



EVOLUTIONARY BIOLOGY

Evidence for the role of selection for reproductively advantageous alleles in human aging

Erping Long^{1,2} and Jianzhi Zhang^{2*}

The antagonistic pleiotropy hypothesis posits that natural selection for pleiotropic mutations that confer earlier or more reproduction but impair the post-reproductive life causes aging. This hypothesis of the evolutionary origin of aging is supported by case studies but lacks unambiguous genomic evidence. Here, we genomically test this hypothesis using the genotypes, reproductive phenotypes, and death registry of 276,406 U.K. Biobank participants. We observe a strong, negative genetic correlation between reproductive traits and life span. Individuals with higher polygenic scores for reproduction (PGS_R) have lower survivorships to age 76 (SV_{76}), and PGS_R increased over birth cohorts from 1940 to 1969. Similar trends are seen from individual genetic variants examined. The antagonistically pleiotropic variants are often associated with cis-regulatory effects across multiple tissues or on multiple target genes. These and other findings support the antagonistic pleiotropy hypothesis of aging in humans and point to potential molecular mechanisms of the reproduction–life-span antagonistic pleiotropy.

INTRODUCTION

Aging or senescence refers to a gradual deterioration of bodily functions that manifests as a decline in reproductive performance and an increase in the death rate with age (1). The prevalence of aging across multicellular organisms has been an evolutionary puzzle because natural selection should favor mutations conferring an extended reproductive life span (2). In 1957, Williams (3) proposed that mutations contributing to aging could be positively selected if they are advantageous early in life, promoting earlier reproduction or more offspring. This antagonistic pleiotropy hypothesis has become one of the leading theories of the evolutionary origin of aging (see Discussion for other theories). Nonetheless, despite observations in life-history evolution of trade-offs consistent with the antagonistic pleiotropy hypothesis (4, 5), the involvement of mutational pleiotropy is usually difficult to establish (6). Experimental studies of model organisms have identified cases of antagonistic pleiotropy between reproduction and life span (6–9), but how frequently pleiotropy occurs between reproduction and longevity and how common this pleiotropy is antagonistic as opposed to concordant are unknown at the genomic scale. Further complicating the situation is the fact that phenotypes measured in the laboratory can often deviate from those measured in nature (6). In humans, cases consistent with the antagonistic pleiotropy hypothesis are known (10, 11). For example, genetic variants associated with coronary artery disease tend to be associated with a higher number of children ever born (NEB) (10). Furthermore, Wang *et al.* (12) observed a significant, negative genetic correlation between NEB and life span among female (but not male) participants of the Framingham Heart Study in some of their analyses. However, evidence for the antagonistic pleiotropy hypothesis in humans have been contested (13–16), and the hypothesis still lacks unambiguous genome-wide support. Furthermore, it is unclear whether

mutations contributing to aging were selectively favored and whether the selection arose from their benefits earlier in life.

One of the primary reasons why testing the antagonistic pleiotropy hypothesis is difficult is that life-history traits tend to be influenced by many small-effect genetic variants (17, 18). The U.K. Biobank has collected the genotypes and various phenotypes of about 0.5 million participants (19), offering an unprecedented opportunity for testing the antagonistic pleiotropy hypothesis in humans. In particular, the data include measures of human reproduction such as age at menarche, number of offspring, and sexual dysfunction as well as exact death information (20). These phenotypic data along with genotypic data allow assessing the genetic relationship between reproduction and life span at the genomic scale and at the level of individual segregating variants. Furthermore, functional genomic data [e.g., expression quantitative trait locus (eQTL)] can be mined to uncover the putative pathways and mechanisms underlying any reproduction–life-span relationship (21, 22). Taking advantage of these resources, we performed multiple analyses outlined in Fig. 1 to address the following questions about the antagonistic pleiotropy hypothesis. First, do genetic variants influencing reproduction more likely affect life span than expected by chance? Second, is the pleiotropy between reproduction and life span largely antagonistic? Third, are pleiotropic mutations promoting reproduction but causing aging favored by natural selection? Last, what are the potential molecular mechanisms linking reproduction and aging?

RESULTS

Negative genetic correlation between reproduction and life span

The genetic correlation between two phenotypic traits is the proportion of variance that the two traits share due to genetic causes and is a measure of the contribution of pleiotropy to the covariation of the traits (23). Because the antagonistic pleiotropy hypothesis of aging posits that the same mutations antagonistically affect reproduction and life span, the hypothesis predicts a negative genetic correlation

¹Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100005, China. ²Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA. *Corresponding author. Email: jianzhi@umich.edu

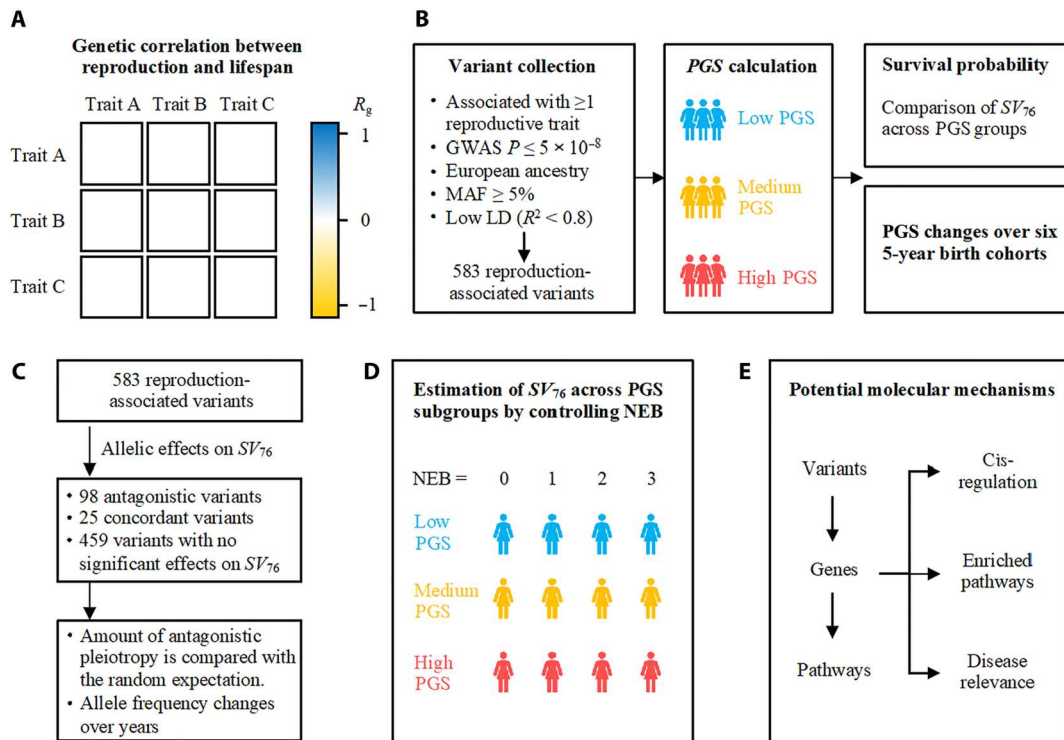


Fig. 1. Schematics summarizing the approaches and analyses used in the present study. (A) Genetic correlation between reproduction and life span identified from a list of previously computed genetic correlations for heritable U.K. Biobank traits. Available traits include three reproductive traits—nAFB, nAFS, and NCF—as well as two parental life-span traits—father’s age at death and mother’s age at death. (B) A total of 583 reproduction-associated variants are collected, and PGSs are calculated for four reproductive traits: (i) nAFB, (ii) nAFS, (iii) nAMC, and (iv) AMP. The probability of survival to age 76 (SV_{76}) is compared across equal-size groups of individuals with relatively low, medium, and high PGS for each of the four reproductive traits. The potential change of PGS over six 5-year birth cohorts from 1940 to 1969 is investigated. (C) Effects of individual variants on life span are estimated from SV_{76} . The observed amount of antagonistic pleiotropy is compared with the random expectation. Potential allele frequency changes over six 5-year birth cohorts from 1940 to 1969 are examined. (D) Potential horizontal pleiotropy between reproduction and life span is explored by controlling the NEB (NEB = 0, 1, 2, or 3). (E) Potential molecular mechanisms of antagonistic pleiotropy are explored by surveying the associated cis-regulatory activity of antagonistically pleiotropic variants, their target genes, enriched pathways, and relevance to human diseases. MAF, minor allele frequency; LD, linkage disequilibrium.

between reproduction and life span. Because the life spans of most U.K. Biobank participants are unknown (as they are still living), we examined the genetic correlation between the reproductive traits of U.K. Biobank participants and the life spans of their parents by searching the genetic correlations between heritable traits in U.K. Biobank previously computed by the Neale group (24) (see Materials and Methods). Note that both the number of offspring and timing of reproduction impact a person’s fitness because earlier reproduction means a shorter generation time and thereby a higher fitness even when the number of offspring is the same. We therefore focused on three reproductive traits available in the list—negative age at first birth (nAFB), negative age at first sex (nAFS), and number of children fathered (NCF); larger values of these traits correspond to higher reproduction (i.e., earlier reproduction and/or more offspring). Note that nAFB is a more accurate measure of the earliest reproduction of a participant than nAFS, but nAFB was measured for female U.K. Biobank participants only. Hence, we considered both nAFB and nAFS here, the latter of which was measured for both male and female participants. We examined two parental life-span traits—father’s age at death and mother’s age at death. Reproduction and parental life span show a significant negative genetic correlation in all pairwise comparisons between the three reproductive traits and two life-span traits (Table 1),

supporting the antagonistic pleiotropy hypothesis. Because NCF is measured in males only, whereas nAFB is measured in females only, our results indicate that the genetic correlation between reproduction and life span is negative in both sexes. Because the U.K. Biobank also recorded the number of full brothers and that of full sisters of each participant, we further examined the genetic correlation between a participant’s parental reproduction and parental life span. A significant genetic correlation was observed (data S1), further supporting the antagonistic pleiotropy hypothesis.

