Genetic evidence for natural selection in humans in the contemporary United States

Jonathan P. Beauchamp

*Department of Economics, Harvard University, Cambridge, MA 02138

Edited by Kenneth W. Wachter, University of California, Berkeley, CA, and approved May 20, 2016 (received for review January 20, 2016)

Recent findings from molecular genetics now make it possible to test directly for natural selection by analyzing whether genetic variants associated with various phenotypes have been under selection. I leverage these findings to construct polygenic scores that use individuals’ genotypes to predict their body mass index, educational attainment (EA), glucose concentration, height, schizophrenia, total cholesterol, and (in females) age at menarche. I then examine associations between these scores and fitness to test whether natural selection has been occurring. My study sample includes individuals of European ancestry born between 1931 and 1953 who participated in the Health and Retirement Study, a representative study of the US population. My results imply that natural selection has been slowly favoring lower EA in both females and males, and are suggestive that natural selection may have favored a higher age at menarche in females. For EA, my estimates imply a rate of selection of about −1.5 mo of education per generation (which pales in comparison with the increases in EA observed in contemporary times). Although they cannot be projected over more than one generation, my results provide additional evidence that humans are still evolving—albeit slowly, especially compared with the rapid changes that have occurred over the past few generations due to cultural and environmental factors.

natural selection | human evolution | educational attainment | menarche | polygenic scores

Whether natural selection has been operating and still operates in modern humans—and at what rate—has been the subject of much debate. Until recently, it was often held that human evolution had come to an end about 40,000–50,000 y ago (see, e.g., ref. 1). However, new evidence that has been accumulating over the last decade suggests that natural selection has been operating in humans over the past few thousand years (2–4) and that a number of adaptations—such as lactase persistence (5), resistance to malaria (6), and adaptation to high altitude (7)—have occurred relatively recently. It has also been shown that height (HGT) and body mass index (BMI) have been under selection in Europeans (8).

In parallel, a number of recent studies have sought to examine the association between lifetime reproductive success (LRS)—the number of children an individual ever gave birth to or fathered—and various phenotypes in contemporary human populations. In modern populations with low mortality, fitness can be reasonably approximated by LRS (9, 10), notwithstanding some caveats summarized in Discussion. These studies have typically found that natural selection has been operating in contemporary humans (9, 11–14). It has also been shown that there was significant variance in relative fitness in a preindustrial human population, such that there was much potential for natural selection (15).

However, this literature has analyzed the relationship between phenotypes and LRS, and natural selection occurs only when genotypes that are associated with the phenotypes covary with reproductive success. This literature’s conclusions regarding ongoing natural selection are therefore particularly sensitive to assumptions that are needed to estimate the relationship between genotypes and phenotypes and to the inclusion in the analysis of all correlated phenotypes with causal effects on fitness (16, 17). Some of those assumptions have been criticized and debated (e.g., ref. 18), and it has proven challenging to include all relevant correlated phenotypes in analyses of selection in natural populations (17). (The latter point matters because if, for example, phenotypes 1 and 2 are phenotypically but not genetically correlated and if only phenotype 2 is under selection, an analysis based on phenotypic data that does not include phenotype 2 or that fails to account for the lack of genetic correlation may erroneously conclude that phenotype 1 is also under selection.)

Recent advances in molecular genetics now make it possible to look directly at the relationship between LRS and genetic variants associated with various phenotypes, thus eliminating those potential confounds. Here, I examine the association between relative LRS (rLRS)—the ratio of LRS to the mean LRS of individuals of the same gender born in the same years—and genetic variants associated with various phenotypes, for a sample of females and males in the Health and Retirement Study (HRS). Using rLRS instead of LRS as the measure of fitness helps control for the effects of time trends in LRS and makes it possible to interpret my estimates as rates of natural selection (9, 16). (My results are robust to using LRS instead of rLRS as the measure of fitness.)

The phenotypes I analyze are BMI, educational attainment (EA), fasting glucose concentration (GLU), HGT, schizophrenia (SCZ), plasma concentrations of total cholesterol (TC), and age at menarche (AAM; in females). These phenotypes were selected on the basis of previous evidence showing that selection acts on some of them (see, e.g., ref. 9) and because summary statistics [i.e., the estimated effects of the single-nucleotide polymorphisms (SNPs) on the phenotypes] from previous large-scale genome-wide association studies (GWAS) are available for them (19–25).

The HRS is a representative longitudinal panel study of ~20,000 Americans shortly before and during retirement. It is well suited for this study for several reasons. First, the HRS was designed to be representative of the US population over the age of 50 y (26), which makes it possible to generalize my results to the entire US population of European ancestry born in the years of my study sample. In addition, individuals in the study are in the later stages of their lives, when they have typically completed their lifetime reproduction. Nonetheless, as I explain in Discussion, selection

Significance

I leverage recent advances in molecular genetics to test directly whether genetic variants associated with a number of phenotypes have been under natural selection in the contemporary United States. My finding that natural selection has been slowly occurring for genetic variants associated with educational attainment and (suggestively, in females) for variants associated with age at menarche provides additional evidence that humans are still evolving—albeit slowly and at a rate that cannot account for more than a small fraction of the large changes that have occurred over the past few generations.

