17}\) (Supplementary Fig. 15). In this case, as we were unable to accurately estimate the increased risk of developing individual mental health disorders as a factor of individual \(s_{\text{het}}\) burden, we instead used odds ratios from Ganna et al.\(^{4}\). Only odds ratios for schizophrenia, autism spectrum disorder, and bipolar disorder were retained. As Ganna et al.\(^{4}\) estimated the risk based on total count of high pLI (\(\geq 0.9\)) genes\(^{42}\) with PTVs per individual instead of with \(s_{\text{het}}\), we assumed that an individual carrying one such variant had an \(s_{\text{het}}\) burden of 0.162, or the mean \(s_{\text{het}}\) value of all high pLI (\(\geq 0.9\)) genes. We then converted this into a proportion of individuals with a given mental health disorder at \(s_{\text{het}}\) burden \(x\) by scaling the odds ratio with the following formula:

\[
\log(\text{OR}_{\text{het},\text{r}}) = \frac{\log(\text{OR}_{\text{Ganna}}) \times s_{\text{het}}[X]}{0.162}
\]

To establish a baseline expectation for the prevalence of each mental health disorder at \(s_{\text{het}}\) we used population-level data from Power et al.\(^{20}\) and extrapolated for each trait at increasing \(s_{\text{het}}\) values (Supplementary Fig. 21). To generate an expected mean number of children for simulated individuals with mental health disorders, we used fertility statistics generated by Power et al.\(^{46}\). As Power et al.\(^{46}\) 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 \(s_{\text{het}}\) burden and childlessness, we used a multiple regression model incorporating various combinations of these traits (Extended Data Fig. 8a, Supplementary Fig. 17). First, we calculated the variance in childlessness explained by a null model consisting of age, age\(^2\), birth cohort, WES batch, the first 40 ancestry PCs, and the first 100 IBD-derived 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 \(s_{\text{het}}\) alone, we calculated incremental pseudo-\(R^2\) between this null model and a model additionally incorporating a term for PTV \(s_{\text{het}}\) burden. We then repeated this analysis, except now including an additional covariate (for example, having a partner at home) to determine the reduction in variance explained by PTV \(s_{\text{het}}\) when correcting for the additional covariate. This reduction in variance was then converted to a percent change via the following formula:

\[
\% \text{ reduction in variance explained by } PTV_{\text{het}} = \frac{s_{\text{het}}\text{incremental}^2\text{with covariate}}{s_{\text{het}}\text{incremental}^2\text{without covariate}}
\]

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 (Extended Data Fig. 8a, Supplementary Fig. 17). Percent reduction in variance explained values plotted in Extended Data Fig. 8a and Supplementary Fig. 17 are displayed for \(s_{\text{het}}\) calculated using PTVs 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 (Extended Data Fig. 8a). As income is provided by UK Biobank on a per-household basis (Fig. 2e) and the number of individuals with fluid intelligence data recorded at recruitment is significantly smaller than for other covariates (Fig. 2f), we did not include these as part of our multiple regression model.
Article

Reporting summary
Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability
Raw data produced by this study are available as part of the UK Biobank data return catalogue with application ID 44165: https://biobank.ndph.ox.ac.uk/ukb/docs.cgi?id=1.

Code availability
Code used as part of this project to perform phenotype testing, CNV calling, variant quality control and to generate all main text figures, supplementary figures and supplementary tables is available on github: https://github.com/HurlesGroupSanger/UKBBFertility. All statistical analysis in this manuscript was performed using R v3.6.0.

Acknowledgements
We thank L. Parts, J. Kaplanis, M. Przeworski and G. Davey-Smith for useful discussions and advice on data analysis; M. Oud and J. Veitman for helpful discussions regarding infertility; G. Kalantzis and P. F. Palamara for assistance with correcting for recent ancestry; and the INTERVAL study for sharing genotyping and exome data that enabled us to refine our CNV filtering methodology. This work has been funded by core Wellcome funding to the Wellcome Sanger Institute (grant WT098051) and as part of a Medical Research Council (MRC) Centre Grant to the MRC Centre for Neuropsychiatric Genetics and Genomics (grant MR/L010305/1). K.B. is supported by a fellowship from the Bank of Sweden Tercentenary Foundation (Riksbankens Jubileumsfond). 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., K.E.S., H.C.M. and M.E.H. designed experiments, oversaw the study and wrote the manuscript. E.J.G. and M.E.K.N. annotated and assessed SNV and indel variants. 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. is supported by a fellowship from the Bank of Sweden Tercentenary Foundation (Riksbankens Jubileumsfond). This work has been conducted using the UK Biobank Resource under application numbers 14421 (to G.K.) and 44165 (to H.C.M.).

Competing interests
M.E.H. is a founder of, director of, consultant to, and holds shares in, Congenica Ltd and is a consultant to the AstraZeneca Centre for Genomics Research.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-022-04549-9.