In addition to biological pleiotropy (one mutation affecting multiple traits) and statistical pleiotropy (linked mutations affecting multiple traits), both of which are relevant to the antagonistic pleiotropy hypothesis, genetic correlation could also be caused by two factors unrelated to the antagonistic pleiotropy hypothesis. The first is population stratification and the second is cross-trait assortative mating (25). Our above results should not be affected by population stratification because it had been controlled in genome-wide association analysis (GWAS) based on which genetic correlations were computed (24). There should be no assortative mating between reproductive traits and life span because life span is unknown at the time of mating.

Table 1. Genetic correlation between reproduction and parental life span. The R_g values are from www.nealelab.is/blog/2019/10/10/genetic-correlation-results-for-heritable-phenotypes-in-the-uk-biobank and are all significant at $P < 10^{-7}$ except for that between NCF and father's age at death ($P = 1.4 \times 10^{-5}$) and that between NCF and mother's age at death ($P = 2.0 \times 10^{-5}$).

Genetic correlation coefficient R_g	nAFS	nAFB	NCF	Father's age at death	Mother's age at death
nAFS	1	0.84	0.48	-0.59	-0.75
nAFB	0.84	1	0.45	-0.66	-0.80
NCF	0.48	0.45	1	-0.29	-0.33
Father's age at death	-0.59	-0.66	-0.29	1	0.91
Mother's age at death	-0.75	-0.80	-0.33	0.91	1

Higher polygenic scores for reproduction predict lower probabilities of survival to age 76

Although the life spans of many U.K. Biobank participants are unknown, the deaths of some participants allow reliable estimation of the probability of survival to various ages (up to the age of 76 years), providing an opportunity to assess the potential genetic impact of reproduction on the life span. We focused on 276,406 participants of British ancestry with no kinship to other participants in the database (see Materials and Methods) and considered 891 genetic variants reported to be associated with at least one reproductive trait with genome-wide significance ($P \leq 5 \times 10^{-8}$) (26). After the removal of variants with high linkage disequilibrium [R^2 (coefficient of determination) > 0.8 ; the variant with a relatively high P value is excluded] and those with low minor allele frequencies (< 0.05), 583 genetic variants remained, which are associated with the following reproductive traits: (i) nAFB, (ii) nAFS, (iii) negative age at menarche (nAMC), (iv) age at menopause (AMP), (v) NEB, which is the NCF for a male or number of children mothered for a female, and (vi) polycystic ovary syndrome (PCOS) (data S2 and see Materials and Methods).

An individual's polygenic score (PGS) of a trait reflects the individual's estimated genetic predisposition for the trait; it measures the individual's likelihood of having the trait based exclusively on genetics without taking environmental factors into account. To assess the aggregated effect of the above reproduction-associated genetic variants, we computed the PGSs of four reproduction traits each associated with at least 50 variants: nAFB, nAFS, nAMC, and AMP (see Materials and Methods). nAMC, nAFS, and nAFB measure the onset of reproduction, while AMP measures the end of reproduction, so together they inform the timing and potential amount of reproduction. NEB and PCOS are not considered here because they each have fewer than 50 variants. The survival rate per year from the age of 40 to 76 years and the cumulative survival probability for homozygotes of each variant were calculated (see Materials and Methods). The log-rank method was used to test the difference in survivorship between two genotypes. For example, individuals ranked in the top third in PGS for nAFB (PGS_{nAFB}) have a significantly lower probability of survival to age 76 ($SV_{76} = 0.800$) than that of individuals ranked in the bottom

third in PGS_{nAFB} ($SV_{76} = 0.839$) ($P = 3.5 \times 10^{-4}$; Fig. 2A), supporting the antagonistic pleiotropy hypothesis. Qualitatively similar results were observed when the PGS for each of the other three reproductive traits was examined (Fig. 2B and fig. S1).

Dividing the 276,406 individuals into six 5-year birth cohorts (1940–1944, 1945–1949, ..., and 1965–1969), we investigated the change of the PGS of each of the four reproductive traits over birth cohorts. As shown in Fig. 2C, for each of these traits, PGS steadily increased over time, presumably a result of natural selection for higher reproduction (see below for the analysis of individual variants that controls the potential age effect and genetic drift effect).

Together, the analyses of PGS for reproduction provide evidence for (i) antagonistic pleiotropy between reproduction and life span and (ii) natural selection for higher reproduction (presumably to the detriment of life span), supporting the antagonistic pleiotropy hypothesis of the origin of aging.

Genetic variants exhibiting antagonistic pleiotropy between reproduction and life span

We further tested the antagonistic pleiotropy hypothesis by identifying genetic variants associated with both reproduction and life span. On the basis of the directions of significant effects of an allele on reproduction and life span (see Materials and Methods), we inferred whether a variant shows antagonistic pleiotropy (Fig. 3A), concordant pleiotropy (Fig. 3B), or no pleiotropy. Among the 583 reproduction-associated variants, 123 variants have significant effects on SV_{76} [false discovery rate (FDR) < 0.05], and antagonistic pleiotropy cases significantly outnumber concordant pleiotropy cases (98 antagonistic versus 25 concordant pleiotropy cases; $P < 0.00001$, two-tailed binomial test; Fig. 3C and data S3). As a negative control, we sampled 100 sets of 583 random variants with matched allele frequencies ($\pm 1\%$) and recombination rates (± 0.05 cM/Mb), but none of these sets contained ≥ 123 variants with significant effects on SV_{76} . There is also no significant bias toward antagonistic pleiotropy for these control variants (on average, 13 antagonistically versus 12 concordantly pleiotropic variants per set if the matched allele is considered beneficial to reproduction). Therefore, compared with random polymorphisms, those impacting reproduction are 4.9 times more likely to affect life span and 7.5 times more likely to affect life span antagonistically.

Because average participants in the U.K. Biobank live many more years after the end of reproduction, the antagonistic pleiotropy hypothesis predicts that most mutations that increase reproduction but reduce life span have larger fitness advantages than disadvantages so are selectively favored. To verify this prediction, for each of the 98 variants exhibiting antagonistic pleiotropy between reproduction and life span, we investigated the frequency change of the allele beneficial to reproduction over the six birth cohorts after the correction for different ages of different birth cohorts at the time of the U.K. Biobank recruitment (see Materials and Methods) (27). At an FDR of 0.05, we detected significant allele frequency increases at 12 variants and decreases at 2 variants (Fig. 3D), the former being significantly more prevalent than the latter ($P = 0.006$, one-tailed binomial test; Fig. 3E and data S4). To account for the effect of genetic drift on allele frequency changes (28), we applied a robustness test (see Materials and Methods) and validated 6 of the above 14 cases, all showing allele frequency increases (data S4). These results demonstrate that

among polymorphisms with antagonistic pleiotropy between reproduction and life span, alleles advantageous to reproduction tend to be selectively favored, supporting the antagonistic pleiotropy hypothesis.

Horizontal pleiotropy between reproduction and life span

To understand the antagonistic pleiotropy between reproduction and life span, we investigated the relationship between the PGSs of the four reproductive traits and SV_{76} by controlling the NEB (NEB = 0, 1, 2, or 3). Individuals with four or more children are too few to allow reliable estimation of SV_{76} in each PGS subgroup. Among individuals with the same NEB, SV_{76} is negatively correlated with the PGS for each of the four reproductive traits (Fig. 4), suggesting that the antagonistic pleiotropy between reproduction and life span is, at least in part, independent from the actual amount of reproduction. In other words, there must be horizontal antagonistic pleiotropy between reproduction and life span. We also performed the analysis for males and females separately and observed a negative correlation between PGS and SV_{76} in each stratified subgroup

without significant gender-specific effects, although the negative correlation is sometimes not significant due to reduced sample sizes (data S5).

Given the reproductive PGS (regardless of which of the four reproductive traits is considered), individuals with NEB = 2 had the highest SV_{76} among the four groups of individuals with different NEBs (Fig. 4). Our finding is consistent with the report that females with two children had the lowest mortality risk among all females in the Framingham Heart Study (12), although our analysis further controlled reproductive PGS. These results indicate that the relationship between NEB and longevity is complex; it is positive in the low NEB range (below 2) but negative in the high NEB range (exceeding 2).

