Supplementary Note

"Genome-wide analysis identifies 12 loci influencing human reproductive behavior"

TABLE OF CONTENTS

1. HUMAN REPRODUCTIVE BEHAVIOR MOTIVATION AND PHENOTYPE DEFINITION 4

1.1 PHENOTYPE MOTIVATION 4
1.2 EVOLUTIONARY CAUSES OF GENETIC VARIANCE IN FERTILITY 7
1.3 ADDITIVE GENETIC VARIATION IN FERTILITY 7
1.4 DOMINANT GENETIC VARIATION IN FERTILITY 8
1.5 ENVIRONMENTAL VARIATIONS IN FERTILITY 9
1.6 PHENOTYPE DEFINITION 11
1.7 INSTRUCTIONS FOR CONTRIBUTING COHORTS 11

2. PRIMARY GWAS OF HUMAN REPRODUCTIVE BEHAVIOR 13

2.1 OVERVIEW OF HUMAN REPRODUCTIVE BEHAVIOR ANALYSES 13
2.2 PARTICIPATING COHORTS 14
2.3 GENOTYPING AND IMPUTATION 15
2.4 ASSOCIATION ANALYSES 15
2.5 QUALITY CONTROL 16
FILTERS 16
DIAGNOSTIC CHECKS 18
2.6 META ANALYSES 19

3. BIVARIATE AND CONDITIONAL ANALYSIS OF THE TWO FERTILITY-RELATED TRAITS 21

4. TESTING FOR POPULATION STRATIFICATION 22

4.1 LD SCORE INTERCEPT TEST 22
4.2 STATISTICAL SIGNIFICANCE OF THE POLYGENIC SCORES IN A WF REGRESSION 24

5. SEX-SPECIFIC GENETIC EFFECTS IN HUMAN REPRODUCTIVE BEHAVIOR 25

5.1 SEX-SPECIFIC GWAS META-ANALYSES FOR AFB AND NEB 26

5.2 GENETIC OVERLAP AMONG SEXES USING LD SCORE BIVARIATE REGRESSION 26

5.3 GENETIC OVERLAP AMONG SEXES USING GCTA 27

5.4 BIVARIATE GREML ANALYSIS 27

5.5 ANALYSIS OF DIFFERENCES BETWEEN SAMPLE AND EFFECT SIZES 28

5.6 DISCUSSION 29

6. POLYGENIC SCORES PREDICTION 30

6.1 LINEAR POLYGENIC SCORES FOR AFB AND NEB 30

6.2 LINEAR POLYGENIC SCORES FOR INFERTILITY 31

6.3 ADDITIONAL STATISTICAL MODELS FOR CENSORING AND SELECTION 31

6.4 LINEAR PREDICTION OF AGE AT MENARCHE AND AGE AT MENOPAUSE USING AFB LINEAR SCORE 33

6.5 ASSOCIATION OF MENOPAUSE VARIANTS WITH AFB 33

6.6 DISCUSSION: THE PREDICTIVE POWER OF POLYGENIC SCORES 34

7 GENETIC CORRELATIONS 35

7.1 ESTIMATING GENETIC OVERLAP USING LD SCORE REGRESSION 35

7.2 ESTIMATING THE GENETIC CORRELATION BETWEEN AFB AND NEB 36

7.3 RESULTS: PHENOTYPIC CORRELATIONS WITH HUMAN REPRODUCTIVE BEHAVIOR 37

7.4 DISCUSSION 38

7.4.1 HUMAN DEVELOPMENT 38

7.4.2 CARDIOMETABOLIC TRAITS 38

7.4.3 ANTHROPOMETRIC TRAITS 41

7.4.4 AFB AND EDUCATIONAL ATTAINMENT 42

7.4.5 AFB, PERSONALITY AND NEUROPSYCHIATRIC DISORDERS 42

7.4.6 SMOKING BEHAVIOR 43

7.4.7 LIMITATIONS OF LD SCORE REGRESSION GENETIC CORRELATIONS 43

8. LOOK-UP OF LEAD SNPS IN AFB GWAS FOR AGE AT MENOPAUSE AND AGE AT MENARCHE 45

9. BIOLOGICAL ANNOTATION 47

9.1. IDENTIFYING POTENTIALLY CAUSAL VARIANTS 47

9.2. GENE-BASED GWAS ANALYSIS 48
9.3. EQTL AND MQTL ANALYSES 48
GENOTYPE DATA 48
9.3.2 RNA DATA PREPARATION, SEQUENCING AND QUANTIFICATION 48
9.3.3 METHYLATION DATA GENERATION, MAPPING AND NORMALIZATION 49
9.3.4 EQTL AND MQTL ANALYSIS 50
9.4. FUNCTIONAL VARIANT ANALYSIS USING REGULOMEDB 50
9.4.1 GENE PRIORITIZATION USING FOUR BIOINFORMATICS APPROACHES 50
9.5. FUNCTIONAL NETWORK AND ENRICHMENT ANALYSES 51

10. GENE-ENVIRONMENT INTERACTIONS 52

10.1 COHORT ANALYSIS 53
10.2 EDUCATIONAL ATTAINMENT 54

11. ROBUSTNESS CHECKS 55

12. POSITIVE SELECTION 56

13. ADDITIONAL ACKNOWLEDGMENTS 56

14. LIST OF SUPPLEMENTARY TABLES 80

15. LIST OF SUPPLEMENTARY FIGURES 82

REFERENCES 86

1. HUMAN REPRODUCTIVE BEHAVIOR MOTIVATION AND PHENOTYPE DEFINITION

1.1 Phenotype Motivation

Human reproductive behavior – measured by age at first birth (AFB) and number of children ever born (NEB) – is a core topic of research across the medical, social and biological sciences.[1] Two central indicators are the tempo of childbearing of age at first birth (AFB) and the quantum or number of children ever born (NEB). NEB is also often referred to in biological research as life-time reproductive success,[2] number of offspring[3] or as 'fitness' in evolutionary studies, which is the function of the number of children of a person in relation to the number of children of peers of the same birth cohort.[4,5] Due to improvements in hygiene and the reduction in prenatal, infant and child mortality in industrialized societies, NEB has emerged as the gold standard to measure lifetime reproductive success indicating biological fitness.[5]

AFB and NEB are complex phenotypes related not only to biological fecundity, but also behavioral in that they are driven by the reproductive choice of individuals and their partners, and shaped by the social, cultural, economic and historical environment. Genetic factors influence the first two factors of biological fecundity and choice, with the social and historical environment filtering the types of behavior that are possible (e.g., via contraceptive legislation, social norms).

Although interrelated, AFB and NEB, but also childlessness, are distinct phenotypes. Late AFB, low NEB or remaining childless is not only due to 'involuntary' infertility or factors outside of the individual's control (e.g., inability to find a partner), but also 'voluntary' choices to remain 'childfree'.[6] In the past four decades there has been a rapid postponement by around 4-5 years in the AFB to advanced ages in many industrialized societies[7] and a growth in childlessness, with around 20% of women born from 1965-69 in Southern and Western European countries having no children.[8] The biological ability to conceive a child starts to steeply decline for some women as of age 25, with almost 50% of women sterile by the age of 40.[9] Birth postponement and a lower number of children has been largely

attributed to social, economic and cultural environmental factors (i.e., individual and partner characteristics, socioeconomic status).[7,10] Not surprisingly, this delay has led to an unprecedented growth in infertility (i.e., involuntary childlessness), which impacts between 10-15% of couples in Western countries, with men and women affected equally.[8] An estimated 48 million couples worldwide are infertile,[11] with a large part of subfertility, particularly in men, remaining unexplained.[12] Although therapeutic options for infertility in the form of Assisted Reproductive Technology (ART) are available, they are highly ineffective at later ages and older mothers have considerably more problems during gestation and delivery, also associated with low birth weight and preterm delivery.[13–15] Recent studies have also linked advanced maternal age to a higher risk of schizophrenia in offspring.[16]

Childless individuals (and those with a low NEB) are a heterogeneous group consisting of the involuntary childless (e.g., infertility, sterility) and voluntarily childless or 'childfree' (e.g., out of choice). Although primarily related to biological fecundity, involuntary childlessness may also be due to circumstantial socio-environmental reasons outside of the individual's control, including a lack of ability to find a stable partner,[17] divorce and lack of housing, employment or material resources to start a family.[18] Those who are voluntarily childless are generally considered to have made an active choice or to be endowed with an underlying preference[19] or personality traits that pull individuals towards or away from parenthood.[20] It is difficult to disentangle the voluntary from the involuntary, however, since fertility intentions can be adjusted in relation to circumstances[21] and these modifications are age-related.[22]

A better understanding of the genetic architecture of human reproductive behavior and its relation to the environment would enable the discovery of predictors of infertility, which would in turn greatly improve family planning but also reduce costly and invasive ART treatments. Examination of AFB and NEB may also produce a better understanding of the biology of human reproduction, which in turn may give insight into fundamental biological mechanisms and could have ramifications for the study of many health outcomes, especially the etiology of diseases related to the reproductive tract. Furthermore, it is important to understand whether and which proportion of these traits are driven by genetic, behavioral and environmental factors. Relatively little is known about the relationship between indicators of

women's reproductive lifespan (menarche, menopause) and reproductive success. A smaller and recent study has produced some evidence of the link between age at first sexual intercourse (AFS) with AFB and NEB, with a focus on puberty and development.[23]

By systematically investigating the relationship with genetic variants for a multitude of phenotypes related to human reproduction we can establish to what extent diseases related to the reproductive tract play a role in human reproduction and vice versa, and begin to chart the complex biological and related mechanisms that drive human reproduction. It is therefore crucial to examine not only genetic determinants of more biologically proximate phenotypes (e.g. age at menarche, endometriosis, PCOS) but also human reproductive behavior and success. AFB and NEB represent more accurate and concrete measures of observed reproductive success in comparison with proxies which capture the reproductive life span (e.g., age at menarche, menopause) or infertility measures (e.g., endometriosis, PCOS).