Author contributions: J.P.B. designed research, performed research, analyzed data, and wrote the paper.

The author declares no conflict of interest.

This article is a PNAS Direct Submission.

See Commentary on page 7693.

Email: jonathan.pierre.beauchamp@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1600398113/-/DCSupplemental.
bias due to incomplete genotyping of the study participants and differential survival remains a concern, although genotyped individuals do not appear to differ markedly from nongenotyped individuals in my study sample.

To mitigate the risks of confounding by population stratification, my analyses focus on unrelated individuals of European ancestry and control for the top 20 principal components of the genetic relatedness matrix [which capture the main dimensions along which the ancestry of the individuals in the dataset vary (27)]. To mitigate the risk of selection bias due to differential mortality, and to ensure that the LRS variable is a good proxy for completed fertility, I limit my analyses to individuals born between 1931 and 1953 who were at least 45 y old (for females) or 50 y old (for males) when asked the number of children they ever gave birth to or fathered. I refer to the resulting sample as the “study sample.” I performed my main analyses separately for females and males, as different selection gradients can operate across genders.

In a recent paper, Troup et al. (28) used genetic data to study the relationship between LRS and age at first birth in a sample of females, and found that the two phenotypes are negatively genetically correlated. My analyses complement theirs in several important ways: My analyses cover both females and males and seven different phenotypes; they include childless individuals (who can have an important impact on the gene pool by foregoing reproduction); they translate selected estimates into interpretable measures of the rate at which natural selection has been operating; and they indirectly leverage the statistical power of previous large-scale GWAS to estimate the relationship between rLRS and genetic variants associated with the phenotypes, thus increasing the precision of my estimates for some phenotypes. [An alternative to my score regression approach (described in Genetic Evidence for Natural Selection) is the bivariate GREML method (29) (used by Troup et al.), but it is not well suited for the present study. It has very large error variances for the other estimates of the genetic correlation in samples of moderate size like the HRS, given the low SNP heritability of rLRS (28), according to the GCTA-GREML Power Calculator (30). It also requires a dataset with phenotypic data for every studied phenotype and assumes normally distributed phenotypes (which is not realistic for rLRS).]

Phenotypic Evidence for Natural Selection

I begin by looking at the phenotypic evidence for natural selection in the HRS. The HRS contains phenotypic variables for four of the seven phenotypes I study: BMI, EA, HGT, and TC. (The phenotypic variable for TC is an indicator for a self-reported health problem with high cholesterol, and not plasma concentrations of TC as in the GWAS of TC.) Table S1 reports summary statistics for these and for the other phenotypic variables I use, both for all individuals in the study sample and for the genotyped individuals in the study sample. As can be seen, the two samples look remarkably similar. Table 1 reports estimates from separate regressions of rLRS on each of these phenotypic variables (and on control variables) for the sample of all individuals (genotyped and not genotyped) in the study sample. (The incremental $R^2$ of the score of a phenotype is defined as the difference between the $R^2$ of the regression of the phenotype on controls for sex and birth year, the top 20 principal components of the genetic relatedness matrix, and the score, and the $R^2$ of the same regression but without the score.) The incremental $R^2$ estimates range from 0.012 (for TC) to 0.174 (for HGT) and are all significantly larger than zero; nonetheless, they are all much smaller than estimates of the phenotypes’ heritability in the existing literature (31), implying that the scores are very imperfect proxies for the individuals’ true genetic scores (defined as the sum of the true average causal effects of all their alleles) for the various phenotypes. [The $R^2$ of the polygenic score of a phenotype is bounded by the phenotype’s heritability (32) and depends, in part, on the precision with which the effects of the individual genetic variants were estimated in the GWAS of that phenotype, which, in turn, depends on the GWAS sample size. Future, larger GWAS should allow more precise estimation of the effects of the genetic variants and the construction of more precise scores.]

Table 2 reports estimates from separate regressions of rLRS on the polygenic scores of the various phenotypes [and on control variables, which include the top 20 principal components of the genetic relatedness matrix (27)]. The score of EA is significantly negatively associated with rLRS for both females ($P = 0.002$) and males ($P = 0.013$). The association remains significant after Bonferroni correction for 13 tests (the number of estimates reported in Table 2) for females, but not for males (Bonferroni-corrected $P = 0.022$ for females, $P = 0.174$ for males). The estimates for females and males are similar, and the association is also significant in the sample of females and males together (Table S3; $P = 1.2 \times 10^{-5}$) and remains significant after Bonferroni correction for seven tests (the number of phenotypes) (Bonferroni-corrected $P = 8.3 \times 10^{-5}$). Fig. 2 shows the mean polygenic score of EA as a...
function of LRS, by sex. Both females and males who had no children have a significantly higher mean score of EA than those who had one or more children ($P < 0.005$ in both cases, unpaired $t$ tests). Thus, the negative association between rLRS and the score of EA appears to be driven primarily by score differences between individuals with and without children.