Potential molecular mechanisms

To explore the potential mechanisms involved in the antagonistic pleiotropy between reproduction and life span, we queried an aggregated eQTL database—eQTL catalog (22). This is because the antagonistically pleiotropic variants detected are mostly in noncoding

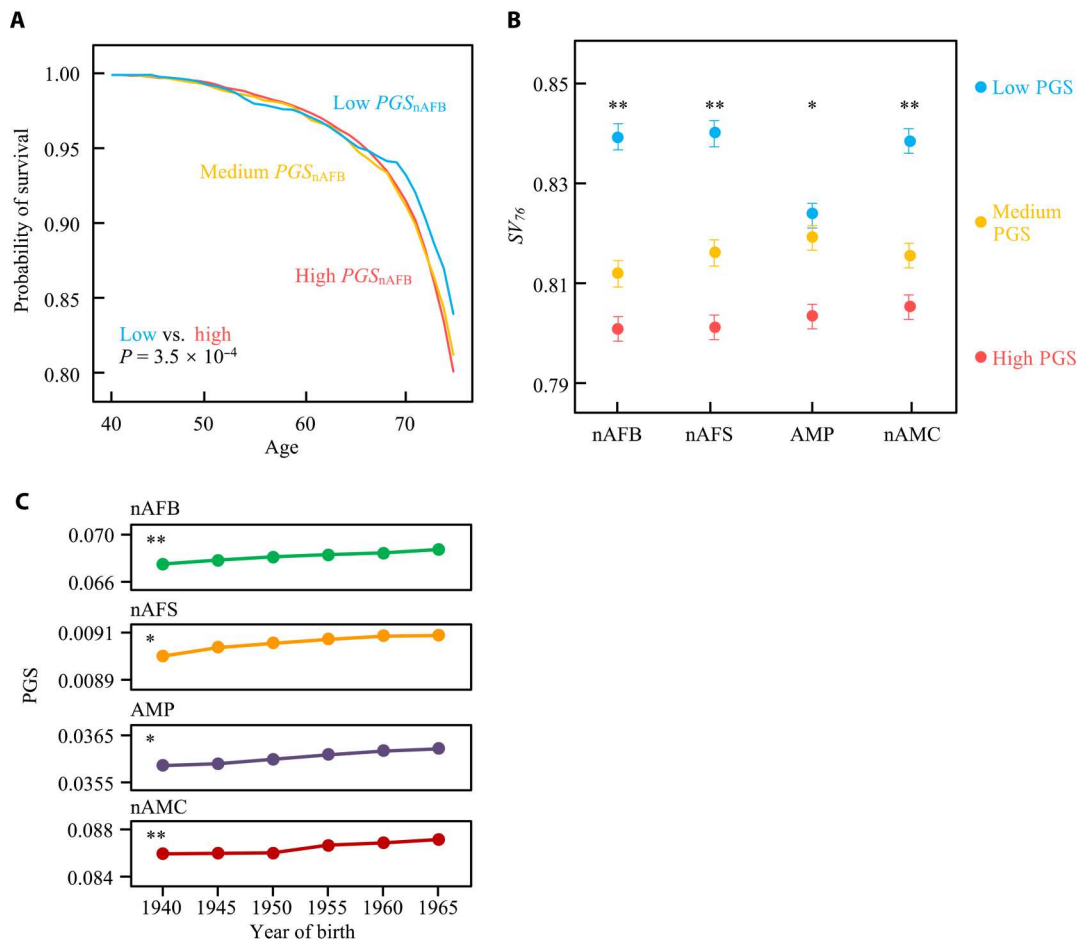


Fig. 2. Higher PGSs for reproduction predict lower probabilities of survival to the age of 76 years (SV_{76}). (A) SV_{76} of three equal-size groups of individuals with relatively low, medium, and high PGS for nAFB (PGS_{nAFB}). $SV_{76} = 0.839, 0.811,$ and 0.800 for relatively low, medium, and high PGS_{nAFB} groups, respectively. Plots of the other three reproductive traits are in fig. S1. (B) SV_{76} of three equal-size groups of individuals with relatively low, medium, and high PGS for each of the four reproductive traits (nAFB, nAFS, nAFC, AMP, and nAMC). Error bars show 95% confidence intervals calculated by Clopper-Pearson exact methods. Log-rank $*P < 0.05$ and $**P < 0.001$ between the relatively low and high PGS groups. (C) Mean PGS for each reproductive trait inferred from genotypes of each of the six 5-year birth cohorts. Error bars showing standard errors are too small to see. $*P < 0.05$ and $**P < 0.001$, linear correlation between mid-point birth year of a cohort and PGS.

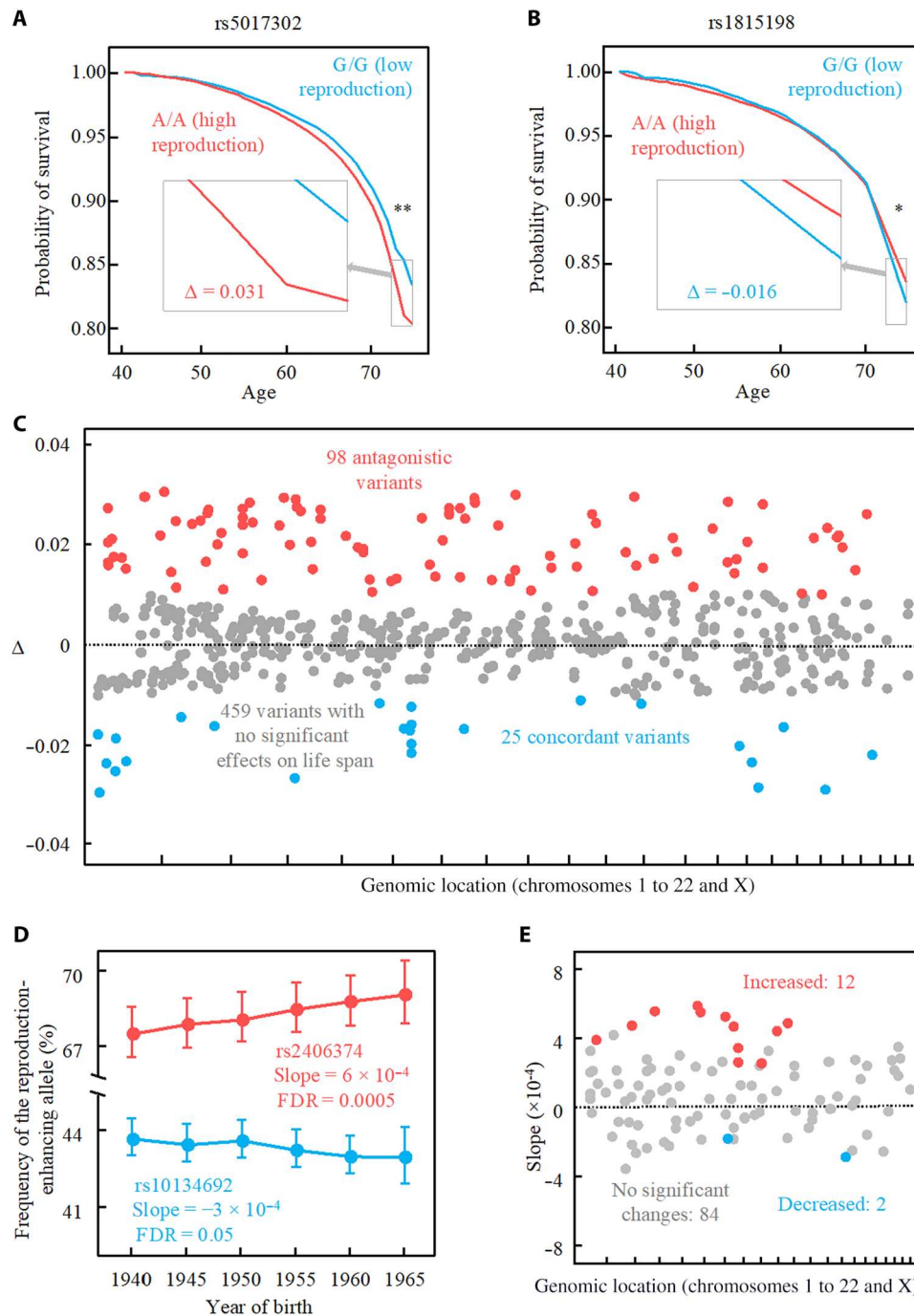


Fig. 3. Reproduction-enhancing alleles tend to lower life span and be selectively favored. (A) An example of a variant that has antagonistic effects on reproduction (nAMC) and probability of survival to age 76 (SV_{76}). $SV_{76} = 0.835$ for G/G (lower reproduction genotype) and 0.804 for A/A (higher reproduction genotype) (FDR = 1.4×10^{-6}). Δ is the difference in SV_{76} between lower and higher reproduction homozygotes for the variant considered. (B) An example of a variant with concordant effects on reproduction (AMP) and SV_{76} . $SV_{76} = 0.820$ for T/T (lower reproduction genotype) and 0.836 for C/C (higher reproduction genotype) (FDR = 0.001). *FDR < 0.05 and **FDR < 0.001. (C) Effects (Δ) on SV_{76} of 583 reproduction-associated variants. Each dot represents one variant, whose genomic coordinate is shown on the x axis (chromosome 1 to 22, and X for each interval from left to right). Red and blue dots indicate variants with significant antagonistic and concordant effects on SV_{76} , respectively, whereas gray dots indicate variants with no significant effects on SV_{76} . The horizontal line indicates no effects on SV_{76} . (D) Two examples of variants showing allele frequency changes over six 5-year birth cohorts (red for an increased trend and blue for a decreased trend). Slope and FDR are from linear regression. Error bars represent standard errors. (E) Allele frequency changes for 98 antagonistically pleiotropic variants. Each dot represents one variant, whose genomic coordinate is shown on the x axis (chromosome 1 to 22). Y axis refers to the slope of the linear regression as in (D). Red and blue dots indicate significantly positive and negative slopes, respectively, whereas gray dots indicate nonsignificant slopes. The horizontal line indicates a slope of 0.