To our knowledge, the current study is the largest meta-GWAS effort on human reproductive behavior, which we launched in early 2012. As mentioned previously, a recently published smaller and related study of cohorts also involved in our study focused on age at first sex (AFS), also linking it to AFB and NEB (among other traits).[23] The AFS study examined how individual variation in pubertal timing and personality characteristics related to high risk-taking and low neuroticism related to reproductive activity and success with AFS measures, measures integrated into our examination of genetic correlations (see Supplementary Note, section 7).

Several studies have shown promising results for fertility-related outcomes related to both infertility and the reproductive life span. Previous research has uncovered a genetic component to reproduction with over 70 genome-wide association studies (GWAS) published for 32 traits and diseases associated with reproduction (for a review see ref. [24]). This includes identification of genes such as those related to age at menarche[25,26,27], menopause[28–32], endometriosis[33–36] and polycystic ovary syndrome[37]. This study is the first step towards understanding the pathways between genes and the complex relationship between reproduction and other phenotypes and the environment.

1.2 Evolutionary causes of genetic variance in fertility

Given the diminishing child mortality rate in contemporary societies, evolutionary biologists have used NEB as a proxy for fitness.[2,5,38] Additive genetic variance in fitness implies natural selection within populations: alleles that lead to higher reproductive success will have a higher frequency in future generations.[39] Researchers have until now arguably given less attention to NEB than it deserves,[1] perhaps due to a frequent erroneous interpretation of Fisher's[40] Fundamental Theorem of Natural Selection. The theorem states that the increase in population mean fitness ascribable to changing allele frequencies is equal to the additive genetic variance in fitness. It has often been misinterpreted, however, to mean that the additive genetic variance in fitness itself should always be close to zero. A close reading of the text shows that Fisher actually argued that fitness is moderately heritable in human populations. The misinterpretation of Fisher's theorem is likely repeated so often due to its intuitive appeal. Naively, it may seem that genes that reduce fitness should have been less frequently passed on, leading to the elimination of genetic variability in traits such as fertility.[40,41] Nevertheless, we find that fitness traits such as NEB and AFB have significant narrow-sense heritabilities – yet these traits are still not as heritable as morphological traits such as height.[38,42–44] Several reasons have been put forward to explain the persistent genetic variance in fertility. One argument is that new mutations suffice to restore any genetic variance lost to selection.[45] For the current study design, additional aspects to consider are sexual antagonistic genetic effects, non-additive genetic effects, environment and gene-environment interaction. As discussed in more detail in the Supplementary Note (Section 5), the current GWAS was conducted separately for both sexes, with a detailed examination explored within that section.

1.3 Additive genetic variation in fertility

Several twin and family studies provide evidence for a genetic component underlying both the tempo (AFB) and quantum (NEB) of human fertility.[1,3] Heritability – the proportion of the variance in a trait explained by genetic variance – is typically assessed by a comparison of the phenotypic correlations of family members of different genetic relatedness (for example genetically identical or monozygotic and genetically fraternal dizygotic twins). The genetic component is the extent to which genetically identical twins are more similar in their fertility behavior. As summarized in Fig. S1.1, heritability estimates for AFB (for women)

are around 0.25 and for NEB ranging from 0.15 to 0.45.[1] A recent meta-analysis of all twin studies conducted until 2012[44] shows average heritability of 0.45 (SE = 0.027, N = 50,265) among 64 reproductive disease traits of women and of 0.36 (SE= 0.054, N = 9,376) among 25 reproductive disease traits of men. These mainly pertain to diseases of the genitourinary system, endocrine, nutritional and metabolic diseases, and only few directly pertain to pregnancy, childbirth and the puerperium.

With the advent of molecular genetic data and complementary analytical tools,[46] it has become feasible to go beyond twin models to produce heritability estimates to apply the same logic to unrelated individuals based on the genetic relatedness matrix across all individuals estimated from common SNPs from the whole genome.[47,48] A recent study combined data from the Lifelines Cohort Study and the TwinsUK to estimate this so called SNP-based heritability as the lower bound of narrow sense heritability.[38] Results show that 10% of the variance in NEB and 15% of the variance in AFB are associated with common additive genetic variance. Given that SNP-based heritability is estimated from the same genomic information as utilized in GWAS studies, these results suggest that we should be able to find genetic variants associated with human fertility when conducting GWAS meta-analyses of sufficient sample size.

1.4 Dominant genetic variation in fertility

GWAS typically assume additive genetic effects. Dominant models, however, are in principle also applicable.[49] Dominant genetic effects and overdominance (heterozygote advantage) are mechanisms which potentially maintain non-additive genetic variation in fertility and other fitness related outcomes.[40] Dominant genetic effects result if the conditional phenotypic mean of the heterozygote is not exactly intermediate between those of the homozygotes. Overdominance refers to the special case of the heterozygote possessing a fitness advantage over both homozygotes. At the equilibrium under selection, overdominance leads to an absence of additive genetic variance. Any deviation from strict additivity within a locus, however, should lead to dominance variance that is in principle detectable.[45]

Previous studies approaching the genetic architecture of human fertility almost exclusively relied on twin designs.[1] Dominant genetic effects are detectable in twin studies if the correlation in a trait among identical twins exceeds twice the correlation of fraternal twins. Correlations amongst family members, however, can by inflated by shared environmental factors and therefore hide dominant effects – a potential reason why previous twin studies did not find effects.[49]

Recently, Zhu and colleagues[49] developed a method to estimate dominant genetic effects based on the genetic relatedness of unrelated individuals. This is a complementary approach to the established GREML analysis, which estimates additive genetic effects on traits. In the article of Zhu and colleagues, they quantify dominant relative to additive variance components for 79 quantitative traits and find little evidence for dominant effects. We applied the GREML model to investigate additive genetic effects on NEB and AFB in combined cohorts of women from the TwinsUK and the Lifelines study in the Netherlands.[38] On a slightly larger sample – with a relaxed relatedness cut-off of 0.05[50] and the exclusion of women younger than 45 for AFB – we replicated previous findings with a SNP-heritability of 0.09 for NEB and 0.17 for AFB. However, we find no evidence for dominant genetic effects δ^2_{SNP} for either NEB (0.1×10^{-06}, SE 0.07, P=0.45) nor AFB (0.02, SE=0.08, P=0.43, see Supplementary Table 28 for results). We can therefore conclude that due to this lack of evidence of dominant genetic effects, it is not problematic that we have excluded dominant models in our GWAS.

1.5 Environmental variations in fertility

Social scientists, such as demographers and sociologists, have attributed later ages of first birth, lower NEB and growing levels of childlessness in many industrialized societies almost exclusively to socio-environmental factors.[7,10] The underlying socio-environmental forces shaping fertility can be divided into four main factors. First, the introduction of efficient and reliable contraceptives in the early 1960s revolutionized human reproductive behavior, namely the ability to control the timing and number of children.[7] The diffusion of the pill in the late 1960s in the United States resulted in an almost immediate postponement in the age of first marriage for college-educated women.[51] Contraception allowed women and couples

to avoid pregnancy and delay entry into parenthood. Contraceptives were generally widely introduced across Western and Northern Europe, Australia and North America in the late 1960s, which is where the majority of cohorts are situated in the current study.

Second, there is a well-documented association between female education and later AFB.[52] Early research demonstrated a strong inverse relationship between education and fertility, with women's increased participation in higher college and University degrees resulting in a significant shift to later AFB.[53–55] A central argument driving childbearing delay was the difficulty to balance student and mother (parent) roles, but also women's opportunity costs in terms of wages and career progression that they forego when having children early.[56–58] A third factor, which is interdependent with educational level, is women's labor force participation and attachment. Research has demonstrated an incompatibility of early AFB and high NEB with paid labor force participation,[59] largely due to work-family conflict[7] and the high motherhood 'wage penalty'. In fact, the postponement of AFB results in substantial increases in earnings, particularly for higher-educated women.[60,61] It is estimated that there is a 7% motherhood wage penalty per child, with a year delay of entry into motherhood increasing career earnings by 9%.[61]

A fourth factor is the Second Demographic Transition, which encompasses cultural and ideational changes surrounding the preferences for and role of children, which is coupled with a shift to more individualistic desires for personal development.[62,63] Since infant mortality rates have fallen sharply in modern societies, extra births are not required for insurance against death and children no longer provide the economic support and labor to support parents that they once did, which dramatically changes preferences and the need to have children.[64,65] Multiple national institutional factors have also been shown as related to the delay of AFB and the decrease in NEB. This includes changes in the educational systems, labor market regulations, gender equity,[66] but also economic uncertainty,[67] the housing market,[68] influence by friendship networks,[69] family networks and social capital,[70] and changes in partnering and mating practices.[71] The empirical relationship of these factors – namely birth cohort and educational level – with genetic risk scores of AFB and NEB is elaborated upon in section 10.