The polygenic score of AAM is also significantly and positively associated with rLRS for females at the 10% level ($P = 0.080$), but this association does not remain significant after Bonferroni correction for 13 tests. I therefore interpret it as being weakly suggestive that genetic variants associated with higher AAM may have been selected for. The polygenic scores of the other phenotypes (BMI, GLU, HGT, SCZ, and TC) are not robustly significantly associated with rLRS. Although these estimates are small in magnitude and insignificant, this could be due to lack of statistical power and because my polygenic scores are imperfect proxies for the true genetic scores, and does not prove that natural selection has not been operating on genetic variants associated with those phenotypes.

According to the Robertson–Price identity (35, 36), the directional selection differential of a “character” is equal to the genetic covariance between the character and relative fitness. (A character is an observable feature of an organism, and its directional selection differential is the change in its mean value due to natural selection in one generation.) As I show in SI Materials and Methods, if we define the polygenic scores as the characters of interest, it follows that the coefficients on the scores reported in Table 2 can be interpreted as the directional selection differentials of the scores themselves, expressed in Haldanes—i.e., each coefficient equals the implied change in the score that will occur due to natural selection in one generation, expressed in SDS of the score per generation. Hence, the estimates from Table 2 imply that natural selection has been operating on the score of EA at a rate of $-0.033$ Haldanes among females and of $-0.031$ Haldanes among males in the study sample. (Even if the mechanism that underlies the negative association between rLRS and the score of EA is that more educated people choose to have fewer children, it would still be the case that natural selection has been operating.) I rescaled these estimates of the directional selection differential of the score of EA to express them in years of education per generation rather than in Haldanes (SI Materials and Methods). My rescaled estimates imply that natural selection has been operating on the score of EA at rates of $-0.022$ [95% confidence interval (CI): $-0.036$ to $-0.009$] and $-0.022$ (95% CI: $-0.040$ to $-0.004$) years of education per generation for females and males—or about $1$ wk of education per generation for both sexes.

I also obtained estimates of the directional selection differential of EA (or, equivalently, of the true genetic score of EA), which is equal (under some assumptions) to the directional selection differential of the polygenic score of EA multiplied by the ratio of the heritability of EA to the $R^2$ of the score of EA (SI Materials and Methods). [I assume the heritability of EA to be 0.40, based on a recent meta-analysis of existing heritability estimates of EA (37).] Generalizing my results from the study sample to the general population, my estimates imply that natural selection has been operating on EA at rates of $-1.30$ (95% CI: $-2.12$ to $-0.54$) and $-1.53$ (95% CI: $-2.85$ to $-0.31$) months of education per generation among US females and males of European ancestry born between 1931 and 1953. As I discuss in Discussion, these rates are small relative to the increases in EA that have been observed over the past few generations.

I performed a number of checks to verify the robustness of my results. First, I repeated the analyses with LRS instead of rLRS. Second, I used polygenic scores constructed with PLINK (34) instead of LDpred. Third, I only included individuals aged no more than 70 y in 2008 (the last year in which individuals were genotyped) and at least 50 y old (for females) or 55 y old (for males) when asked their number of children—to mitigate the risk of selection bias due to differential mortality and to ensure that almost every individual had completed fertility when asked his or her number of children. Fourth, I included the HRS0 cohort of individuals born between 1924 and 1930 together with the study sample. (As detailed in Materials and Methods, I define cohorts based on the individuals’ birth years; the study sample includes the HRS1, HRS2, and HRS3 cohorts, but excludes the HRS0 cohort because of possible selection checks. In all cases, the results for EA are robust. Further, for the score of EA for both females and males and for the score of AAM (for females), the estimates are not significantly different from one another at the 5% level across the HRS1, HRS2, and HRS3 cohorts (Table S3 and $t$ tests of the interactions between the coefficients on the scores and cohort dummies).

Following Lande and Arnold (16), I also estimated quadratic regressions of rLRS on all of the polygenic scores and their squares and interactions together and on control variables, to test for nonlinear selection (Table S5 and SI Materials and Methods). I found no convincing evidence that nonlinear selection has been operating on the genetic variants associated with the various phenotypes.

![Fig. 1. Predictive power ($R^2$) of polygenic scores of the various phenotypes. “Previously reported” denotes the highest previously reported $R^2$ of scores from prediction analyses from the articles reporting the GWAS whose summary statistics were used to construct the scores of BMI (19), EA (20), HGT (22), and SCZ (23) [the $R^2$ of the score was not reported for GLU, AAM, and TC]. The previously reported $R^2$ for SCZ is the $R^2$ on the liability scale. “Estimated in HRS” denotes estimates of the incremental $R^2$ of the LDpred scores used in this article and for which there are phenotypic variables in the HRS, with percentile confidence intervals estimated with the nonparametric bootstrap with 1,000 bootstrap samples.](https://www.pnas.org/doi/10.1073/pnas.1600398113)