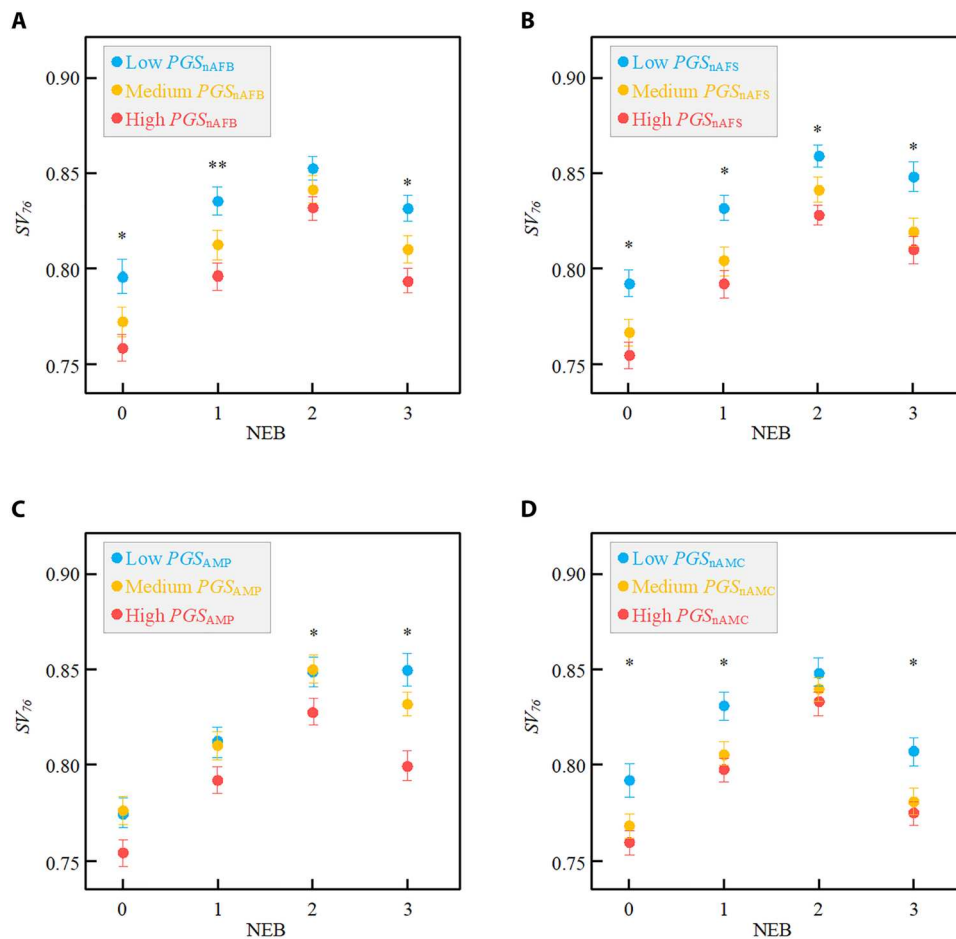


Fig. 4. Probability of survival to age 76 (SV_{76}) as a function of the NEB and PGSs for four reproductive traits. The four reproductive traits include nAFB (A), nAFS (B), AMP (C), and nAMC (D). Each dot represents individuals with the highest, medium, or lowest third of PGS considered. Error bars show 95% confidence intervals calculated by Clopper-Pearson exact methods. * $P < 0.05$ and ** $P < 0.001$ between the low and high PGS groups. Gender-specific results are in data S5.

regions and may play cis-regulatory roles in regulating target gene (eGene) expression. We divided the 583 reproduction-associated variants into three sets, depending on whether the variants have significant antagonistic effects on life span (antagonistic pleiotropy set), significant concordant effects on life span (concordant pleiotropy set), or no significant effects on life span (control set). Relative to the control set, antagonistic and concordant pleiotropy sets both contain a higher fraction of variants associated with cis-regulatory activities (eQTLs with empirical genome-wide significance) (Table 2). We defined three different types of cis-regulatory events: multicontext (associated with cis-regulatory activities in

multiple tissues/cell lines), multi-eGene (associated with the expressions of multiple eGenes), and discordant (discordant effects on an eGene across tissues/cell lines). When compared with the control set, antagonistic and concordant pleiotropy sets are enriched with multicontext and multi-eGene events (Table 2). In addition, the antagonistic pleiotropy set is enriched with discordant events (Table 2). Note that these differences are not due to potential disparities in statistical power in eQTL mapping across the three sets of variants because we found no significant differences in minor allele frequencies across them. These results suggest that a common molecular mechanism of antagonistic pleiotropy between reproduction

Table 2. Number of reproduction-associated variants associated with cis-regulatory activities. * $P < 0.05$ and ** $P < 0.001$ in a chi-squared test via comparison with the bottom set of variants.

Set name	No. of variants	Cis-regulatory activity	Multi-context events	Multi-eGene events	Discordant events
Antagonistic effect on longevity	98	75*	62**	54*	24*
Concordant effect on longevity	25	18*	15*	15*	6
No effect on longevity	459	221	124	162	64

and life span is for a cis-regulatory mutation to influence the expressions of the same target gene in different tissues (sometimes with opposite directions) and to influence the expressions of multiple target genes.

A total of 230 candidate eGenes from the antagonistic pleiotropy set reached empirical genome-wide significance in eQTL analysis. Significantly enriched pathways for these eGenes included those related to hormone (estrogen receptor signaling), immune response (IL-15 production), and age-related pathology (amyloid processing and protein ubiquitination pathway) (data S6). Additional mechanistic insights can be gained from examining individual antagonistically pleiotropic variants. For example, variant rs12203592 at chromosome 6p25.3 has its major T allele associated with a younger age at first sexual intercourse and an increased risk of mortality and shows a rise in the frequency of the T allele over 25 years (data S3). Previous GWAS suggested that the T allele is additionally associated with decreased parental life span (29) and increased risk of multiple cancers such as melanoma (30) and lung cancer (31). This variant is also associated with multicontext, multi-eGene (*IRF4* and *EXOC2*), and discordant cis-regulatory events. Specifically, the T allele is correlated with higher *IRF4* expressions in the lung and whole blood but lower *IRF4* expression in skin tissues. Higher *IRF4* expressions in the human lung were reported to promote endogenous DNA damage in lung fibroblasts (31). Another antagonistically pleiotropic variant—rs34811474 at chromosome 4p15.2—has its reproduction-enhancing G allele associated with an increased risk of osteoarthritis, an age-related degenerative disease (32). The rs7719499 G allele, in linkage disequilibrium ($R^2 = 0.73$, European population, 1000 Genome Project) with the reproduction-enhancing allele of the antagonistically pleiotropic variant rs13164856, is associated with an increased risk of cardiovascular diseases (33). These examples show that susceptibilities to various diseases underlie the antagonistically pleiotropic effects of reproduction-enhancing alleles, consistent with findings in several previous studies (10, 34, 35).

DISCUSSION

Using the U.K. Biobank, we performed a series of trait-level and variant-level analyses to test the antagonistic pleiotropy hypothesis of the evolution of human aging at the genomic scale. At the trait level, we observed a strong negative genetic correlation between reproduction and parental life span as well as that between parental reproduction and parental life span, found that the probability of survival to the age of 76 is negatively correlated with PGSs for reproduction, and detected increases in these PGSs over 25 years. At the variant level, we found that alleles associated with higher reproduction tend to be associated with lower survival to the age of 76 and that frequencies of some of these alleles have increased over the years in a pattern consistent with the action of positive selection. These findings together provide strong genome-wide evidence for the antagonistic pleiotropy hypothesis of aging in humans. That there was ongoing selection for human reproduction as recent as a few decades ago suggests that the optimal genotype for reproduction had not been reached, possibly due to rapid environmental changes. We found that at least some antagonistically pleiotropic variants exhibit horizontal pleiotropy, meaning that they influence life span and reproduction through different pathways (as opposed to affecting life span as a result of affecting reproduction). Analysis

of an eQTL database suggests that the pleiotropic effects may be often realized through cis-regulatory effects in multiple tissues and/or on multiple target genes.

The antagonistic pleiotropy hypothesis of aging posits that at least some mutations antagonistically affect reproduction and life span, without requiring that this antagonistic pleiotropy is enriched. However, relative to comparable mutations selected at random, we found that mutations impacting reproduction were 4.9 times more likely to influence life span and 7.5 times more likely to have that influence reflect antagonistic pleiotropy. In other words, mutational antagonism between reproduction and life span far exceeds that expected for a random pair of traits. In terms of the potential biological reasons behind this phenomenon, it is worth mentioning the disposable soma theory of aging (36), which posits that organisms have limited resources, such that a greater investment in reproduction would lead to a lower investment in DNA repair maintenance, causing accumulation of somatic mutations and aging. The recent report of an inverse relationship between somatic mutation rate per year and life span across 16 mammalian species (37) is consistent with this theory. Although this theory does not specifically invoke pleiotropy, the reproduction–life-span tradeoff arising from resource limitation could be an explanation for antagonistic pleiotropy. Note that some mechanisms mediating the reproduction–life-span antagonistic pleiotropy in our study are probably human specific. For instance, a mutation that negatively affects educational attainment might simultaneously increase reproduction (38) and reduce longevity (39). By contrast, pure environmental factors that antagonistically affect reproduction and longevity (e.g., socioeconomic status) are irrelevant to the antagonistic pleiotropy hypothesis of aging, nor do they confound our genetic analyses such as genetic correlation, PGS calculations, and natural selection.

Given the recently raised concern of a potential publication bias toward empirical evidence for the reproduction-longevity trade-off (40), it is worth discussing the mutation accumulation theory of aging proposed by Medawar (41). This theory asserts that aging is an inevitable consequence of the reduction in the efficacy of natural selection with age, that is, a mutation killing youth will be strongly selected against but a lethal mutation exerting its effect only after reproduction will experience limited or even no negative selection. Over generations, late-acting deleterious mutations will accumulate, leading to an increase in mortality rates late in life. Because U.K. Biobank participants were at least 40 years old at the time of participation, our life-span analysis was largely concentrated on the post-reproductive life span; consequently, our finding of antagonistically pleiotropic variants influencing the life span is consistent with the mutation accumulation theory. Nevertheless, the mutation accumulation theory posits that late-acting deleterious mutations accumulate by genetic drift, while the antagonistic pleiotropy theory asserts that they are selectively favored due to the positive effects on reproduction. We found that the frequencies of reproduction-promoting alleles of some antagonistically pleiotropic variants and the PGSs for reproduction increased over years, providing strong support for the antagonistic pleiotropy theory. This, however, does not reject the mutation accumulation theory because the two theories are not mutually exclusive, that is, the fixations of some aging-promoting alleles may follow the antagonistic pleiotropy theory (i.e., by positive selection), while others may follow the mutation accumulation theory (i.e., by genetic drift).