1.6 Phenotype definition

The current study measures human reproductive choice by the two phenotypes of: age at first birth (AFB) and number of children ever born (NEB). AFB is the self-reported age when subjects had their first child. In most cohorts this was asked directly (e.g. *"How old were you when you had your first child?"*). Alternatively, it could also be calculated based on several survey questions (such as the date of birth of the subject and date of birth of the first child). Supplementary Table 2 describes in detail the exact question asked for each cohort and if applicable, whether and how it varies in the way it was asked to men and women. Often these questions were part of a medical questionnaire about women's reproductive health. In a large number of cohorts, this means that only women were asked this question. For this reason, the sample size for AFB for women is considerably larger than for men. Note that only people who have had at least one child (parous) are eligible to be included for the analysis of this phenotype.

Number of children ever born (NEB) was the self-reported number of children. This phenotype was either asked directly (e.g. "*How many children do you have?"* or "*How many natural (biological) children have you ever had, that is, all children who were born alive?*", or "*How many children have you had - not counting any step, adopted, or foster children, or any who were stillborn?*") or it was calculated based on several survey questions (such as pregnancy histories and outcomes, number of deliveries). In most cases it was possible to distinguish between biological (live born or stillborn) and adopted or step-children. When it was possible to distinguish between cases, we used the number of live born biological children. We included cases for NEB if they finished their reproductive career (aged at least 45 for women and 55 for men at time of study) and were thus unlikely to have future biological children.

1.7 Instructions for contributing cohorts

The instructions given to cohorts who agreed to participate in our study is described in detail in the original Analysis Plan that was posted on the Open Science Framework preregistration site, described in detail in Supplementary Note Section 2.1 and uploaded December 9, 2013 at: https://osf.io/53tea/. For ease of analysis, we advised that AFB should be treated as a

continuous measure. When possible, we asked analysts to use the more direct question: How old were you when you had your first child? Another variant of this question is: What is the date of birth of your first child? In the case of the latter, we advised them to create the AFB variable by subtracting the date of birth of the first child from the date of birth of the subject.

Analysts then normalized the raw measure of the age at first birth for sex/ birth cohort specific means and standard deviations. In other words, we asked them to compute a mean and standard deviation separately for men and women by birth cohort category (generally ten-year intervals) and then subtract the mean value for that group from the respondent's value. They should then divide the result by the standard deviation. This was used as the final AFB variable measured in sex/cohort specific Z-score and is our regressand.

Analysts were asked to include birth year of the respondent (represented by birth year – 1900), its square and cubic to control for non-linear birth cohort effects. Combined analyses that included both men and women also needed to include interactions of birth year and its polynomials with sex. Some cohorts only used birth year and not its polynomials because of multi-collinearity issues/convergence of the GWA analysis.

2. PRIMARY GWAS OF HUMAN REPRODUCTIVE BEHAVIOR

2.1 Overview of human reproductive behavior analyses

The genome-wide association study (GWAS) of human reproductive behavior is based on the summary statistics that were uploaded to a central server by cohort-level analysts. As outlined in more detail in Section 1 of the Supplementary Note, our analysis includes the two phenotypes of age at first birth (AFB) and number of children ever born (NEB), with analysts producing results for women, men and combined analyses of both sexes, also including birth cohort as a covariate. The summary statistics were then subsequently quality-controlled and meta-analyzed by two separate independent centers at the University of Oxford and Erasmus University Rotterdam.

We follow the QC protocol of the GIANT consortium's recent study of human height[72] and employed the software packages QCGWAS[73] and EasyQC[74], which allowed us to harmonize the files and identify possible sources of errors in association results. This procedure entailed that diagnostic graphs and statistics were generated for each set of GWAS results (i.e., for each file). In the case where apparent errors could not be amended by stringent QC, cohorts were excluded from the meta-analysis (see the bottom of Supplementary Table 1 for a list of excluded cohorts).

The lead PI of each cohort confirmed that the results on these analyses were approved by the local Research Ethics Committee and/or the relevant Institutional Review Board. All participants fell under the written informed consent protocol of each participating study. The entire project was also approved by the local Research Ethics Committee of the PI.

We first circulated three documents to interested cohorts at the end of April 2012, which included: (a) Rationale for a GWAS of Fertility Behavior, (b) GWAS Fertility Behavior Analysis Plan; and, (c) Collaboration Agreement for Fertility GWAS Meta-analyses. This was after a meeting and approval from the REPROGEN working group of the CHARGE consortium on Dec. 9, 2011 that we were not competing with or unduly replicating existing efforts. Preliminary results were presented at various CHARGE meetings between the years

of 2012-2015. This study was initially set up as a two-stage GWAS with a large discovery and smaller replication phase. Due to an increasing influx of new data, we opened the participation to cohorts that had genome-wide data, but also to cohorts that had Metabochip data. We also included a list of 15 independent SNPs with $P<10^{-06}$ for cohorts that did not have genome-wide data available but could perform *de-novo* replication on a limited number of SNPs. Agreements at a later stage included data from RPGEH (Kaiser Permanente Research Program on Genes, Environment, and Health, REPEGH/GERA), *N*(AFB women)=31,898, *N*(NEB women)=39,576), deCODE (*N*(AFB pooled)=60,602, *N*(NEB pooled)=65,228), and UK Biobank (*N*(AFB women)=40,082, *N*(NEB pooled)=88,094). Given the resulting well-powered total sample size of *N*≈250k for *AFB* and *N*≈340k for *NEB,* we chose to merge the discovery and replication cohorts into a single large discovery phase, as in other recent well-powered GWAS efforts.[72,75,76] We also opted to include only cohorts with genome-wide data in the meta-analysis, leaving the remaining cohorts that performed *de-novo* replication for follow-up analysis.

2.2 Participating Cohorts

A total of 62 cohorts contributed to this study. Cohorts with acceptable measures of AFB and/or NEB were eligible to participate. Some measured one or both of the phenotypes, and there was also variation by whether the question was asked to women and/or also men. Supplementary Table 1 provides an overview of the study-specific details of all analyses conducted for the traits of interest. Cohorts of unrelated individuals uploaded separate results for men and women. In addition to sex-specific association results, family-based cohorts uploaded pooled results. As described in the Supplementary Note (Section 1), particularly AFB is less frequently asked of men. The total number of association-result files per trait is as follows. We have 28 files for AFB men, 57 for AFB women, 72 for AFB pooled, 50 for NEB men, 67 for NEB women, and 102 for NEB pooled.

As Supplementary Table 1 shows, most cohorts were included in the meta-analysis (i.e., 62 cohorts are included, constituting 26 files for AFB men, 50 for AFB women, 64 for AFB pooled, 47 for NEB men, 60 for NEB women, and 91 for NEB pooled) and some only in the follow-up analyses (9 cohorts, constituting 2 files for AFB men, 5 for AFB women, 6 for

AFB pooled, 3 for NEB men, 5 for NEB women, and 9 for NEB pooled). We had to exclude the association results of two cohorts – ABCFS (AFB women, *N*=410, NEB women, N=410) and Longenity (AFB women, *N*=285; NEB women, N=352) – from the meta- and follow-up analyses due to unresolvable issues with the cohort's association results that came up in the quality control procedures. For more details regarding the reasons for exclusion, see SI Section 2.6.

2.3 Genotyping and Imputation

Supplementary Table 1 gives an overview of the study-specific details on pre-imputation quality control filters applied to the genotype data, subject-level exclusion criteria, imputation software used, and the reference sample for imputation. Due to the fact that we started our study in 2012 before 1000G imputation, our analysis plan recommended using resulted imputed using the HapMap 2 CEU (r22.b36) reference sample.[77]

2.4 Association analyses

Cohorts were asked to only include participants of European ancestry, with no missing values on all relevant covariates (sex, birth year, and cohort specific covariates), who were successfully genotyped genome-wide (e.g., genotyping rate greater than 95%), and who passed cohort-specific quality controls (e.g., no genetic outliers).

Cohorts used the fully imputed set of HapMap Phase 2 autosomal SNPs, and estimated an additive linear model, including top principal components to control for population stratification and cohort specific covariates if appropriate. They were specifically instructed to control for population stratification for ancestry principal components with reference to Price et al. (2006).[78] In addition, cohorts were requested to include the birth year of the respondent (represented by birth year – 1900), its square and cubic to control for non-linear birth cohort effects. Analyses pooling data across sexes also needed to include interactions of birth year and its polynomials with sex. Some cohorts only used birth year and not its polynomials because of multi-collinearity issues/convergence of the GWA analysis. Omission of these nonlinear birth year effects is unlikely to lead to biased inferences, since

genotypes are not usually considered as truly associated with birth year. However, inferences might be less accurate (i.e., have larger standard errors), since omission of nonlinear birth year effects can lead to larger residual variation.