**Table 2. Estimates from separate regressions of rLRS on the polygenic score of each phenotype**

| Score     | Females    | Males    |
|-----------|------------|----------|
| Score of BMI | 0.006 (0.010) | 0.016 (0.013) |
| Score of EA  | $-0.033**$ (0.010) | $-0.031**$ (0.012) |
| Score of GLU | 0.009 (0.010) | $-0.013$ (0.013) |
| Score of HGT | $-0.011$ (0.014) | $-0.005$ (0.018) |
| Score of SCZ | $-0.001$ (0.011) | 0.009 (0.013) |
| Score of TC  | $-0.012$ (0.011) | $-0.003$ (0.013) |
| Score of AAM | 0.018* (0.011) | —       |
| $N$        | 3,416      | 2,571    |

This table shows estimates of the coefficients on the polygenic scores and their SEs (in parentheses) from separate regressions of rLRS on the polygenic score of each phenotype and on control variables, for the study sample. All regressions for each sex had the same number of observations. ($^{*}P < 0.10; ~^{**}P < 0.05; ~^{***}P < 0.01.$)
Fig. 2. Mean polygenic score of EA as a function of LRS, for females and males in the study sample.

Discussion
My results suggest that natural selection has been operating on the genetic variants associated with EA, and possibly with AAM. Although I find no evidence that natural selection has been operating on the genetic variants associated with other phenotypes or that nonlinear selection has been operating. I emphasize that this could be because my polygenic scores are imperfect proxies for the true genetic scores, which limits the statistical power of my analyses.

My estimates of the negative associations between rLRS and both phenotypic EA and the polygenic score of EA are consistent with previous findings of negative associations between LRS and phenotypic EA in samples of females (38-41), males (38, 41), and females and males together (42) in contemporary Western populations, although positive phenotypic associations have also been reported for males (39). My estimates are also consistent with concurrent findings of a negative phenotypic association for females and males together and of a negative correlation between LRS and a score of EA (constructed with the summary statistics from an earlier, smaller GWAS of EA) (43). To my knowledge, few articles have investigated the relationship between phenotypic AAM and LRS in contemporary Western populations. Kirk et al. (44) find a quadratic relationship that is suggestive of stabilizing selection, but they find no genetic covariation (using behavioral genetic techniques in a sample of twins). Consistent with the results of my regressions of rLRS on phenotypic BMI and HGT (Table 1), there is previous phenotypic evidence of positive selection for weight and negative selection for HGT in females (9, 45). Previous studies have also documented a positive (9) and an inverted-U (45) relationship between LRS and phenotypic HGT in males, and have found evidence of negative selection for TC and of stabilizing selection for GLU in females (11), in contemporary Western human populations.

Consistent with the results from previous studies with phenotypic data (e.g., ref. 9), my results suggest that natural selection has been operating slowly relative to the rapid changes that have occurred over the past few generations, presumably due to cultural and environmental factors. For instance, my estimate of a directional selection differential of EA of about −1.5 mo of education per generation pales in comparison with the increase of 6.2 y in the mean level of EA that took place for native-born Americans born between 1876 and 1951 (46) (which is equivalent to about 2 y of education per generation). Moreover, although I find suggestive evidence that genetic variants associated with higher AAM may have been selected for, AAM has substantially decreased in contemporary Western populations (47). Also, although I find no evidence of selection for the genetic variants associated with BMI and HGT, both phenotypes have markedly increased over the past century (48). Thus, although natural selection is still operating, the environment appears to have achieved an “evolutionary override” (28) on the measurable phenotypes I study.

As shown in Okbay et al. (20), the association between the score of EA and EA is not likely to be driven by the effects of culture, the environment, or population stratification, and is likely to reflect the true causal effects of multiple genetic variants. For instance, in cohorts that are independent of those used in the GWAS of EA, the score remains significant in regressions of EA on the score when family fixed effects are also included. Moreover, estimates from an LD score regression (49)—which disentangles the signal due to the genetic variants’ causal effects from the signal due to confounding biases—suggest that stratification is not a major source of bias in the GWAS summary statistics of EA. Okbay et al. also analyzed the summary statistics of EA and obtained sizeable and significant estimates of the genetic correlation between EA and several neuropsychiatric and cognitive phenotypes, as well as of the genetic variance of EA accounted for by SNPs annotated to the central nervous system relative to other SNPs. Thus, although it is not possible to rule out with certainty that my results are (at least partly) confounded by stratification, stratification is unlikely to be an important concern.

Several additional caveats should be kept in mind when interpreting my results. First, LRS is not a perfect proxy for individual genetic contribution. Among other possible reasons for this, a tradeoff between the quantity and quality of children has been documented in preindustrial human societies and may still exist in modern societies (50). In the presence of such a tradeoff, the number of grandchildren or third-generation descendants is a better measure of fitness—although most datasets (including the HRS) lack such data, and it has been shown that LRS and number of grandoffspring were perfectly genetically correlated in a postindustrial human population (10). Also, in growing populations, individuals who successfully produce earlier in life tend to have higher fitness (51), but rLRS does not account for fertility timing. In the case of EA, individuals with high EA typically have children at a more advanced age, which may further reduce their fitness. Alternative measures of fitness—such as the intrinsic rate of increase (the exponentiated Lotka’s r)—account for fertility timing, but they require data on the age at birth of every offspring and do not always perform better in natural populations (52). A second caveat is that it is not possible to translate my estimates into projected evolutionary changes over more than one generation, because my results do not account for the effects of all phenotypes studied to correlate genetically with the phenotypes I study and that also have causal effects on fitness (16). Selection on phenotypes that are genetically correlated with the phenotypes of interest impacts their genetic covariance, which, in turn, impacts the selection gradients on the phenotypes of interest in future generations. Furthermore, because the cultural environment changes through time, the selection gradients that existed from 1931 to 1953 may not apply to earlier and subsequent periods, which makes long-term projections problematic. For instance, it has been shown that the demographic transition has significantly changed the selective forces in some populations (41, 53-55).