Correspondence and requests for materials should be addressed to Matthew E. Hurles.

Peer review information Nature thanks Loic Yengo and the other, anonymous reviewers for their contribution to the peer review of this work. Peer review reports are available.

Reprints and permissions information is available at http://www.nature.com/reprints.
Extended Data Fig. 1 | Odds ratio estimates for the association of s_{het} burden with having children for all variant classes. Identical plot to main text Fig. 1a, but with additional data for synonymous, missense, and duplication s_{het} scores, separated into females (violet) and males (jade). Asterisks indicate significance after Bonferroni correction for 20 tests (P < 2.5 × 10^{-3}; Methods).
Extended Data Fig. 2 | Effect size estimate for the association of $s_{\text{het}}$ burden with number of children for individuals with children. Shown are the effect size estimates for the association of $s_{\text{het}}$ burden with number of children, separated into females (purple) and males (jade), but with all childless individuals in the UK Biobank removed. Like Main Text Fig. 1a, the regression used to generate the displayed result used the raw number of children, live births for females and children fathered for males, rather than a binary value for having children. Asterisks indicate significance after Bonferroni correction for 20 tests ($P < 2.5 \times 10^{-3}$; Methods).
**Extended Data Fig. 3 | Odds ratio estimate for \( s_{\text{het}} \) burden stratified by age group.** Shown are odds ratio estimates for the association of \( s_{\text{het}} \) burden with having children, stratified by participant age (y-axis) and separated into females (violet) and males (jade). Age range intervals are left-open. Dash of the line indicates whether the estimate comes from \( s_{\text{het}} \) burden calculated from deletions (long dash), PTVs (short dash), or from a fixed effects meta-analysis (no dash). Also shown for reference are the results for all individuals regardless of age (All Ages), which is identical to the result shown in main text Fig. 1b. Asterisks indicate significance after Bonferroni correction for 20 tests \((P < 2.5 \times 10^{-3}; \text{Methods})\).
Extended Data Fig. 4 | Mediation of the relationship between sHof burden and childlessness by various disorders. Plotted is the deletion and PTV meta-analysis \(-\log_{10} P\) value for the association between sHof 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). Remaining ICD-10 codes at different levels on the ICD-10 hierarchy are not displayed here. Results are ordered first by ICD-10 chapter (x-axis) and then by increasing \(-\log_{10} P\) value (y-axis). The arrow for code Z37 indicates the point is below the scale of the y-axis with \(-\log_{10} P\) value indicated in parentheses. Visual outliers are labelled and do not imply a significant change in the effect size of sHof on childlessness.
Extended Data Fig. 5 | Mediation of the association between \( s_{het} \) burden and childlessness by various disorders. Depicted are the results of our primary association between childlessness and individual \( s_{het} \) burden corrected for presence/absence of approximately 2,000 different disorders, diseases, and health factors queried from (A,B) hospital episode statistics and (C,D) complete health outcomes data as represented by the ICD-10 medical coding system separately for (A,C) males and (B,D) females (see main text methods). Shown on the x-axis is the \(-\log_{10} p\) value for the association of \( s_{het} \) with having children, corrected for a given diagnostic code. On the y-axis is the \(-\log_{10} p\) value for having a given medical code on likelihood of having children; \( p\) values are placed above or below \( y = 0 \) based on the direction of effect, with disorders which are associated with having children above and those associated with not having children below. Codes were chosen for labeling to highlight outliers and not based on any statistical criteria. Codes with points at the top or bottom of plots have \(-\log_{10} p\) values \( \geq 100 \). Color of points and text is based on the ICD-10 chapter. Please note that text labels do not necessarily represent the full official name of a given ICD-10 code.
Extended Data Fig. 6 | Association of eight relevant phenotypes/demographic measures with the likelihood of having children among UK Biobank participants. Shown are the results of a logistic regression estimating the odds ratio for the relationship of (A) having a partner at home, (B) ever having had sex (C) completing university, (D) having a severe mental health disorder, (E) household income, (F) fluid intelligence, (G) Townsend deprivation index, and (H) engaging in same sex sexual behaviour with likelihood of having children, separated into females (violet) and males (jade). 95% confidence intervals for all plots are included, but may be invisible at the resolution of the figure. Please note that the scales of the x-axis for plots (A) and (B) are different from plots (C-H) due to the relatively stronger association of these traits with having children.
Extended Data Fig. 7 | Effect size estimates for the association between s_{het} burden and Townsend Deprivation Index. Shown are the effect size estimates for the association between s_{het} burden and Townsend Deprivation Index, separated into females (purple) and males (jade). Units are unnormalized Townsend Deprivation Indices for each individual in the UK Biobank. Asterisks indicate significance after Bonferroni correction for 20 tests (p < 2.5 × 10^{-3}; Methods).
Extended Data Fig. 8 | The role of individual phenotypes in the relationship between s\textsubscript{het} burden, childlessness, and fitness. (A) Odds ratio estimates for the association of cumulative deleterious variation for a combined meta-analysis (deletions + PTVs) with childlessness (middle), corrected for a combination of whether or not a study participant has a mental health (MH) disorder, a partner at home, a university degree, infertility (as ascertained from Complete Health Outcomes Data; Methods), 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 either nominal (*) or Bonferroni-corrected (**) significance level with childlessness for each covariate independently when correcting for PTV s\textsubscript{het} burden. For all possible combinations of these traits, see Supplementary Fig. 17. As indicated by coloured boxes, all models include s\textsubscript{het} burden and were run separately for males and females. The marginal bar plot to the right gives the proportion of the variance in childlessness explained by s\textsubscript{het} burden as calculated for PTVs only, scaled to the model which only includes s\textsubscript{het} burden (i.e. the model on the bottom of the plot). (B) Predicted reduction in overall fitness as a factor of increasing s\textsubscript{het} burden. Displayed is the expected reduction in fitness as a factor of increasing s\textsubscript{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 (Fig. 1b; Methods). The dashed line represents the theoretical reduction in fitness as predicted by s\textsubscript{het} \textsuperscript{4}. 
Extended Data Fig. 9 | The association of \( s_{het} \) burden with childlessness.
Identical to Extended Data Fig. 8B, except in this instance, the y-axis represents predicted childlessness as a factor of individual \( s_{het} \) burden, rather than predicted reduction in fitness. Values at \( x = 0 \) represent actual mean childlessness among UK Biobank males (jade) and females (violet) with an \( s_{het} \) burden of 0.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted. Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection The majority of the data collection for this study was performed by the UK Biobank consortium and has been previously published/described (Brycroft et al. Nature, 2018; Van Hout et al. Nature Genetics, 2020; Szustakowski et al. Nature Genetics, 2021). Collection of data used for the assessment of cognition in the Swedish population has been both previously described (Kolk and Barclay, Proc. Royal Society B, 2019) and described extensively in the methods of the manuscript.