2.5 Quality Control

In this section, we summarize the main steps and diagnostic tests of the Quality Control (QC) procedure. The quality control was conducted in two separate independent analysis centers (Oxford/Groningen and Rotterdam). Once data were submitted, each study was independently subjected to quality control in the two analyses centers according to standard protocols. We followed the QC protocol of the GIANT consortium's recent study of human height[72] and the SSGAC's study of educational attainment.[76,79]

Since this study began, QC procedures have become more stringent. Recently, a comprehensive set of guidelines for GWAS QC was released.[7] For the cohorts initially included in the study a first round of QC was performed using the R package QCGWAS[73]. We updated the QC protocol based on the GIANT consortium's and SSGAC's protocols. The updated QC protocol was applied to all cohorts using the R package EasyQC.[74] Findings of the first round of QC were used as a starting point for the updated QC.

In the QC procedure, diagnostic graphs and statistics were generated for each set of GWAS results (i.e., for each result file uploaded by the cohort analysts). Most errors (e.g., coded allele reported as other allele and vice versa) could be easily addressed. When apparent errors could not be amended by combining stringent QC with file-specific inspections and corrections, cohorts were excluded from the meta-analysis. For details on cohort inclusion and exclusion, see Supplementary Table 1.

Filters

We harmonized base pair positions of the markers across files using NCBI build 37. For each result file, a given marker was excluded in case:

1. The combination of chromosome and base-pair position could not be uniquely linked to the HapMap Phase II CEU panel.
2. The marker had missing or incorrect values. Specifically,
    a. the effect allele and other allele were missing,

b. the association *p*-value was missing or outside the unit interval,
c. the effect estimate was missing or reported to have infinite magnitude,
d. the standard error (SE) of the effect estimate was missing, negative, or infinite,
e. the allele frequency was missing or outside the unit interval,
f. the sample size was not reported, or zero or below,
g. the reported callrate was outside unit interval,
h. the reported imputation quality was negative, and
i. the reported imputed dummy was not binary.

3. The marker was not a SNP, not biallelic, non-autosomal, and/or monomorphic.
4. The sample size was below 30.
   This filter is to guard against spurious associations due to overfitting of the model.
5. The minor allele count was 6 or below.
   This filter is to guard against spurious associations with low-frequency SNPs in small samples. The risk of spurious associations has shown to be particularly high for SNPs that are extremely rare[7].
6. Minor allele frequency (MAF) was below 1%.
   For all the cohorts, we dropped SNPs with a MAF below 1%. For small cohorts we applied more stringent filters based on diagnostic tests and figures.
7. The SE of the effect estimate was greater than $100/\sqrt{N}$.
   Based on the approximation to the expected standard error by Winkler *et al.*[7], we calculated that an SE greater than $100/\sqrt{N}$ is at least 40% greater than the expected SE of the estimated effect of a SNP with a MAF of 1% for a trait with standard deviation of 10. Since in our analyses we only consider SNPs with MAF≥1% and traits with a standard deviation below 10, an effect estimate with an SE greater than $100/\sqrt{N}$ can be considered to be unreasonably large.
8. The R^2 of the marker with respect to the phenotype was greater than 10%.
   We excluded SNPs for which the estimated R^2 was greater than 10% (Supplementary Information in Rietveld *et al.*[79]) because such an R^2 would defy all upper bounds on reasonable effect sizes of SNPs.
9. The marker was imputed while imputation quality was missing.
10. The marker was imputed while imputation quality was below 0.4.
    For all the cohorts, we dropped imputed SNPs with an imputation quality below 0.4. For several cohorts we apply more stringent filters based on diagnostic tests and figures.
11. The callrate was below 95%.
12. The SNP was genotyped and not in Hardy-Weinberg Equilibrium (HWE).
    We excluded genotyped SNPs if they fail the HWE chi-squared test. Violation of HWE will lead to lower chi-squared *p*-values as sample size increase, the threshold is therefore sample-size dependent. We applied an HWE *p*-value threshold of 10^{-03} in case $N < 1{,}000$, 10^{-04} in case $1{,}000 \leq N < 2{,}000$, 10^{-05} in case $2{,}000 \leq N < 10{,}000$, and no filter in case $N \geq 10{,}000$).

Diagnostic checks

For the SNPs remaining after applying the filters of steps 1 – 12, we generated five key diagnostic graphs:

1. Allele frequency (AF) plots. – to identify errors in allele frequencies and strand orientations.

The AF plot shows the expected AF (based on the HapMap II CEU2 reference panel or the 1000 Genomes Phase 1 European panel in case of 1000 genomes imputed data) versus the reported AF.

2. Reported P-values versus P-values of the Z-scores (PZ) plots – to assess the consistency of the reported P-values with respect to those implied by the effect estimates and the corresponding standard errors.
3. Quantile-Quantile (QQ) plots – to check for evidence of unaccounted population stratification.
4. Reported Standard Error versus Expected Standard Error (SE) plots – to assess whether the reported standard errors behave in line with the approximation of the expected standard errors provided by Winkler et al.[74] , implemented as a QC step by Okbay *et al*.[80]

These diagnostic plots were examined by two independent analysts. If problems were detected which could not be resolved by more stringent thresholds, we applied the following *ad hoc* filters (descending order in terms of frequency used).

1. MAF filters more stringent than the generic MAF filter (e.g., 5% instead of 1%).
2. Imputation quality filters more stringent than the generic filter (e.g., 0.8 instead of 0.4).
3. Filter on the absolute difference between expected (based on the HapMap II CEU2 reference panel or the 1000 Genomes Phase 1 European panel in case of 1000 genomes imputed data) and reported allele frequencies. This filter helps to remove clear outliers in the AF-plots (e.g., strand-ambiguous SNPs that are likely to have been reverse-coded).
4. Filter on the absolute difference between the reported log(*P*-value) and the log(*P*-value) derived from the report Z-score. This filter helps to remove clear outliers in the PZ-plots. Such outliers can arise when software such as SNPTEST[13] switches to another estimation method, for reasons such as poor convergence of the estimates.

For a list of the filters used per cohort, per association file, see Supplementary Table 27, which reports the total number of markers prior and post-QC when applying the described generic and specific filters, for each set of association results.

The AF plots for ABCFS (*N*=410 for AFB and NEB) shows a strong anti-diagonal that persists when considering only genotyped markers, implying that reverse-coded SNPs are likely to have been used for imputation, thereby yielding poorly imputed SNPs. Consequently, we exclude the ABCFS result files from the meta-analyses. In addition, for Longenity (*N*=285 for AFB and *N*=352 for NEB) many SNPs have far greater standard errors

for the effect estimates than expected, as well as callrates substantially below 95%. When applying QC to Longenity, only several hundreds of SNPs are left after QC. Consequently, we also exclude Longenity results from the meta-analyses.

2.6 Meta analyses

Cohort association results (after applying the QC filters) were combined using sample-size weighted meta-analysis, implemented in METAL.[81] Sample-size weighting is based on Z-scores and can account for different phenotypic measurements among cohorts.[82] The two QC centers agreed in using sample-size weighting to allow cohorts to introduce study-specific covariates in their cohort-level analysis. Only SNPs that were observed in at least 50% of the participants for a given phenotype-sex combination were passed to the meta-analysis. SNPs were considered genome-wide significant at *P*-values smaller than 5×10^{-08} (α of 5%, Bonferroni-corrected for a million tests. The meta-analyses were carried out by two independent analysts. Comparisons were made to ensure concordance of the identified signals between the two independent analysts. The PLINK clumping function[83] was used to identify the most significant SNPs in associated regions (termed "lead SNPs").

The total sample size of the meta-analysis is *N*=251,151 for AFB pooled and *N*=343,072 for NEB pooled. Although considered to be separate from our main pooled results, we also performed separate meta-analyses for

- AFB women (*N*=189,656),
- AFB men (*N*=48,408),
- NEB women (=225,230),
- NEB men (*N*=103,909)

The sex-specific results are discussed in more detail in Supplementary Note, Section 5. To understand the magnitude of the estimated effects, we used an approximation method to compute unstandardized regression coefficients based on the Z-scores of METAL output obtained by sample-size-weighted meta-analysis, allele frequency and phenotype standard deviation. Further details of the approximation procedure are available in the Supplementary Information of Rietveld et al.[79]

Figure S2.1.1. to Figure S2.13.2 contains the forest plots and regional association plots of all genome-wide significant SNPs, the latter created by LocusZoom plots.[84] The forest plots provide a visualization of the effect size estimates for each cohort and the summary meta-analysis (red rectangle) in addition to the 95% confidence intervals. As would be expected, small cohorts have larger confidence intervals. LocusZoom plots provide a graphic depiction of the local association results and include information about the locus, the location and orientation of the genes it includes, LD coefficients and the local estimates of recombination rates.