Lastly, there are several reasons why my results in the study sample of genotyped individuals might not be fully generalizable to the entire US population of European ancestry born between 1931 and 1953. First, the HRS only targets individuals who survived until age 50 y, and about 10% of female and 15% of male Americans born in 1940 died before reaching age 50 y, based on data from the US Social Security Administration (56). Second, in the study sample, only 85% of the participants were still alive in 2008 (the last year when they could be genotyped), 69% were asked to be genotyped, and 59% consented to be genotyped. That being said, a comparison of the summary statistics for all individuals in the study sample and for the genotyped individuals in the study sample (Table S1) suggests that there are no important differences between the two samples, and the results of the phenotypic regressions are very similar across the two samples (Table 1 and Table S2).
In sum, and keeping those limitations in mind, my results strongly suggest that genetic variants associated with EA have slowly been selected against among both female and male Americans of European ancestry born between 1931 and 1953, and that natural selection has thus been occurring in that population—albeit at a rate that pales in comparison with the rapid changes that have occurred in recent generations. My results also suggest that genetic variants positively associated with AAM may have been positively selected for among females in that population. As larger GWAS are conducted and better estimates of genetic variants’ effects on various phenotypes become available, polygenic scores will become more precise. The eventual completion of a GWAS of LRS will also make it possible to use other methods in order to more rigorously estimate the genetic covariance between LRS and other phenotypes. Future studies that address the above-mentioned limitations will be able to leverage these developments to replicate my results and to obtain more precise estimates of the rate at which natural selection has been and is occurring in humans.

Materials and Methods
The Study Sample and the Cohorts. The HRS is a longitudinal panel study for which a representative sample of ~20,000 Americans have been surveyed every 2 years since 1992. The main analyses focus on individuals born between 1931 and 1953. To reduce the risks of confounding by population stratification, I restrict the analyses to unrelated individuals of European ancestry (i.e., non-Hispanic White individuals). To ensure that the LRS variable is a good proxy for completed fertility, I only include females who were at least 45 y old and who have not, I only include individuals who were enrolled in the HRS and asked the number of children they ever gave birth to or fathered. Further, to ensure that the sample of individuals who have not, I only include individuals who were enrolled in the HRS and asked the number of children they ever gave birth to or fathered in 2008 or earlier. This left 6,414 females and 5,436 males with phenotypic data and 3,416 unrelated females and 2,571 unrelated males who have been successfully genotyped and who passed the quality control filters described in SI Materials and Methods (for which I could thus construct polygenic scores). I refer to the resulting sample as the “study sample.”

For some specifications, I divided the study sample into three nonoverlapping cohorts based on the individuals’ birth years. This allowed me to test the robustness of my results across cohorts (my definition of the cohorts resembles the definition used by the HRS, which recruited its different cohorts at different times). Table S6 summarizes the three cohorts—which I label HRS1 (birth years 1931–1941), HRS2 (birth years 1942–1947), and HRS3 (birth years 1948–1953)—as well as the HRS0 cohort of individuals born between 1924 and 1930. To mitigate the risk of selection bias based on mortality, I excluded the HRS0 cohort of individuals born between 1924 and 1930 from the study sample (Table S6 and SI Materials and Methods), but my main results are robust to the inclusion of this cohort. For the Rest of the paper “individual” always refers to the respondent (or proxy) who passes the quality control steps and the construction of the polygenic scores, and Table S4 summarizes the three cohorts as well as the sample sizes across the SNPs used to construct the scores and the sources for each phenotype’s summary statistics.

Association Analyses. For each of BMI, EA, HGT, and TC—for which polygenic variables are available in the HRS—I regressed rLRS on the corresponding polygenic variable, separately for females and males; those regressions included birth year dummies and HRS-defined cohort dummies and were estimated by ordinary least squares (OLS). For all phenotypes, I also regressed rLRS on the polygenic score of the phenotype in various samples; the regressions also included birth year dummies, HRS-defined cohort dummies, and the top 20 principal components of the genetic relatedness matrix [to control for population stratification (27); see also section 5 of the supplemental material of ref. 59], and were also estimated by OLS. For the Rest of the paper “variable” always refers to the selected phenotype. I only included the respondent with the lowest person number (PN, an HRS identifier) in each household, as spouses very often have the same number of children, which induces a complex correlation structure between the error terms (the results for the score of EA are robust to alternative ways of selecting one respondent per household). In most results tables, I report the coefficient estimates and SEs, with asterisks to indicate statistical significance. P values for the main results are reported in Table S2. The SEs and P values of my estimates from regressions of rLRS on the LDpred scores do not account for the uncertainty stemming from the selection of the Gaussian mixture weights for the LDpred scores; however, the fact that my main results are robust to the use of the PLINK scores instead of the selected LDpred scores implies that my results are not driven by this weight selection procedure.