While our analysis focused on the genetic relationship (i.e., mutational pleiotropy) between reproduction and life span, the phenotypic relationship between reproduction and life span is debated. For example, a negative relationship between female fertility and longevity was reported in Chinese oldest-old individuals (42), while an opposite trend was found in the Amish (43). In our data, given PGSs for reproduction, individuals with two children have a higher probability of survival to 76 than those with 0, 1, or 3 children. Hence, the phenotypic relationship between reproduction and life span is complex and nonmonotonic. Furthermore, the phenotypic correlation does not equal causality because of potential confounding factors such as socioeconomic status.

Human life expectancy, birth rate, and reproductive behavior have all changed markedly in the last few decades (44–46). Specifically, more than half of humans live in areas of the world where birth rates have declined, along with increased incidences of contraception, abortion, and reproductive disorder (47). The global human life expectancy at birth, on the other hand, has steadily increased from 46.5 years in 1950 to 72.8 years in 2019 (48, 49). These trends of phenotypic changes are primarily driven by substantial environmental shifts including changes of lifestyles and technologies and are opposite to the phenotypic changes caused by natural selection of the genetic variants identified in this study. This contrast indicates that, compared with environmental factors, genetic factors play a minor role in the human phenotypic changes studied here. A potential consequence of the environmental shift is that alleles with increased frequencies over time will tend to correlate with a longer life span, which would make our conclusion about antagonistic pleiotropy more conservative. Another potential consequence is that the extended life expectancy exposes the antagonistic pleiotropic effects of genetic variants that did not incur costs in the past when life was, on average, shorter (50).

Our study has several limitations. First, U.K. Biobank participants may be a biased sample of the general population because rates of all-cause mortality and the total cancer incidence are lower in the U.K. Biobank than in the general population (51). Therefore, the survival probability estimated here may be higher than that in the general population, but this factor presumably does not influence the allelic comparisons within the U.K. Biobank. Second, because the life expectancy of our study samples is beyond 76 years (49), the genetic effect on life span is likely underestimated in our study, rendering our conclusion on the negative life-span impact of reproduction-enhancing mutations conservative. However, if life span ≤ 76 years and that >76 years are controlled by largely nonoverlapping loci, our finding would be limited to the former. Third, our observation of allele frequency changes over the six 5-year birth cohorts are consistent with the action of natural selection but does not directly document allele frequency changes between parental and offspring generations because generations are overlapping and only members of the first cohort are likely to be parents of some of the members of the last cohort. Fourth, a substantial fraction of antagonistically pleiotropic variants (23 of 98 = 23.5%) do not have a match in the eQTL database, and we do not know how they influence reproduction and life span. This fraction is on par with the finding that even multiple bulk tissue eQTL data could not explain ~40% of the GWAS-identified variants (21, 31). Future cell type-specific eQTL and other molecular trait (e.g., methylation) data might help elucidate the mechanisms by which these antagonistically pleiotropic variants

act (52). Fifth, our study is solely based on GWAS findings and health records from European-descent individuals. Therefore, one needs to be cautious when generalizing our findings to other populations because of their different social and environmental factors and genetic backgrounds. Future validation of our findings in other populations, especially those currently underrepresented in biobanks (53), is highly desired.

MATERIALS AND METHODS

Study participants and data used

The U.K. Biobank data comprise about 0.5 million participants aged 40 to 70 years, recruited between 2006 and 2010 in 22 assessment centers throughout the United Kingdom and followed up for a variety of health conditions from their recruitment date to 17 February 2016 or their date of death (19). Participants provided a blood sample, from which DNA was extracted and genotyped using the U.K. BiLEVE Axiom array or Affymetrix Axiom array (19). We used the imputed genotypes available from the U.K. Biobank; full details can be found in the U.K. Biobank imputation document (http://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/impute_ukb_v1.pdf, accessed on 16 April 2022). The current study was approved by the U.K. Biobank (reference no. 48678), and the analyses presented were based on data from 488,377 individuals accessed through the U.K. Biobank (www.ukbiobank.ac.uk) on 12 February 2022.

From the entire set of 488,377 individuals with genotype information, we removed individuals with any of the following conditions to prevent population stratification: self-report of non-white British ethnicity, genetic principal components indicative of non-European ancestry, at least one relative in the data identified by genetic kinship, outlying level of genetic heterozygosity, and withdrawal of informed consent. Our final analysis included 276,406 unrelated individuals of European ancestry.

Reproduction-associated variants

We collected reproduction-associated variants from National Human Genome Research Institute GWAS catalog at genome-wide significance ($P \leq 5 \times 10^{-8}$) by searching three types of reproductive traits: sexual maturation, reproductive behavior, and infertility risk (full lists in data S7). GWASs performed or replicated in European populations were included, resulting in 891 unique variants that are associated with at least one reproductive trait.

We then tested linkage disequilibrium among these 891 variants in the European population (1000 Genomes Project, phase 3) by LDlink (54). For two variants with high linkage disequilibrium ($R^2 > 0.8$), the one with the relatively low P value (more significant in GWAS) was retained, while the other was discarded. The variants with minor allele frequencies < 0.05 were removed. The final dataset consisted of 583 genetic variants associated with six reproductive traits: (i) nAMC, (ii) AMP, (iii) nAFS, (iv) nAFB, (v) NEB, and (vi) PCOS (data S1). nAMC, nAFS, and nAFB measure the onset of reproduction, while AMP measures the end of reproduction, so together they inform the timing and potential amount of reproduction. NEB measures the actual number of children. PCOS is one of the most common causes of infertility with a relatively large genetic component (55). By coincidence, each of these 583 variants is associated with only one of the above six reproductive traits. Hence, the direction of the allelic effect of each of these variants on reproduction was unambiguously determined.

Estimation of survival probabilities

We estimated the survival probability based on Cox proportional hazard models (56). The death records in the U.K. Biobank were updated quarterly with the U.K. National Health Service (NHS) Information Centre for participants from England and Wales and with NHS Central Register, Scotland for participants from Scotland. The latest date of death among all registered deaths in the downloaded data is 31 October 2018, and we used this date to approximate the time of last death entry and assumed that we have no mortality or viability information for the volunteers after this date. Specifically, we used five entries—age at recruitment, date of recruitment, year of birth, month of birth, and age at death—to calculate the number of individuals (N_i) who are ascertained from age i to age $i + 1$, and the occurrence of death observed (O_i) from these N_i individuals during the interval of age i to age $i + 1$. Using this information, we calculated the ascertained age for each individual. The death rate per year is then calculated as $h_i = O_i/N_i$, and the probability of survival to age i from age 40 is $S_i = \prod_{n=40}^{i-1} (1 - h_n)$. The U.K. Biobank data allow estimation of h_{40} , h_{41} , ..., and h_{75} with $N_i > 800$. We estimated h_i for the homozygotes of each variant.

Allele frequency changes over years and controls of the age effect and genetic drift effect

We computed the frequency of the allele beneficial to reproduction at each relevant variant in each birth cohort using PLINK v2 (57). Because different birth cohorts had different ages at the time of recruitment by the U.K. Biobank (year 2006–2010; ~66 years old for the earliest birth cohort and ~41 years old for the latest cohort) and because some alleles affect longevity, a fair comparison of allele frequencies across birth cohorts require an equal age of participants of different cohorts. To this end, we used the probability of survival to correct the allele frequency to the age of 41 for each birth cohort. For example, let $N_{A/A}$ be the number of individuals with genotype A/A in the birth cohort of 1940–1944 (age ~ 66 at the U.K. Biobank recruitment) and let $P_{66A/A}$ be the probability of survival to age 66 from 41 for the genotype A/A. The number of A/A individuals at age 41 was inferred to be $N_{A/A}$ divided by $P_{66A/A}$, and the allele frequency was then calculated using corrected numbers of individuals of all genotypes. For each considered variant, allele frequency in a birth cohort was linearly regressed against the mid-point of the birth years of the cohort (1942.5, 1947.5, ..., 1967.5), as described previously (27). P -values from linear regressions were corrected for multiple testing by the Benjamini-Hochberg procedure. To consider the effect of genetic drift on allele frequency changes (28), we used a robustness test by further dividing individuals into 15 2-year birth cohorts (1940–1941, 1942–1943, ..., 1966–1967, and 1968–1969). The allele frequency difference between the first (1940–1941) and second (1942–1943) cohort is denoted as d_1 , and the difference between the third (1944–1945) and fourth (1946–1947) cohort is denoted as d_2 , and so on. This resulted in seven independently estimated d values. We considered a variant to have changed its allele frequency consistently over the years when all seven d values are of the same sign because this event has a chance probability of only $0.5^6 = 1/64 = 0.016$.

Polygenic scores

We computed the PGSs for four reproductive traits (nAFB, nAFS, AMP, and nAMC) each with >50 associated variants. All included variants were independent from one another in contributing to the

trait of concern and reached genome-wide significance (see above). PLINK v2 (57) was used to calculate the PGS of each individual with the following formula

$$\text{PGS}_j = \frac{\sum_i^N S_i G_{ij}}{2M_j}$$

where G_{ij} is the number of reproduction-enhancing alleles at variant i in individual j , S_i is the effect size (beta) of variant i , N is the number of variants included in the calculation, and M_j is the number of nonmissing variants observed in individual j .