3. BIVARIATE AND CONDITIONAL ANALYSIS OF THE TWO FERTILITY-RELATED TRAITS

As joint analysis of correlated traits may boost power for mapping functional loci, we applied a recently developed multi-trait analysis method[85] to test the association between each variant and the two correlated traits AFB and NEB simultaneously using multivariate analysis of variance (MANOVA). The analysis was performed based on the genome-wide meta-analysis summary statistics of each single trait. The joint analysis did not reveal additional genome-wide significant loci (λ=0.995), however, such bivariate analysis, accounting for the correlation between the two phenotypes, improved the strength of two signals on chromosomes 1 and 5, indicating possible pleiotropic architecture between the AFB and NEB (Supplementary Figure 30).

The analysis also provides a conditional association test of the genetic effect of each variant on AFB including NEB as a covariate, and that on NEB including AFB as a covariate. The conditional analysis also did not reveal additional genome-wide significant loci (Supplementary Figure 31). Nevertheless, adjusting for NEB eliminated the three genome-wide significant loci on chromosomes 1, 2 and 6 for AFB, and adjusting for AFB eliminated the two genome-wide significant loci on chromosomes 1 and 14 for NEB, which may indicate underlying pleiotropic effects on both traits across these loci.

4. TESTING FOR POPULATION STRATIFICATION

Population stratification can severely bias GWAS estimates for causal variants and lead to false positives. This can occur if a particular variant of a SNP is more common in a particular subpopulation and if there are mean differences in the phenotype of interest between subpopulations due to factors that do not involve that SNP. As described in Supplementary Note Section 2, all cohorts in the GWAS of AFB and NEB included the top principal components[78] in their analyses to account for population stratification. Even despite this inclusion, residual stratification could still remain and affect the results.

To test the extent of this problem, we used two methods to assess if our GWAS results for AFB and NEB exhibit signs of population stratification. First, we used the LD Score intercept method described in Bulik-Sullivan et al..[86] Second, we conduct a series of individual and within-family (WF) regressions using polygenic scores (PGS) as predictors[87–89] on a dataset of DZ twins (STR and TwinsUK). Within-family regressions are based on family differences in PGS for AFB and NEB and are therefore are not affected by population stratification. We compare the coefficients of individual and WF regression using different p-value thresholds for the construction of PGS. Polygenic scores are based on independent results (i.e. meta-analysis results excluding STR and TwinsUK). Additional information on how we computed our PGS are available in Section 7 of the Supplementary Note.

4.1 LD Score Intercept Test

The LD Score intercept test uses GWAS summary statistics for all measured SNPs. LD Score regression is a method that can disentangle inflation in the chi-square statistics that is due to a true polygenic signal throughout the genome from inflation that is due to confounding biases such as cryptic relatedness and population stratification. The inflation due to a true polygenic signal impacts the slope of the LD regression, whereas inflation due to population stratification and other confounding biases affects the intercept of the regression.

We used the LDSC software[86,90] to estimate the intercepts in LD Score regressions with the summary statistics of our GWAS of: (i) AFB (pooled sample), (ii) NEB (pooled sample), (iii) AFB (women), (iv) AFB (men), (v) NEB (women), and, (vi) NEB (men). We estimated a separate LD Score regression for each of the phenotypes using the summary statistics from the meta-analyses based on all available data.

For each phenotype, we used the "eur_w_ld_chr/" files of LD Scores computed by Finucane et al.[91] available on https://github.com/bulik/ldsc/wiki/Genetic-Correlation. These LD Scores were computed with genotypes from the European-ancestry samples in the 1000 Genomes Project using only HapMap3 SNPs. Only HapMap3 SNPs with MAF > 0.01 were included in the LD Score regression.

Because genomic control (GC) will tend to bias the intercept of the LD Score regression downward, we did not apply GC to the summary statistics we used to estimate the LD Score regression. Furthermore, we excluded the deCODE cohort from the data for the estimation of the LD Score intercept for AFB and NEB, since the cohort-level regression estimates for deCODE did not directly correct for the high level of relatedness in the sample (their standard procedure is to apply GC). Our intercept estimates from the LD Score regressions are thus unbiased measures of the amount of stratification there is in the data (excluding deCODE) that we used for the GWAS of each phenotype.

Supplementary Note Figures 4.1 and 4.2 show LD Score regression plots based on the summary statistics from the GWAS of AFB, and NEB. For AFB, we estimated a LD Score intercept of 1.0216 (SE=0.008) and for NEB, 1.009 (SE = 0.006). In all six cases, the intercept estimates are not significantly different from 1. By comparison, the mean χ^2 statistics for all the SNPs in the LD Score regressions are 1.239 for AFB and 1.141 for NEB. Under the null hypothesis that there is no confounding bias and that the SNPs have no causal effects on the phenotypes, the mean χ^2 statistics would be one, thus mean χ^2 statistics greater than one indicate that some SNPs are associated with the phenotypes. These estimates imply that about 9% of the observed inflation in the mean χ^2 statistics for AFB and about 6% of the

inflation for NEB is accounted for by confounding bias (due to relatedness, or other confounds) rather than a polygenic signal.

As described in Section 2 of the Supplementary Information, we applied the standard single GC correction to produce our main estimates. Once a single GC is applied, the LD score regression estimates indicate negligible confounding bias due to population stratification. The LD score intercept for AFB is 0.9618 (SE= 0.0077) and for NEB 0.9763 (SE=0.0068). We can therefore conclude that the amount of inflation in our final results due to confounding bias is likely to be negligible.

4.2 Statistical Significance of the Polygenic Scores in a WF regression

To test the robustness of our all-SNP polygenic scores calculated with a set of SNPs meeting several different threshold P-values (5e-08, 5e-07,5e-06, 5e-05, 5e-04, 5e-03, 5e-02, 5e-01, all SNPs), we estimated WF regressions of AFB and NEB on each polygenic score in samples that are independent from those used to construct the scores. For each WF regression, we also compared the estimated coefficient on the polygenic score to the corresponding coefficient from individual-level regression.

For both the individual-level and WF regression, we standardized NEB and AFB on birthyear, birthyear squared, birthyear cubic, sex and the first 10 PCAs[a]. Our regressions are based on 7,944 twin couples for AFB and 9,220 twin couples for NEB. Supplementary Note Figures 4.1, 4.2 and Supplementary Tables 30, 31 report the results.

The regression analyses show that WF regression coefficients for both AFB and NEB are statistically different from zero when the p-value threshold is sufficiently far from zero. When including all SNPs, both coefficients for AFB and NEB are larger than zero, confirming that the GWAS analyses uncovered true polygenic signals. Overall, these results indicate a minimum effect of population stratification and the existence of polygenic signals.

[a] Details on the construction of polygenic scores is available in section 6 of the Supplementary Note.

5. SEX-SPECIFIC GENETIC EFFECTS IN HUMAN REPRODUCTIVE BEHAVIOR

Sex-specific genetic effects have been proposed as an important source of variation for complex human traits.[92,93] For this reason we also ran sex-specific GWAS meta-analyses for both AFB and NEB and examined the genetic overlap among sexes using LD score bivariate regression and GCTA. Sex-specific effects refer to large differences in average phenotypes or biological processes known to differ between the sexes (e.g., hormonal effects). Since AFB and NEB are not only biological but also socio-behavioral phenotypes, it is likewise important to make a distinction between sex- versus gender-specific effects. Sex refers to biological differences between males and females, which often have their underpinnings in human reproduction.[94] Gender refers to the socially constructed differences between men and women that may give rise to particular behavioral outcomes (e.g., gender-specific social norms regarding alcohol consumption or occupational choice). There is growing evidence that biological (sex) and social (gender) processes are interrelated, which in turn impacts the phenotypes we are examining.[95] Although we recognize the importance of these distinctions, it is beyond the scope of the current study to disentangle sex- versus gender-effects. Rather in this section, we emphasize similarities and differences in the sex-specific GWAS results and examine the sex-specific genetic overlap of these traits.

There are several key sex-specific differences in AFB and NEB. Women in contemporaneous populations have a comparatively lower age at first birth than men, which is attributed factors such as the persistent age gap between partners.[96] Fecundability is strongly influenced by sex-specific hormonal processes and gender-specific diseases. Sex can modify both penetrance and expressivity of a wide variety of traits.[97,98] Sex-genotype interactions can also theoretically act to maintain genetic variation in a population.[99] Sexual antagonism, which is the existence of opposite genotypic effects among sexes, has been often theorized as one of the possible explanations for genetic differences in fertility.[100] In other words, particular genes might influence men and women differently and could thus still be transmitted to the next generation. Genes that contribute to the fecundability of men may therefore be inherited via women's lineage and those for women via men's lineage.[101]

5.1 Sex-specific GWAS meta-analyses for AFB and NEB

In addition to the pooled GWAS results presented in the main text, we also ran sex-specific GWAS meta-analyses for AFB and NEB. The sample size for sex-specific analysis is: AFB women, N=189,656; AFB men, N=48,408; NEB women N=225,230; NEB men N=103,909. Our results indicate 6 genome-wide significant ($P<5x10^{-08}$) independent SNPs for AFB women and 1 genome-wide significant independent SNP for NEB men. We do not find any genome-wide significant loci for AFB men and NEB women. Among the 6 hits for AFB women, 5 are also significant in the AFB pooled analysis, while 1 hit on chr8 (rs2721195; chr8: 145677011) is specific for women. We find a single independent SNP for NEB men (rs13161115; chr5:107050002) that reaches genome-wide statistical significance (P-value$<5x10^{8}$), which is not significant in the NEB pooled analysis. Supplementary Figure 34 shows the Miami plots for AFB and NEB sex-specific analyses. Supplementary Figure 35 depicts the QQ plots of men and women's meta-analyses for AFB and NEB. The figure shows a noteworthy departure from the null hypothesis of no statistical association, in particular for the analysis of AFB women.