Directional Selection Differentials. Based on the Robertson–Price identity (35, 36), the directional selection differential of a character is equal to its genetic covariance with relative fitness. As I show in SI Materials and Methods, it follows that the estimates of the coefficients on the polygenic scores reported in Table 2 can be interpreted as directional selection differentials of the scores, expressed in Haldanes (1 Haldane is 1 SD per generation). SI Materials and Methods also show how to rescale the estimates of the directional selection differential for the score of EA to express them in years of education per generation instead of in Haldanes, and shows how to obtain estimates of the directional selection differential of EA (or, equivalently, of the true genetic score of EA—rather than of the polygenic score of EA) expressed in years of education per generation (under some assumptions). I used the nonparametric bootstrap method with 1,000 bootstrap samples to obtain percentile confidence intervals for the estimates of the directional selection differentials.

Institutional Review of This Project. This project was reviewed and approved by the National Bureau of Economic Research (NBER) Institutional Review Board. The Harvard University Committee on the Use of Human Subjects also reviewed the protocol for this project and determined that it is not human subjects research.

ACKNOWLEDGMENTS. I thank David Cesarini, Joseph Henrich, Lawrence Katz, Iain Mathieson, Steven Pinker, Alkes Price, Stephen Stearns, and Peter

Materials and Methods
The Study Sample and the Cohorts. The HRS is a longitudinal panel study for which a representative sample of ~20,000 Americans have been surveyed every 2 years since 1992. The main analyses focus on individuals born between 1931 and 1953. To reduce the risks of confounding by population stratification, I restrict the analyses to unrelated individuals of European ancestry (i.e., non-Hispanic White individuals). To ensure that the LRS variable is a good proxy for completed fertility, I only include females who were at least 45 y old and who have not, I only include individuals who were enrolled in the HRS and asked the number of children they ever gave birth to or fathered. Further, to ensure that the sample of individuals who have not, I only include individuals who were enrolled in the HRS and asked the number of children they ever gave birth to or fathered in 2008 or earlier. This left 6,414 females and 5,436 males with phenotypic data and 3,416 unrelated females and 2,571 unrelated males who have been successfully genotyped and who passed the quality control filters described in SI Materials and Methods (for which I could thus construct polygenic scores). I refer to the resulting sample as the “study sample.”

For some specifications, I divided the study sample into three nonoverlapping cohorts based on the individuals’ birth years. This allowed me to test the robustness of my results across cohorts (my definition of the cohorts resembles the definition used by the HRS, which recruited its different cohorts at different times). Table S6 summarizes the three cohorts—which I label HRS1 (birth years 1931–1941), HRS2 (birth years 1942–1947), and HRS3 (birth years 1948–1953)—as well as the HRS0 cohort of individuals born between 1924 and 1930. To mitigate the risk of selection bias based on mortality, I excluded the HRS0 cohort of individuals born between 1924 and 1930 from the study sample (Table S6 and SI Materials and Methods), but my main results are robust to the inclusion of this cohort. For the Rest of the paper “individual” always refers to the respondent (or proxy) who passes the quality control steps and the construction of the polygenic scores, and Table S4 summarizes the three cohorts as well as the sample sizes across the SNPs used to construct the scores and the sources for each phenotype’s summary statistics.

Association Analyses. For each of BMI, EA, HGT, and TC—for which polygenic variables are available in the HRS—I regressed rLRS on the corresponding polygenic variable, separately for females and males; those regressions included birth year dummies and HRS-defined cohort dummies and were estimated by ordinary least squares (OLS). For all phenotypes, I also regressed rLRS on the polygenic score of the phenotype in various samples; the regressions also included birth year dummies, HRS-defined cohort dummies, and the top 20 principal components of the genetic relatedness matrix [to control for population stratification (27); see also section 5 of the supplemental material of ref. 59], and were also estimated by OLS. For the Rest of the paper “variable” always refers to the selected phenotype. I only included the respondent with the lowest person number (PN, an HRS identifier) in each household, as spouses very often have the same number of children, which induces a complex correlation structure between the error terms (the results for the score of EA are robust to alternative ways of selecting one respondent per household). In most results tables, I report the coefficient estimates and SEs, with asterisks to indicate statistical significance. P values for the main results are reported in Table S2. The SEs and P values of my estimates from regressions of rLRS on the LDpred scores do not account for the uncertainty stemming from the selection of the Gaussian mixture weights for the LDpred scores; however, the fact that my main results are robust to the use of the PLINK scores instead of the selected LDpred scores implies that my results are not driven by this weight selection procedure.