Data analysis We have made a public github repository which holds all code that was created by our research team to perform the analysis and generate all figures and tables. See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Our current data availability statement is: "Raw data produced by this study are available as part of the UK Biobank data returns catalogue with application ID 44165: https://biobank.ndph.ox.ac.uk/ukb/docs.cgi?id=1." UK Biobank requires research utilizing UK Biobank human participants to be deposited to their own secure platform to ensure patient confidentiality. Linking of this raw data to specific figures/analyses is handled in our code availability statement."
Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences  
- Behavioural & social sciences  
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

**Sample size**
The only filtering of individuals based on sample size was to restrict UK Biobank participants to unrelated individuals of broadly European genetic ancestry (n = 328,899) and, for SNV/InDel analyses, to the subset of participants with whole exome sequencing data (N = 139,480). Collection of participants is described in the relevant UK Biobank papers.

**Data exclusions**
We excluded a small subset of individuals (~2k) from CNV analyses who had an excessive number of variant calls indicating low array quality. When performing linear/logistic regressions, we excluded a set of ~14k individuals that did not have IBD sharing data as provided by Nait Saada et al. as described in the methods.

**Replication**
All findings are reproducible, and code to reproduce our result is included as part of our github repository included in our code availability statement.

**Randomization**
We did not separate individuals into case/control groups as our primary regression uses a continuous dependent variable which was calculated independently for each participant. Males and females were separated based off of sex reported at individual recruitment.

**Blinding**
Blinding was not relevant to our study as no patient interaction was performed by our researchers.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

| Materials & experimental systems      | Methods                         |
|---------------------------------------|---------------------------------|
| n/a Involved in the study             | n/a Involved in the study       |
| - Antibodies                          | - ChiP-seq                      |
| - Eukaryotic cell lines               | - Flow cytometry                |
| - Palaeontology and archaeology       | - MRI-based neuroimaging        |
| - Animals and other organisms         |                                 |
| - Human research participants         |                                 |
| - Clinical data                       |                                 |
| - Dual use research of concern        |                                 |

Human research participants

Policy information about studies involving human research participants

| Population characteristics | Individuals included in this study were recruited by UK Biobank. We restricted UK Biobank participants to unrelated, broadly European genetic ancestry individuals. |
|-----------------------------|----------------------------------------------------------------------------------------------------------------------------------|
| Recruitment                 | Individuals were recruited by the UK Biobank. Recruitment information can be obtained from the UK Biobank papers/website [https://www.ukbiobank.ac.uk/] |
| Ethics oversight            | The UK Biobank approved our study design (application numbers 14421 and 44165)                                                  |

Note that full information on the approval of the study protocol must also be provided in the manuscript.