Table 1 (in the main text) shows the sex-specific signals respectively for AFB and NEB. The effects of all significant hits in AFB have the same direction for both men and women. The single locus found in NEB men (rs13161115) has an opposite effect on NEB for women, although the p-value associated with its effect size in NEB for women does not reach statistical significance.

5.2 Genetic overlap among sexes using LD score bivariate regression

We used LD score bivariate regression[8] to estimate the genetic correlation among men and women based on the sex-specific summary statistics of AFB and NEB meta-analysis results. For each phenotype, we used the "*eur_w_ld_chr/"* files of LD Scores computed by Finucane et al. and made available on https://github.com/bulik/ldsc/wiki/Genetic-Correlation. These LD Scores were computed with genotypes from the European-ancestry samples in the 1000 Genomes Project using only HapMap3 SNPs. Only HapMap3 SNPs with MAF>0.01 were included in the LD Score regression. Our estimates indicate a genetic correlation of r_g=0.86

(SE=i0.052) among sexes for AFB and r_g=0.97 (SE=0.095) for NEB. These results indicate a large genetic overlap among sexes for both AFB and NEB, which is statistically different from zero. We additionally test whether these genetic correlations support the null hypothesis of complete genetic overlap among sexes (r_g=1). We reject this null hypothesis for AFB, indicating sex-specific genetic variants for AFB. We do not find any evidence of sex-specific signals for NEB.

5.3 Genetic overlap among sexes using GCTA

We additionally estimate the degree of genetic overlap among sexes using Genomic-Relatedness-Matrix Maximum Likelihood (GREML)[46] on six cohorts for which we have direct access to genotypic data.[46,47,102–104] For the GREML analyses, we combine data from six cohorts: HRS, EGCUT, QIMR Lifelines Cohort Study, TwinsUK and STR (N_{women}=20,966; N_{men}=17,024, see Supplementary Table 33 for descriptive statistics). We used GCTA[46] to construct a Genome-wide Relatedness Matrix (GRM) $\boldsymbol{A^{n\times n}}$ and estimate the models. For quality control (QC), we included in the analysis only HapMap3 SNPs with an imputation score larger than 0.6. We additionally excluded SNPs with a missing rate larger than 5%, MAF lower than 1% and which failed the Hardy-Weinberg equilibrium test for a threshold of 10^{-06}. We applied these QC steps for each cohort and repeated again on the merged dataset. After QC, 847,278 SNPs could be utilized to estimate the GRM between individuals.

5.4 Bivariate GREML analysis

First, we fit a bivariate GREML model as proposed by Lee et al.[104] treating the fertility traits of men and women as different traits.[102] To account for potential country heterogeneity, we estimated genetic variation from within cohorts only ($\sigma^2_{g\_wc}$), setting the GRM between individuals from different cohorts equal to zero.[50] This allows us to avoid the potential bias due to differences in allele frequency across different countries. The GRM can be depicted as a block matrix composed by six within-cohort GRMs ($\mathbf{A}_{\boldsymbol{g_{wc}}}$) containing only values for pairs of individuals within cohorts.

The variance-covariance matrix of the bivariate model is shown as:

$$\boldsymbol{V}\begin{bmatrix}\boldsymbol{f_{men}}\\ \boldsymbol{f_{women}}\end{bmatrix}=\begin{bmatrix}\boldsymbol{A_{wc\_men}}\sigma^2_{g\_wc\_men}+\mathbf{I}\sigma^2_{e\_wc\_men} & \boldsymbol{A_{wc\_men\_women}}\sigma^2_{g\_wc\_men\_women}\\ \boldsymbol{A_{wc\_men\_women}}\sigma^2_{g\_wc\_men\_women} & \boldsymbol{A_{wc\_women}}\sigma^2_{g\_wc\_women}+\mathbf{I}\sigma^2_{e\_wc\_women}\end{bmatrix}$$

whereas $\boldsymbol{f_{men}}$ **and** $\boldsymbol{f_{women}}$ are vectors of length $N_{men}\ and\ N_{women}$ of fertility phenotypes (NEB or AFB), with N being the respective sample size of the subsets, $\boldsymbol{A_{wc\_men\_women}}$ is the within population GRM for all individuals, $\boldsymbol{A_{wc\_men}}$ is the within cohorts GRM for men, and $\boldsymbol{A_{wc\_women}}$ for women. The parameter $\sigma^2_{g\_wc\_men}$ is an estimate of the genetic variance component for men and $\sigma^2_{g\_wc\_women}$ and $\sigma^2_{g\_wc\_men\_women}$ the genetic covariance across sexes. **I** is the identity matrix, and $\sigma^2_{e\_wc\_women}$, $\sigma^2_{e\_wc\_men}$ the respective, sex-specific residual variances within cohorts. We present the variance components standardized for the phenotypic variance σ^2_p. The correlation of the genetic factors are estimated as:

$$r_{\sigma^2_{g\_wc\_men\_women}}=\sigma^2_{g\_wc\_men\_women}/\sqrt{\sigma^2_{g\_wc\_men}*\sigma^2_{g\_wc\_\mathrm{women}}}$$

We find significant heritability for NEB and both sexes $\sigma^2_{g\_wp}/\sigma^2_P$ = 0.13 (SE=0.057, P=0.01) for men, and 0.08 (SE=0.04, P=0.01) for women (see Supplementary Table 34 for full results). This means that around 10% of the variance in NEB is explained by common SNPs for both sexes. The estimated genetic correlation across sexes is 0.98 (SE=0.44) and a likelihood ratio-test against a perfect genetic correlation across sexes has a p-value of 0.5. We therefore cannot reject the null-hypothesis that genetic effects are the same across sexes.

For AFB we find a very similar pattern of sex specific SNP-based heritabilities of around 0.10 and a genetic correlation of 1.00 (SE=0.67, P=0.5 when testing against 1). These results also cannot reject the null-hypothesis that genetic effects on AFB are the same across sexes.

5.5 Analysis of differences between sample and effect sizes

Table 1 in the main text did not include the Ns of the sex-specific analyses. It is, however, important to place the p-value of women and men in context and clarify why the effect size for some loci is similar in men and women but the p-value is not. This could reflect a difference in sample size, or it may reflect a difference in variance. Supplementary Table 32 shows all of the sex-specific sample sizes, p-values, z-scores and the p-value differences

between males and females by each SNP. It indicates sex-specific effects and a statistical test showing the differences between effect sizes.

The statistical test is based on the differences between male and female Z-scores:

$$Z_{diff} = \frac{\frac{Z_1}{\sqrt{N_1}} + \frac{Z_2}{\sqrt{N_2}}}{\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}} \sim N(0,1)$$

Supplementary Table 32 reports the P-value differences of this Z-score test. Despite the fact that p-values differ among the sexes, it seems plausible that the differences are mainly due to variation in sample size and not attributed to different effect sizes. Our results show that the only locus that has a statistically different effect between men and women after taking into account the number of test conducted is *rs13161115* in chromosome 5, where the effect is significant only in men and the direction of the effect differs among sexes.

5.6 Discussion

Sex-genotype interactions and sexual antagonistic effects may affect the transmission of traits across generations and has been proposed as a possible source of genetic variation in fertility traits.[101] Fecundability is strongly influenced by sex-specific hormones and infertility causes differ between men and women.[105] Our results show little differences in the genetic architecture of the fertility traits (AFB, NEB) of our study between men and women. Out of 12 independent loci for AFB and NEB, only two have a sex-specific effect. Moreover, all the signals found for AFB and two out of three signals in NEB, have a consistent direction across the sexes. We found a high genetic correlation among men and women for both AFB and NEB, both using LDscore bivariate regression and GREML bivariate analysis. This suggests that most of the genetic effect of fertility due to common SNPs is shared across sexes. However, using LDscore regression, we reject the null hypothesis of r_g=1 for AFB (P=0.007). A possible explanation of why we have not found more evidence for sex-genotype interactions may attributed to the fact that we analyzed only common variants and that we restrict our analysis to autosomal chromosomes. Moreover, our sex-specific meta-analysis may be underpowered to discover sex-specific loci.

When we compare Table 1 and 2, we note that in addition to the chr 5 locus for NEB, the chr

2 locus for AFB also shows a discrepancy between a sex-specific effect in the GWAS (women only) versus the (known) function of a candidate gene (AFF3). It would be premature to draw any firm conclusions since little is known about the role of AFF3 (chr 2) and EFNA5 (chr 5) in reproduction. For a substantial number of loci there are differences in the p-value between men and women, but the effect size suggests the association is present in both sexes. Only four loci seem to have a convincing null effect in men (rs1160544, rs10056247, rs2721195) or women (rs1316111). We would encourage functional follow-up studies on these points to further our understanding of human reproduction.