Genetic correlation survey

We surveyed the genetic correlations computed for the significantly heritable phenotypes in the U.K. Biobank that are available at www.nealelab.is/blog/2019/10/10/genetic-correlation-results-for-heritable-phenotypes-in-the-uk-biobank. Methodological details have been published (24).

eQTL survey

We surveyed uniformly processed eQTLs across 69 distinct cell/tissue types from 21 available public studies aggregated by eQTL catalog (22). Significant eQTLs were defined using the empirical genome-wide significance threshold as previously described (21, 22). Antagonistically pleiotropic variants were nominated with cis-regulatory activities if they displayed significant eQTL signals for at least one of the significant target genes (eGenes). In addition, we defined three types of cis-regulatory events: (i) multicontext events: significant eQTL signals in more than one tissue/cell type; (ii) multi-eGene events: significant eQTL signals for more than one eGene; and (iii) discordant events: discordant directions of allelic effects for a single eGene across tissues/cell types. The nominated eGenes from antagonistically pleiotropic variants were subjected to pathway enrichment analysis by Ingenuity Pathway Analysis. Unique pathways enriched in antagonistically pleiotropic variants were defined by comparing with the enriched pathways from eGenes of reproduction-associated variants with no significant effects on life span (data S6).

Supplementary Materials

This PDF file includes:

Fig. S1

Legends for data S1 to S7

Other Supplementary Material for this manuscript includes the following:

Data S1 to S7

REFERENCES AND NOTES

1. L. Partridge, N. H. Barton, Optimally, mutation and the evolution of ageing. *Nature* **362**, 305–311 (1993).
2. M. Reichard, Evolutionary perspectives on ageing. *Semin. Cell Dev. Biol.* **70**, 99–107 (2017).
3. G. C. Williams, Pleiotropy, natural selection, and the evolution of senescence. *Evolution* **11**, 398–411 (1957).
4. S. C. Stearns, Trade-offs in life-history evolution. *Funct. Ecol.* **3**, 259–268 (1989).
5. S. C. *Stephen The Evolution of Life Histories* (Oxford Univ. Press, 1992).
6. S. N. Austad, J. M. Hoffman, Is antagonistic pleiotropy ubiquitous in aging biology? *Evol. Med. Public Health* **2018**, **2018**, 287–294