Directional Selection Differentials. Based on the Robertson–Price identity (35, 36), the directional selection differential of a character is equal to its genetic covariance with relative fitness. As I show in SI Materials and Methods, it follows that the estimates of the coefficients on the polygenic scores reported in Table 2 can be interpreted as directional selection differentials of the scores, expressed in Haldanes (1 Haldane is 1 SD per generation). SI Materials and Methods also show how to rescale the estimates of the directional selection differential for the score of EA to express them in years of education per generation instead of in Haldanes, and shows how to obtain estimates of the directional selection differential of EA (or, equivalently, of the true genetic score of EA—rather than of the polygenic score of EA) expressed in years of education per generation (under some assumptions). I used the nonparametric bootstrap method with 1,000 bootstrap samples to obtain percentile confidence intervals for the estimates of the directional selection differentials.

Institutional Review of This Project. This project was reviewed and approved by the National Bureau of Economic Research (NBER) Institutional Review Board. The Harvard University Committee on the Use of Human Subjects also reviewed the protocol for this project and determined that it is not human subjects research.

ACKNOWLEDGMENTS. I thank David Cesarini, Joseph Henrich, Lawrence Katz, Iain Mathieson, Steven Pinker, Alkes Price, Stephen Stearns, and Peter
Visscher for helpful comments. I also thank Dan Benjamin and David Laibson for helpful comments on the GWAS of EA directed to the SSGAC policy. I acknowledge the remaining authors of that paper in the SSGAC. I thank Ayus Okbay for constructing these scores on behalf of the SSGAC. I contributed to the GWAS of EA reported in Okbay et al. (20); in accordance with SSGAC policy, I acknowledge the remaining authors of that paper in the SI Appendix.