6. POLYGENIC SCORES PREDICTION

We performed out-of-sample prediction using cohorts for which we have direct access to genotypic data. We calculated polygenic scores for AFB and NEB, based on GWA meta-analysis results and used regression models to predict the same phenotypes in four independent cohorts: HRS, Lifelines, STR and TwinsUK. We ran ordinary least-squares (OLS) regression models and report the R^2 as a measure of goodness-of-fit of the model. In addition, we tested how well our polygenic scores for NEB could predict childlessness at the end of the reproductive period (using age 45 for women and 55 for men). Since age at first birth is observed only in parous women, we adopt an additional statistical model to account for censoring and selection. Finally, we also tested the predictive value of our polygenic scores for AFB for age at menarche (using TwinsUK) and age at menopause (using Lifelines).

6.1 Linear polygenic scores for AFB and NEB

We ran meta-analyses of the pooled AFB and NEB phenotypes, excluding each of the independent cohorts. Using these summary statistics, we constructed linear polygenic scores using the effect sizes from the original meta-analysis.[1] We constructed all scores using the software PLINK and PRSice[2,3] based on best call genotypes imputed to 1000G. For each phenotype, we calculated nine different scores using different p-value thresholds: 5e-08, 5e-07, 5e-06, 5e-05, 5e-04, 5e-03, 0.05, 0.5 and 1. Results are clumped using the genotypic data as a reference panel for LD structure.

We first regressed each phenotype on birthyear, its square and cubic to control for nonlinear trends in fertility, and the first 10 principal components, following the analysis plan distributed to the cohorts. If the cohort included both men and women, we included sex as a covariate in the regression models. Next, we regressed the residuals from the previous regression on the polygenic score. We performed a set of Ordinary Least Squares (OLS) regressions where we calculated R^2 as an indicator of goodness-of-fit of the regression model. For twin studies (STR and TwinsUK), we included only one MZ twin in the analysis and used clustered standard errors at the family level. To obtain 95% confidence intervals (CI) around the incremental $R^{2's}$, bootstrapping was performed with 1,000 repetitions.

The results of the polygenic score analyses are depicted in Supplementary Figure 2. The sample-size-weighted mean predictive power of the AFB score constructed with all SNPs is 0.9%, while the NEB score predictive power is 0.2%.

6.2 Linear polygenic scores for infertility

We used the score for NEB in an additional analysis to predict the probability to remain childless at the end of the reproductive period. Despite its limited predictive power for number of offspring, our analysis shows that an increase of one standard deviation of the polygenic score is associated with a decrease of around 9% in the probability to remain childless for women, with no significant differences among men (see Supplementary Table 21). The results are consistent across different cohorts.

6.3 Additional statistical models for censoring and selection

There are two limitations when studying the genetic determinants of AFB. The first is that this measurement is assessed only for men and women who ever became parents and does not take into consideration that a proportion of respondents are still at risk of having a child (i.e., did not have a child yet by the date of the interview) or will remain childless. This problem is commonly referred in the statistical literature as ‘right censoring’, since the outcome is not observed for all respondents, despite the fact that part of the respondent are still ‘at risk’ of

experiencing childbirth.[106] The second problem is statistical selection. Individuals with a measurement of AFB may be genetically different from individuals who remain childless. If childless individuals are different from the general population, the association results on AFB may be biased by selection problems. To investigate these two issues further, we estimated additional statistical models.

To control for right-censored data, we estimated semi-parametric Cox regression models[4] in which we estimate the effect of the polygenic score (PGS) on increasing the hazard of having a child conditional at each age. In other words, it is a model that estimates the impact of AFB PGS on yearly AFB, which will allow us to assess whether an increase in the AFB PGS is associated with a reduced risk of childbearing at each yearly age interval. This class of models takes into account censoring and is widely used to study fertility timing.[107] Our results show that the calculated PGS for AFB based on all SNPs is associated with an increased risk of childbearing at any age. The median AFB for men in the pooled sample is 28 and 26 for women. The hazard ratio of the PGS for AFB is 0.92 for women and 0.97 for men. This means that an increase of one standard deviation in the PGS is associated with an increase of 8% of AFB for women and 3% for men. Results for different cohorts and sex are depicted in Supplementary Table 22. Since this is a survival model that handles right-censoring (i.e., that the event of AFB did not occur by the observation time), the interpretation is that an increase in one standard deviation of the AFB PGS is associated with a reduction of 8% and 3% respectively for women and men in the hazard ratio of reproduction.

To control for selection, we estimated bivariate Heckman selection models in which we estimate the probability to be 'eligible' or at risk for AFB in a two-step procedure. Since we are interested in possible genetic differences among men and women who ever had children with respect to childless individuals, we used the PGS for NEB to model the probability to be at risk for AFB. The results from the Heckman selection models indicate slightly lower coefficients than OLS regression models but no substantial differences (see Supplementary Table 35 for details).

6.4 Linear prediction of age at menarche and age at menopause using AFB linear score

As an additional test, we examined whether the aforementioned PGS scores for AFB and NEB can predict related fertility traits such as age at menopause and age at menarche. We used the age at menopause measurement included in the Lifelines cohort. Age at menopause is measured with the question: "At which age have you had your last menstrual period?*"* We excluded women from the sample who reported to have had their last menstruation before age 30 or after age 60. The median age at natural menopause (ANM) in the sample is 45. Our results show that the PGS for AFB is associated with a later ANM. Since a substantive proportion of the sample of women in Lifelines is still in the pre-menopausal period, we estimated a proportional hazard model (Cox regression) in which we estimate ANM as a function of PGS for AFB. Our estimates indicate that having higher predisposition for AFB is associated with a later ANM. The hazard ratio estimate 0.97 indicates that an increase of one standard deviation of the PGS for AFB is associated with a decrease of ANM of about 3%. We used TwinsUK to model age at menarche. Our estimates indicate that an increase of one standard deviation on the PGS of AFB is associated with an increase of 0.06 years on age at menarche.[5] Results are depicted in Supplementary Table 23.

6.5 Association of menopause variants with AFB

We also examined whether menopause variants are associated with AFB. We calculated a PGS of age at menopause based on the recent GWAS results from Day et al. (2015)[108] and applied them to LifeLines and TwinsUK. The results for this analysis can be found in Supplementary Table 36 and shows no predictive power of the menopause genotype on AFB. This is consistent with the lookup exercise presented in S7.2, where none of our loci were significantly associated with age at Menopause. There might be several reasons why the LD score regression indicates a positive genetic correlation but we do not find evidence for specific loci. First, one or both of the studies may be underpowered and thus unable to identify a sufficiently large number of variants. Second, the correlation between the two traits may be spurious and mediated by other traits (e.g., age at menarche). We agree that it would be very interesting to pursue this in further research.

6.6 Discussion: The predictive power of polygenic scores

We acknowledge that the predictive power of the polygenic scores created from a meta-analysis of over 60 GWASs is only a fraction of what has been found in previous twin and family[1] and even GREML studies.[38] Several reasons have been noted for this 'missing heritability' problem,[109] including non-additive genetic effects,[49] epistatic effects,[110] rare variants and inflated estimates from twin studies due to *differential sharing of environmental factors in monozygotic and dizygotic twin pairs*.[111] Other factors that can explain the lower magnitude of effects are also plausible. Firstly, as we elaborate in Section S1.5, human reproductive behavior is not only biological, but also strongly related to environmental factors, and we should therefore not expect to find large independent genetic effects. We do not expect the PGS score to explain part of the variance attributable to environmental factors (i.e., the C and E in twin studies), but rather argue that these environmental factors are likely much stronger than genetic factors for these behavioral outcomes. As argued recently elsewhere,[39] it is vital to note that deep genetic analyses need to be united with strong and direct phenotypic measures. Although AFB and NEB are robustly measured, they inherently include a mix of voluntary (choice) and involuntary (infertility) measures. To overcome this problem, future innovations must unite rich genetic data with equally rich and precise phenotypic data collected precisely and continuously over several generations.

A second factor is that when studying phenotypes with behavioral component, GWAS discoveries are potentially limited by heterogeneity across birth cohorts and populations (e.g., countries) and particularly prone to gene-environment interaction. Fertility behavior has been demonstrated to be strongly environmentally determined and modified (e.g., by the introduction of effective contraception).[18] Although we examine gene-environment interaction across birth cohorts in Sweden in the Supplementary Note (section 10.1), in future research we will explore whether gene-environment interaction plays a role across birth cohorts and countries, with preliminary evidence suggesting that this is the case.[112] This is in line with recent research that has shown cohort differences in the genetic influence on smoking over time.[113]

7 GENETIC CORRELATIONS

7.1 Estimating genetic overlap using LD score regression

The estimates of the LD score regression reported in the main text was based on the LD-score regression method, which was developed by Bulik-Sullivan et al. (2015).[90] Here we describe in more detail how these estimates were computed and the genetic correlation we estimated between AFB and NEB and 27 publicly-available GWAS results (Supplementary Table 25 and graphed in Figure 3 in the main text). We focus on infertility traits, developmental traits, anthropometric traits, neuropsychiatric conditions and selected behavioral traits. LD score regression works even in the presence of sample overlap and only requires summary statistics and a reference panel from which to estimate SNP's "LD score", which measures the amount of genetic variation tagged by a SNP.