7. J. R. Delaney, C. J. Murakami, B. Olsen, B. K. Kennedy, M. Kaerberlein, Quantitative evidence for early life fitness defects from 32 longevity-associated alleles in yeast. *Cell Cycle* **10**, 156–165 (2011).
8. J. Chen, D. Senturk, J. L. Wang, H. G. Muller, J. R. Carey, H. Caswell, E. P. Caswell-Chen, A demographic analysis of the fitness cost of extended longevity in *Caenorhabditis elegans*. *J. Gerontol. A Biol. Sci. Med. Sci.* **62**, 126–135 (2007).
9. D. Wu, Z. Wang, J. Huang, L. Huang, S. Zhang, R. Zhao, W. Li, D. Chen, G. Ou, An antagonistic pleiotropic gene regulates the reproduction and longevity tradeoff. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2120311119 (2022).
10. S. G. Byars, Q. Q. Huang, L. A. Gray, A. Bakshi, S. Ripatti, G. Abraham, S. C. Stearns, M. Inouye, Genetic loci associated with coronary artery disease harbor evidence of selection and antagonistic pleiotropy. *PLoS Genet.* **13**, e1006328 (2017).
11. J. A. Rodríguez, U. M. Marigorta, D. A. Hughes, N. Spataro, E. Bosch, A. Navarro, Antagonistic pleiotropy and mutation accumulation influence human senescence and disease. *Nat. Ecol. Evol.* **1**, 0055 (2017).
12. X. Wang, S. G. Byars, S. C. Stearns, Genetic links between post-reproductive lifespan and family size in Framingham. *Evol. Med. Public Health* **2013**, **2013**, 241–253
13. S. G. Byars, K. Voskarides, Genes that improved fitness also cost modern humans: Evidence for genes with antagonistic effects on longevity and disease. *Evol. Med. Public Health* **2019**, 4–6 (2019).
14. S. N. Austad, J. M. Hoffman, Response to genes that improved fitness also cost modern humans: Evidence for genes with antagonistic effects on longevity and disease. *Evol. Med. Public Health* **2019**, 7–8 (2019).
15. E. Long, J. Zhang, Retesting the influences of mutation accumulation and antagonistic pleiotropy on human senescence and disease. *Nat. Ecol. Evol.* **3**, 992–993 (2019).
16. J. A. Rodríguez, X. Farré, G. Muntané, U. M. Marigorta, D. A. Hughes, N. Spataro, E. Bosch, A. Navarro, Reply to: Retesting the influences of mutation accumulation and antagonistic pleiotropy on human senescence and disease. *Nat. Ecol. Evol.* **3**, 994–995 (2019).
17. M. C. Mills, F. C. Tropf, D. M. Brazel, N. van Zuydam, A. Vaez; eQTLGen Consortium; BIOS Consortium (Biobank-based Integrative Omics Study); Human Reproductive Behaviour Consortium, T. H. Pers, H. Snieder, J. R. B. Perry, K. K. Ong, M. den Hoed, N. Barban, F. R. Day, Identification of 371 genetic variants for age at first sex and birth linked to externalising behaviour. *Nat. Hum. Behav.* **5**, 1717–1730 (2021).
18. N. Barban, R. Jansen, R. de Vlaming, A. Vaez, J. J. Mandemakers, F. C. Tropf, X. Shen, J. F. Wilson, D. I. Chasman, I. M. Nolte, V. Tragante, S. W. van der Laan, J. R. B. Perry, A. Kong; BIOS Consortium, T. Ahluwalia, E. Albrecht, L. Yerges-Armstrong, G. Atzmom, K. Auro, K. Ayers, A. Bakshi, D. Ben-Avraham, K. Berger, A. Bergman, L. Bertram, L. F. Bielak, G. Björnsdóttir, M. J. Bonder, L. Broer, M. Bui, C. Barbieri, A. Cavadinio, J. E. Chavarró, C. Turman, M. P. Concas, H. J. Cordell, G. Davies, P. Eibich, N. Eriksson, T. Esko, J. Eriksson, F. Falahi, J. F. Felix, M. A. Fontana, L. Franke, I. Gandin, A. J. Gaskins, C. Gieger, E. P. Gunderson, X. Guo, C. Hayward, C. He, E. Hofer, H. Huang, P. K. Joshi, S. Kanoni, R. N. Laan, S. Kiechl, A. Kifley, A. Kluttig, P. Kraft, V. Lagou, C. Lecoeur, J. Lahti, R. Li-Gao, P. A. Lind, T. Liu, E. Makalic, C. Mamasoula, L. Matteson, H. Mbarek, P. F. Mc Ardle, G. M. Mahon, S. F. W. Meddens, E. Mihailov, M. Miller, S. A. Missmer, C. Monneraue, P. J. van der Most, R. Myhre, M. A. Nalls, T. Nutile, K. I. Panagiotou, E. Porcu, I. Prokopenko, K. B. Raj, J. Rich-Edwards, C. A. Rietveld, A. Robino, L. M. Rose, R. Rueedi, K. Ryan, Y. Saba, D. Schmidt, J. A. Smith, L. Stolk, E. Streeten, A. Tonjes, G. Thorleifsson, S. Ulivi, J. Wedenoja, J. Wellmann, P. Willeit, J. Yao, L. Yengo, J. H. Zhao, W. Zhao, D. V. Zernakova, N. Amin, H. Andrews, B. Balkau, N. Barzilay, S. Bergmann, G. Biino, H. Bisgaard, K. Bonnelykke, D. I. Boomsma, E. Buring, H. Campbell, S. Cappellani, M. Ciullo, S. R. Cox, F. Cucca, T. Daniela, G. Davey-Smith, I. J. Deary, G. Dedoussis, P. Deloukas, C. M. van Duijn, E. J. C. de Geus, J. G. Eriksson, D. A. Evans, J. D. Faul, S. C. Felicitá, P. Froguel, P. Gasparini, G. Girotto, H.-J. Grabe, K. H. Greiser, P. J. F. Groenen, H. G. de Haan, J. Haerting, T. B. Harris, A. C. Heath, K. Heikkilä, A. Hofman, G. Homuth, E. G. Holliday, J. Hopper, E. Hyppönen, B. Jacobsson, V. W. J. Jaddoe, M. Johannesson, A. Jugessur, M. Kähönen, E. Kajantie, S. L. R. Kardia, B. Keavney, I. Kolcic, P. Koponen, P. Kovacs, F. Kronenberg, Z. Kutalik, M. L. Bianca, G. Lachance, W. Iacono, S. Lai, T. Lehtimäki, D. C. Liewald, L. C. Study, C. Lindgren, Y. Liu, R. Luben, M. Lucht, R. Luoto, P. Magnus, P. K. E. Magnusson, N. G. Martin, M. M. Gue, R. M. Quillan, S. E. Medland, C. Meisinger, D. Mellström, A. Metspalu, T. Michela, L. Milani, P. Mitchell, G. W. Montgomery, D. Mook-Kanamori, R. de Mutsert, E. A. Nohr, C. Ohlsson, J. Olsen, K. K. Ong, L. Paternoster, A. Pattie, B. W. J. H. Penninx, M. Perola, P. A. Peyser, M. Pirastu, O. Polasek, C. Power, J. Kaprio, L. J. Raffel, K. Rääkkönen, O. Raitakari, P. M. Ridker, S. M. Ring, K. Roll, I. Rudan, D. Ruggiero, D. Rujescu, V. Salomaa, D. Schlessinger, H. Schmidt, R. Schmidt, N. Schupf, J. Smit, R. Sorice, T. D. Spector, J. M. Starr, D. Stöckl, K. Strauch, M. Stumvoll, M. A. Swertz, U. Thorsteinsdóttir, A. R. Thurik, N. J. Timpson, A. Tönjes, J. Y. Tung, A. G. Uitterlinden, S. Vaccargiu, J. Viikari, V. Vitart, H. Völzke, P. Vollenweider, D. Vuckovic, J. Waage, G. G. Wagner, J. J. Wang, N. J. Wareham, D. R. Weir, G. Willemsen, J. Willeit, A. F. Wright, K. T. Zondervan, K. Stefansson, R. F. Krueger, J. J. Lee, D. J. Benjamin, D. Cesarini, P. D. Koellinger, M. den Hoed, H. Snieder, M. C. Mills, Genome-wide analysis identifies 12 loci influencing human reproductive behavior. *Nat. Genet.* **48**, 1462–1472 (2016).
19. C. Bycroft, C. Freeman, D. Petkova, G. Band, L. T. Elliott, K. Sharp, A. Motyer, D. Vukcevic, O. Delaneau, J. O'Connell, A. Cortes, S. Welsh, A. Young, M. Effingham, G. M. Vean, S. Leslie, N. Allen, P. Donnelly, J. Marchini, The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
20. J. E. Harrison, S. Weber, R. Jakob, C. G. Chute, ICD-11: An international classification of diseases for the twenty-first century. *BMC Med. Inform. Decis. Mak.* **21**, 206 (2021).
21. The GTEx Consortium, The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).
22. N. Kerimov, J. D. Hayhurst, K. Peikova, J. R. Manning, P. Walter, L. Kolberg, M. Samoviá, M. P. Sakhivel, I. Kuzmin, S. J. Trevanion, T. Burdett, S. Jupp, H. Parkinson, I. Papatheodorou, A. D. Yates, D. R. Zerbino, K. Alasoo, A compendium of uniformly processed human gene expression and splicing quantitative trait loci. *Nat. Genet.* **53**, 1290–1299 (2021).
23. W. van Rheenen, W. J. Peyrot, A. J. Schork, S. H. Lee, N. R. Wray, Genetic correlations of polygenic disease traits: From theory to practice. *Nat. Rev. Genet.* **20**, 567–581 (2019).
24. B. Bulik-Sullivan, H. K. Finucane, V. Anttila, A. Gusev, F. R. Day, P.-R. Loh, R. G. Consortium, P. G. Consortium; Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium, L. Duncan, J. R. B. Perry, N. Patterson, E. B. Robinson, M. J. Daly, A. L. Price, B. M. Neale, An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
25. R. Border, G. Athanasiadis, A. Buil, A. J. Schork, N. Cai, A. I. Young, T. Werge, J. Flint, K. S. Kendler, S. Sankaraman, A. W. Dahl, N. A. Zaitlen, Cross-trait assortative mating is widespread and inflates genetic correlation estimates. *Science* **378**, 754–761 (2022).
26. A. Buniello, J. A. L. Mac Arthur, M. Cerezo, L. W. Harris, J. Hayhurst, C. Malangone, A. M. Mahon, J. Morales, E. Mountjoy, E. Sollis, D. Suveges, O. Vrousou, P. L. Whetzel, R. Amode, J. A. Guillen, H. S. Riat, S. J. Trevanion, P. Hall, H. Junkins, P. Flicek, T. Burdett, L. A. Hindorf, F. Cunningham, H. Parkinson, The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019).
27. E. Long, J. Zhang, Natural selection contributes to the myopia epidemic. *Natl. Sci. Rev.* **8**, nwa175 (2021).
28. A. F. Feder, S. Kryazhimskiy, J. B. Plotkin, Identifying signatures of selection in genetic time series. *Genetics* **196**, 509–522 (2014).
29. P. R. H. J. Timmers, J. F. Wilson, P. K. Joshi, J. Deelen, Multivariate genomic scan implicates novel loci and haem metabolism in human ageing. *Nat. Commun.* **11**, 3570 (2020).
30. K. Y. Sarin, Y. Lin, R. Daneshjou, A. Ziyatdinov, G. Thorleifsson, A. Rubin, L. M. Pardo, W. Wu, P. A. Khavari, A. Uitterlinden, T. Nijsten, A. E. Toland, J. H. Olafsson, B. Sigurgeirsson, K. Thorisdóttir, E. Jorgensen, A. S. Whittemore, P. Kraft, S. N. Stacey, K. Stefansson, M. M. Asgari, J. Han, Genome-wide meta-analysis identifies eight new susceptibility loci for cutaneous squamous cell carcinoma. *Nat. Commun.* **11**, 820 (2020).
31. J. Byun, Y. Han, Y. Li, J. Xia, E. Long, J. Choi, X. Xiao, M. Zhu, W. Zhou, R. Sun, Y. Bossé, Z. Song, A. Schwartz, C. Lusk, T. Rafnar, K. Stefansson, T. Zhang, W. Zhao, R. W. Pettit, Y. Liu, X. Li, H. Zhou, K. M. Walsh, I. Gorlov, O. Gorlova, D. Zhu, S. M. Rosenberg, S. Pinney, J. E. Bailey-Wilson, D. Mandal, M. de Andrade, C. Gaba, J. C. Willey, M. You, M. Anderson, J. K. Wiencke, D. Albanes, S. Lam, A. Tardon, C. Chen, G. Goodman, S. Bojeson, H. Brenner, M. T. Landi, S. J. Chanock, M. Johannsson, T. Muley, A. Risch, H.-E. Wichmann, H. Bickeböller, D. C. Christiani, G. Rennett, S. Arnold, J. K. Field, S. Shete, L. L. Marchand, O. Melander, H. Brunnstrom, G. Liu, A. S. Andrew, L. A. Kiemeny, H. Shen, S. Zienoldiny, K. Grankvist, M. Johannsson, N. Caporaso, A. Cox, Y.-C. Hong, J.-M. Yuan, P. Lazarus, M. B. Schabath, M. C. Aldrich, A. Patel, Q. Lan, N. Rothman, F. Taylor, L. Kachuri, J. S. Witte, L. C. Sakoda, M. Spitz, P. Brennan, X. Lin, J. M. Kay, R. J. Hung, C. I. Amos, Cross-ancestry genome-wide meta-analysis of 61,047 cases and 947,237 controls identifies new susceptibility loci contributing to lung cancer. *Nat. Genet.* **54**, 1167–1177 (2022).
32. I. Tachmazidou, K. Hatzikotoulas, L. Southam, J. Esparza-Gordillo, V. Haberland, J. Zheng, T. Johnson, M. Koprulu, E. Zengini, J. Steinberg, J. M. Wilkinson, S. Bhatnagar, J. D. Hoffman, N. Buchan, D. Suveges; arcOGEN Consortium, L. Yerges-Armstrong, G. D. Smith, T. R. Gaunt, R. A. Scott, L. C. M. Carthy, E. Zeggini, Identification of new therapeutic targets for osteoarthritis through genome-wide analyses of UK Biobank data. *Nat. Genet.* **51**, 230–236 (2019).
33. G. Kichaev, G. Bhatia, P.-R. Loh, S. Gazal, K. Burch, M. K. Freund, A. Schoech, B. Pasaniuc, A. L. Price, Leveraging polygenic functional enrichment to improve GWAS power. *Am. J. Hum. Genet.* **104**, 65–75 (2019).
34. K. Voskarides, Combination of 247 genome-wide association studies reveals high cancer risk as a result of evolutionary adaptation. *Mol. Biol. Evol.* **35**, 473–485 (2018).
35. K. Voskarides, K. Koutsofi, M. Pozova, TP53 mutant versus wild-type zebrafish larvae under starvation stress: Larvae can live up to 17 days post-fertilization without food. *Zebrafish* **19**, 49–55 (2022).
36. T. B. Kirkwood, Evolution of ageing. *Nature* **270**, 301–304 (1977).
37. A. Cagan, A. Baez-Ortega, N. Brzozowska, F. Abascal, T. H. H. Coorens, M. A. Sanders, A. R. J. Lawson, L. M. R. Harvey, S. Bhosle, D. Jones, R. E. Alcantara, T. M. Butler, Y. Hooks, K. Roberts, E. Anderson, S. Lunn, E. Flach, S. Spiro, I. Januszczak, E. Wrigglesworth,