1. Gould SJ (2000) The spice of life. Leader Leader 15(Winter):14–19.
2. Voight BF, Kudaravalli S, Wen X, Pritchard JK (2006) A map of recent positive selection in the human genome. PLoS Biol 4(3):e72.
3. Mathieson I, et al. (2015) Genome-wide patterns of selection in 230 ancient Eurasians. Nature 528(7583):499–501.
4. Pickrell JK, et al. (2009) Signals of recent positive selection in a worldwide sample of human populations. Genome Res 19(5):826–837.
5. Tishkoff SA, et al. (2007) Convergent adaptation of human lactase persistence in Africa and Europe. Nat Genet 39(1):31–40.
6. Kiviatkowsk DP (2005) How malaria has affected the human genome and what human genetics can teach us about malaria. Am J Hum Genet 77(1):171–192.
7. Yi X, et al. (2010) Sequencing of 50 human exomes reveals adaptation to altitude. Science 329(5987):75–77.
8. Turchin MC, et al.; Genetic Investigation of ANthropometric Traits (GIANT) Consortium (2012) Evidence of widespread selection on standing variation in Europe at height-associated SNPs. Nat Genet 44(9):1015–1019.
9. Stearns SC, Byars SG, Govindaraju DR, Ewbank D (2010) Measuring selection in contemporary human populations. Nat Rev Genet 11(9):611–622.
10. Zietsch BP, Kuja-Halkola R, Walum H, Verweij KJH (2014) Perfect genetic correlation between number of offspring and grandoffspring in an industrialized human population. Proc Natl Acad Sci USA 111(13):4912–4916.
11. Byars SG, Ewbank D, Govindaraju DR, Stearns SC (2010) Colloquium papers: Natural selection in a contemporary human population. Proc Natl Acad Sci USA 107(Suppl 1):1787–1792.
12. Milot E, et al. (2011) Evidence for evolution in response to natural selection in a contemporary human population. Proc Natl Acad Sci USA 108(41):17040–17045.
13. Elguero E, et al. (2015) Malaria continues to select for sickle cell trait in Central Africa. Proc Natl Acad Sci USA 112(22):7051–7054.
14. Stupl G, Barrett L, Tropf FC, Mills M (2015) Does natural selection favour taller stature among the tallest people on Earth? Proc Biol Sci 282(1806):20150211.
15. Courtright A, Pettay JE, Jokela M, Rotkirch A, Lummaa V (2012) Natural and sexual selection in a monogamous historical human population. Proc Natl Acad Sci USA 109(21):8044–8049.
16. Lande R, Arnold SJ (1983) The measurement of selection on correlated characters. Evolution 37(6):1210–1226.
17. Morrissey MB, Kruek L, Wilson AJ (2010) The dynamics of applying the breeder’s equation in observational studies of natural populations. J Evol Biol 23(11):2277–2288.
18. Kamin Li, Goldberg AS (2002) Twin studies in behavioral research: A skeptical view. Theor Popul Biol 61(1):83–95.
19. Locke AE, et al.; Lifelines Cohort Study, ADOPGen Consortium; AGEM-BMI Working Group; CARDIOMapplusC4D Consortium; CKDGen Consortium; GLGC; ICBP; MAGIC Investigators; MuTHER Consortium; MiGen Consortium; PAGE Consortium; ReproGen Consortium; GENIE Consortium; International Endogene Consortium (2013) Genetic studies of body mass index yield new insights for obesity biology. Nature 518(7538):197–206.
20. Okbay A, et al. (2016) Genome-wide association study identifies 74 loci associated with educational attainment. Nature 533(7604):539–542.
21. Sonnega A, et al. (2014) Cohort profile: The Health and Retirement Study (HRS). J Biosoc Sci 46(4):542–552.
22. Price AL, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38(8):904–909.
23. Ripke S, et al.; Schizophrenia Working Group of the Psychiatric Genomics Consortium, Discovery, Biology, and Risk of Inherited Variants in Breast Cancer (DRIVE) study (2015) Modeling linkage disequilibrium increases accuracy of polygenic risk scores. Am J Hum Genet 97(4):576–592.
24. Perrin S, Chang C (2015) PLINK 1.9. Available at https://www.cog-genomics.org/plink2.
25. Robertson A (1966) A mathematical model of the culling process in dairy cattle. Animi Prod 8(1):95–108.
26. Price GR (1970) Selection and covariance. Nature 227(5275):520–521.
27. Brangan AR, McCallum KJ, Freese J (2013) Variation in the heritability of educational attainment: An international meta-analysis. Soc Forces 92(1):109–140.
28. Weeden J, Abrams MC, Green MC, Sabini J (2006) Do high-status people really have fewer children? Education, income, and fertility in the contemporary U.S. Hum Nat 17(4):377–392.
29. Fieder H, Huber S (2007) The effects of sex and childlessness on the association between status and reproductive output in modern society. Evol Hum Behav 28(3):392–398.
30. Huber S, Bookstein FL, Fieder M (2010) Socioeconomic status, education, and reproduction in modern women: An evolutionary perspective. Am J Hum Biol 22(5):578–587.
31. Skirbekk V (2008) Fertility trends by social status. Demogr Res 18(5):145–180.
32. Hopscroft RL (2006) Sex, status, and reproductive success in the contemporary United States. Evol Hum Behav 27(2):104–120.
33. Conley D, et al. (2016) Assortative mating and differential fertility by phenotype and genotype across the 20th century. Proc Natl Acad Sci USA 113(24):6647–6652.
34. Kirk KM, et al. (2011) Natural selection and quantitative genetics of life-history traits in Western women: A twin study. Evolution 65(2):422–435.
35. Stupl G, Barrett L (2016) Evolutionary perspectives on human height variation. Biol Rev Camb Philos Soc 91(1):206–234.
36. Golden CD, Katz LF (2009) The Race Between Education and Technology (Harvard Univ Press, Cambridge, MA).
37. Wyshak G, Frisch RE (1982) Evolutionary evidence for a secular trend in age of menarche. N Engl J Med 306(17):1033–1035.
38. Cole TJ (2003) The secular trend in human physical growth: A biological view. Ecol Hum Biol 1(2):161–168.
39. Bulik-Sullivan BK, et al.; Schizophrenia Working Group of the Psychiatric Genomics Consortium (2015) LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 47(3):291–295.
40. Lawson DW, Mace R (2011) Parental investment and the optimization of human family size. Philos Trans R Soc Lond B Biol Sci 366(1563):333–343.
41. Jones JH, Bird RB (2014) The marginal valuation of fertility. Evol Hum Behav 35(1):65–71.
42. Brommer JE, Gustafsson L, Pietilänen H, Merilä J (2004) Single-generation estimates of individual fitness as proxies for long-term genetic contribution. Am Nat 163(4):505–517.
43. Courtial A, et al. (2013) The demographic transition influences variance in fitness and selection on height and BMI in rural Gambia. Curr Biol 23(10):884–889.
44. Moorad JA (2013) A demographic transition altered the strength of selection for fitness and age-specific survival and fertility in a 19th century American population. Popul Biol 67(6):1622–1634.
45. Vogl TS (2016) Differential fertility, human capital, and development. Rev Econ Stud 83(1):365–401.
46. Bell FC, Miller ML (2005) Life Tables for the United States Social Security Area 1900–2100. Actuarial Study (Sec Soc Admin, Baltimore), Vol 120.
47. Bulik-Sullivan B, et al. (2014) Meta-analysis of genetic correlations across human diseases and traits. Nat Genet 47(1):1236–1241.
48. University of Washington (2012) Quality control report for genotypic data. Available at hsrnline.isr.umich.edu/stedocs/geneics/HRHS_QC_REPORT_MAR2012.pdf.
49. Augustin, S (2013).
50. Rietveld CA, et al.; Social Science Genetics Association Consortium (2014) Replicability and robustness of genome-wide-association studies for behavioral traits. Psychol Sci 25(11):1975–1986.
51. Van Os J, Jones PB (2001) Neuroticism as a risk factor for schizophrenia. Psychol Med 31(6):1129–1134.
52. Kendler KS, Ohsson H, Sundquist J, Sundquist K (2015) IQ and schizophrenia in a Swedish national sample: their causal relationship and the interaction of IQ with genetic risk. Am J Psychiatry 172(2):259–265.
53. Okasha M, McCarron P, McEvoy M, Smaill FG (2001) Age at menarche: secular trends and association with adult anthropometric measures. Ann Hum Biol 28(1):68–78.
54. Efron B, Tibshirani RJ (1994) An Introduction to the Bootstrap (CRC, New York).