- H. Jenkins, T. Dallas, N. Masters, M. W. Perkins, R. Deaville, M. Druce, R. Bogeska, M. D. Milsom, B. Neumann, F. Gorman, F. Constantino-Casas, L. Peachey, D. Bochynska, E. St. John Smith, M. Gerstung, P. J. Campbell, E. P. Murchison, M. R. Stratton, I. Martincorena, Somatic mutation rates scale with lifespan across mammals. *Nature* **604**, 517–524 (2022).
38. A. Kong, M. L. Frigge, G. Thorleifsson, H. Stefansson, A. I. Young, F. Zink, G. A. Jonsdottir, A. Okbay, P. Sulem, G. Masson, D. F. Gudbjartsson, A. Helgason, G. Bjornsdottir, U. Thorsteinsdottir, K. Stefansson, Selection against variants in the genome associated with educational attainment. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E727–E732 (2017).
39. I. Sasson, Trends in life expectancy and lifespan variation by educational attainment: United States, 1990–2010. *Demography* **53**, 269–293 (2016).
40. A. A. Cohen, C. F. D. Coste, X.-Y. Li, S. Bourg, S. Pavard, Are trade-offs really the key drivers of ageing and life span? *Funct. Ecol.* **34**, 153–166 (2020).
41. P. B. Medawar, An unsolved problem of biology. *Med. J. Aust.* **1**, 854–855 (1953).
42. Q. Zhu, S. Fu, Q. Zhang, J. Tian, Y. Zhao, Y. Yao, Female fertility has a negative relationship with longevity in Chinese oldest-old population: A cross-sectional study. *Front. Endocrinol.* **11**, 616207 (2020).
43. P. F. M. Ardlie, T. I. Pollin, J. R. O’Connell, J. D. Sorkin, R. Agarwala, A. A. Schäffer, E. A. Streeten, T. M. King, A. R. Shuldiner, B. D. Mitchell, Does having children extend life span? A genealogical study of parity and longevity in the Amish. *J. Gerontol. A Biol. Sci. Med. Sci.* **61**, 190–195 (2006).
44. N. E. Skakkebaek, R. Lindahl-Jacobsen, H. Levine, A.-M. Andersson, N. Jørgensen, K. M. Main, Ø. Lidegaard, L. Priskorn, S. A. Holmboe, E. V. Bräuner, K. Almstrup, L. R. Franca, A. Znaor, A. Kortenkamp, R. J. Hart, A. Juul, Environmental factors in declining human fertility. *Nat. Rev. Endocrinol.* **18**, 139–157 (2022).
45. E. Long, J. Zhang, The coupon collection behavior in human reproduction. *Curr. Biol.* **30**, 3856–3861.e1 (2020).
46. J. W. Vaupel, F. Villavicencio, M.-P. Bergeron-Boucher, Demographic perspectives on the rise of longevity. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2019536118 (2021).
47. S. E. Vollset, E. Goren, C.-W. Yuan, J. Cao, A. E. Smith, T. Hsiao, C. Bisignano, G. S. Azhar, E. Castro, J. Chalek, A. J. Dolgert, T. Frank, K. Fukutaki, S. I. Hay, R. Lozano, A. H. Mokdad, V. Nandakumar, M. Pierce, M. Pletcher, T. Robalikh, K. M. Steuben, H. Y. Wunrow, B. S. Zlavog, C. J. L. Murray, Fertility, mortality, migration, and population scenarios for 195 countries and territories from 2017 to 2100: A forecasting analysis for the Global Burden of Disease Study. *Lancet* **396**, 1285–1306 (2020).
48. GBD 2017 DALYs; HALE Collaborators, Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1859–1922 (2018).
49. M. Roser, E. Ortiz-Ospina, H. Ritchie, *Life Expectancy* (Our World in Data, 2013); <https://ourworldindata.org/life-expectancy> (2013).
50. S. Corbett, A. Courtiol, V. Lummaa, J. Moorad, S. Stearns, The transition to modernity and chronic disease: Mismatch and natural selection. *Nat. Rev. Genet.* **19**, 419–430 (2018).
51. A. Fry, T. J. Littlejohns, C. Sudlow, N. Doherty, L. Adamska, T. Sprosen, R. Collins, N. E. Allen, Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am. J. Epidemiol.* **186**, 1026–1034 (2017).
52. J. L. Min, G. Hemani, E. Hannon, K. F. Dekkers, J. Castillo-Fernandez, R. Luijk, E. Carnero-Montoro, D. J. Lawson, K. Burrows, M. Suderman, A. D. Bretherick, T. G. Richardson, J. Klughammer, V. Iotchkova, G. Sharp, A. A. Khleifat, A. Shatunov, A. Iacoangeli, W. L. M. Ardlie, K. M. Ho, A. Kumar, C. Söderhäll, C. Soriano-Tárraga, E. Giralte-Steinhauer, N. Kazmi, D. Mason, A. F. M. Rae, D. L. Corcoran, K. Sugden, S. Kasela, A. Cardona, F. R. Day, G. Cugliari, C. Viberti, S. Guarrera, M. Lerro, R. Gupta, S. Bollepalli, P. Mandaviya, Y. Zeng, T.-K. Clarke, R. M. Walker, V. Schmoll, D. Czamara, C. Ruiz-Arenas, F. I. Rezwan, R. E. Marioni, T. Lin, Y. Awaloff, M. Germain, D. Aissi, R. Zwamborn, K. van Eijk, A. Dekker, J. van Dongen, J.-J. Hottenga, G. Willemsen, C.-J. Xu, G. Barturen, F. Català-Moll, M. Kerick, C. Wang, P. Melton, H. R. Elliott, J. Shin, M. Bernard, I. Yet, M. Smart, T. Gorrie-Stone, B. I. O. S. Consortium, C. Shaw, A. A. Chalabi, S. M. Ring, G. Pershagen, E. Melén, J. Jiménez-Conde, J. Roquer, D. A. Lawlor, J. Wright, N. G. Martin, G. W. Montgomery, T. E. Moffitt, R. Poulton, T. Esko, L. Milani, A. Metspalu, J. R. B. Perry, K. K. Ong, N. J. Wareham, G. Matullo, C. Sacerdote, S. Panico, A. Caspi, L. Arseneault, F. Gagnon, M. Ollikainen, J. Kaprio, J. F. Felix, F. Rivadeneira, H. Tiemeier, M. H. van IJzendoorn, A. G. Uitterlinden, V. W. V. Jaddoe, C. Haley, A. M. M. Intosh, K. L. Evans, A. Murray, K. Räikkönen, J. Lahti, E. A. Nohr, T. I. A. Sørensen, T. Hansen, C. S. Morgen, E. B. Binder, S. Lucae, J. R. Gonzalez, M. Bustamante, J. Sunyer, J. W. Holloway, W. Karmaus, H. Zhang, I. J. Deary, N. R. Wray, J. M. Starr, M. Beekman, D. van Heemst, P. E. Slagboom, P.-E. Morange, D.-A. Trégouët, J. H. Veldink, G. E. Davies, E. J. C. de Geus, D. I. Boomsma, J. M. Vonk, B. Brunekreef, G. H. Koppelman, M. E. Alarcón-Riquelme, R.-C. Huang, C. E. Pennell, J. van Meurs, M. A. Ikram, A. D. Hughes, T. Tillin, N. Chaturvedi, Z. Pausova, T. Paus, T. D. Spector, M. Kumari, L. C. Schalkwyk, P. M. Visscher, G. D. Smith, C. Bock, T. R. Gaunt, J. T. Bell, B. T. Heijmans, J. Mill, C. L. Relton, Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. *Nat. Genet.* **53**, 1311–1321 (2021).
53. E. Long, M. García-Closas, S. J. Chanock, M. C. Camargo, N. E. Banovich, J. Choi, The case for increasing diversity in tissue-based functional genomics datasets to understand human disease susceptibility. *Nat. Commun.* **13**, 2907 (2022).
54. T. A. Myers, S. J. Chanock, M. J. Machiela, LDlinkR: An R package for rapidly calculating linkage disequilibrium statistics in diverse populations. *Front. Genet.* **11**, 157 (2020).
55. K. Mykhalchenko, D. Lizneva, T. Trofimova, W. Walker, L. Suturina, M. P. Diamond, R. Azziz, Genetics of polycystic ovary syndrome. *Expert Rev. Mol. Diagn.* **17**, 723–733 (2017).
56. D. R. Cox, Regression models and life-tables. *J. R. Stat. Soc.* **34**, 187–202 (1972).
57. C. C. Chang, C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, J. J. Lee, Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).

Acknowledgments: We thank D. Jiang and S. Song for valuable comments. E.L. thanks the Bioinformatics Center of Institute of Basic Medical Sciences for computing support. This research has been conducted using the UK Biobank Resource under Application Number 48678. **Funding:** This work was supported in part by the research grant R35GM139484 from the U.S. National Institutes of Health to J.Z. E.L. was supported by the National Natural Science Foundation of China (Excellent Youth Scholars Program, 32300483 and 82090011) and Chinese Academy of Medical Sciences Innovation Fund (2023-I2M-3-010). **Author contributions:** E.L. and J.Z. designed the research. E.L. conducted the research and analyzed the data. E.L. and J.Z. wrote the paper. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** We performed our analyses using the following publicly available software packages: LDlink (<https://ldlink.nci.nih.gov/>), IPA (www.qiagen.com/us/products/discovery-and-translational-research/next-generation-sequencing/informatics-and-data/interpretation-content-databases/ingenuity-pathway-analysis), and PLINK version 2.0 (www.cog-genomics.org/plink/2.0/). The following publicly available datasets were used in this work: Genetic correlations in U.K. Biobank (www.nealelab.is/blog/2019/10/10/genetic-correlation-results-for-heritable-phenotypes-in-the-uk-biobank/), GWAS catalog (www.ebi.ac.uk/gwas/), and eQTL catalog (www.ebi.ac.uk/eqtl/). The individual-level genotype and phenotype data were acquired from the U.K. Biobank (www.ukbiobank.ac.uk/about-biobank-uk/) under the license of Project 48678. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 7 March 2023

Accepted 7 November 2023

Published 8 December 2023

10.1126/sciadv.adh4990

Evidence for the role of selection for reproductively advantageous alleles in human aging

Erping Long and Jianzhi Zhang

Sci. Adv. **9** (49), eadh4990. DOI: 10.1126/sciadv.adh4990

View the article online

<https://www.science.org/doi/10.1126/sciadv.adh4990>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Copyright © 2023 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).