The approach requires GWAS summary statistics for all SNPs in our GWAS and a reference sample from which the LD can be estimated in order to estimate the LD score regression.[86] The method is written formally based on the following relationship:

$$E\left[z_{1j}z_{2j}\right] = \frac{\sqrt{N_1N_2}}{\mathrm{M}}\ell_j\rho_g + intercept,$$

Where z_{kj} is the z-score of SNP *j* from the GWAS of trait *k* (*k*=1,….,20), N_k is the sample size of the GWAS of trait k, $\ell_j$is the LD Score of SNP *j*, *M* the number of SNPs included in the GWAS, ρ_g the genetic covariance between traits *1* and *2*, with the regression intercept represented by *intercept*. The slope from the regression of $\hat{z}_{1j}z_{2j}$ on $\sqrt{N_1N_2}\ell_j$ can be used to estimate the genetic covariance between the two traits. We are also able to estimate the heritabilities of the two traits, h_{g1}^2 and h_{g2}^2 from the univariate LD score regressions of traits 1 and 2. It therefore follows that an estimate of the genetic correlation is:

$$\hat{r}_g = \frac{\widehat{\rho_g}}{\sqrt{\hat{h}_{g1}^2\hat{h}_{g2}^2}}$$

We use the file of LD scores computed by Finucane et al.[91] using genotypic data from a European-ancestry population (eur_w_ld_chr). LD Scores are computed with genotypes from the European-ancestry samples in the 1000 Genomes Project using only HapMap3 SNPs. We additionally follow the common convention of restricting our analyses to SNPs with MAF > 0.01, thus ensuring that all analyses are performed using a set of SNPs that are imputed with reasonable accuracy across all cohorts that contributed towards meta-analyses.

The standard errors (SEs) produced by the LDSC python software package uses a block jackknife over the SNPs. This influences the interpretation. Conventional standard errors are interpreted as measuring the variability of the estimate holding the covariates constant, but drawing on a new set of individuals. In this technique, SEs are interpreted as the variability of the estimate holding the sample constant, but drawing a new set of SNPs.

7.2 Estimating the genetic correlation between AFB and NEB

The negative relationship of late AFB with lower NEB[7,10,114] is well-established and consistent in advanced societies. Behavioral genetic models, based on twin or family studies show that this correlation is partially genetic, suggesting that natural selection favored a younger age at first birth over the Twentieth century.[1,38,115]

A recent study on genetic basis of fertility traits using molecular genetic data shows that common genetic variants influence NEB and AFB in a large sample of unrelated women.[38] Their results indicate a significant negative genetic correlation (r_g=–0.62, SE=0.27) between AFB and NEB. This finding implies that individuals with genetic predispositions for an earlier AFB had a reproductive advantage. We replicated the analysis of Tropf et al.[38] on a large sample of women from the Women General Health Study (WGHS, sample size N=40,120). We found a negative genetic correlation (r_g=–0.26, SE=0.13) between AFB and NEB. The results were limited to women and applied to a limited sample. We extend this work using LD score bivariate regression[86,90] on AFB and NEB on both men and women to identify the extent of cross-trait genetic correlation.

The LD score bivariate estimates indicate high negative correlation r_g=–0.66 (SE=0.03, p-value=$1.03x10^{-102}$) between AFB and NEB. This result is consistent both in men and women and is in line with the phenotypic correlation. Genetic correlation of fertility traits among women is slightly higher (r_g=–0.66, SE=0.04) than men (r_g=–0.58, SE=0.07). Overall these results show a considerable genetic overlap between NEB and AFB (as found in section 3). However, since the genetic overlap is statistically different from 1 for both men and women, these results indicate the existence of trait-specific genetic components.

7.3 Results: phenotypic correlations with human reproductive behavior

As discussed in the main text, we used information from 27 publicly available GWAS results to examine phenotypic correlations between AFB and NEB (Supplementary Table 25 and Figure 3 in the main text). These included: nine developmental traits, some of which are directly related to the reproductive span (age at menarche,[116] age at menopause,[117] Tanner stage,[118] age at voice breaking for males,[119] polycystic ovary syndrome (PCOS),[120] age at first sexual intercourse,[23] DZ Twinning,[121] birth length,[122] birth weight[123]), four behavioral traits (years of education,[76,79] cigarettes per day,[124] ever smoked,[124] age onset smoking[124]), seven personality and neuropsychiatric traits (neuroticism,[125] openness, schizophrenia,[126] bipolar disorder,[127] subjective well-being,[80] Alzheimer's disease,[128] autism[129]), four cardiometabolic traits (LDL cholesterol,[130] triglycerides,[130] type 2 diabetes,[131] fasting insulin levels[132]), and three anthropometric traits (BMI,[133] height,[87] waist-hip ratio[134]).

As shown in Fig. 3 and Supplementary Table 25 (P-values in bold indicate Bonferroni correction (P-value<0.05/27=$1.85x10^{-03}$)), AFB is positively correlated with years of education, age at menarche, age at menopause, age at voice breaking, age at first sexual intercourse and adult height, while it is negatively correlated with PCOS, adult BMI and waist-hip ratio, triglycerides, diabetes and fasting insulin level. Once multiple testing is controlled for, years of education and age at first sexual intercourse are the only traits significantly correlated with NEB (P-value<$2.25x10^{-03}$), and the direction is negative for both traits.

7.4 Discussion

7.4.1 Human development

AFB was shown to be positively correlated with the development measures of age at menarche, age at menopause, age at voice breaking and age at first sexual intercourse. A later age of menarche (AOM) has been associated with subfecundity and infertility in adulthood. A recent large cohort study of 73,107 women[135] demonstrated that women who reached menarche later than 15 years (compared to a reference group of girls with an AOM at 13 years) had a higher risk of infertility. Women younger than 11 years at AOM had lower odds of subfecundity and all results remained significant also after adjusting for women's age of pregnancy. Some studies, however, have also found a significant relationship between early AOM with diminished functional ovarian reserve later in life among infertile women.[136] There is also evidence of a small increased risk of endometriosis associated with early AOM.[137]

Stolk et al. (2012)[138] linked age at menopause to genes implicated in DNA repair and immune function. A recent study reported genetic correlations indicating shared aetiologies in both sexes between the timing of puberty and BMI, lipid levels, type 2 diabetes and cardiovascular disease.[139] Fertility timing has been positively associated with age at menarche and age at first intercourse. Although previous research has largely focused on identifying genes related to menopause and menarche that mark the end the beginning and end of the reproductive career, it is also possible that observed fertility (AFB, NEB) influences the subsequent age at menopause and ovarian aging. Exploring these overlaps and associations would be an interesting area for future research.

Results from a genetic study of age at first sexual intercourse (AFS) linked AFS to variation in pubertal timing, but also personality characteristics related to high risk-taking and low neuroticism.[23] We examine the link with AFS and neuropsychiatric disorders in a later section (Section 7.4.5).

7.4.2 Cardiometabolic traits

Having more AFB-increasing alleles was also significantly associated with a lower genetic scores for triglycerides, Type 2 Diabetes and fasting insulin level. Pregnancy for women results in considerable alterations in the cardiovascular system.[36] Reproductive events are associated with alterations in blood lipids and blood pressure and may therefore influence determinants of coronary heart disease. As with diabetes, there are mixed findings regarding the link between age at birth, parity and coronary heart disease (CHD). Some studies have linked the number of children and CHD risk with the prevalence lowest among those with 2 children with a linear increase with each additional child.[22] These researchers have argued that it is not the pregnancy per se that has a biological impact but rather that the lifestyle risk factors associated with childrearing leads to obesity which in turn results in increased CHD in both sexes. Yet, they maintain the argument that biological responses of pregnancy may have additional adverse effects in women.

Other studies attempted to elucidate the mechanisms linking multiparity to cardiovascular disease demonstrating that repeated pregnancies induce long-term changes in cardiovascular regulation in women due to the changes in vascular compliance and endothelium-dependent vasoconstriction, which in turn increase the risk for CHD in multiparous women.[36] A recent study related early puberty timing to higher risks for both Type 2 Diabetes and cardiovascular disease.[27] It may be however, that just as with the studies on GDM (gestational diabetes mellitus) described shortly, retrospective and cross-sectional approaches may have limitations related to selectivity and unobserved confounding factors. A prospective study in the US found that a younger age at menarche was only weakly associated with CHD and that nulliparous women only had a slightly higher rate of CHD compared to parous women. They also found no change in the risk with an increasing number of births or any association with the age at first birth concluding that there is no clear link between reproductive history and risk of CHD.[140] Further research is required to establish whether there is a true *causal* link and underlying genetic and biological mechanisms to explain the association between reproductive history and cardiometabolic traits.

There does, however, appear to be a link with the cardiometabolic traits that we measure in this study with infertility. Total cholesterol, triglycerides, LDL cholesterol levels and fasting insulin levels have been shown to be statistically higher in groups with endometriosis

compared to controls.[141] Endometriosis is estimated to occur in 5-10% of premenopausal women with ~50% experiencing problems conceiving.[34] A recent study also revealed a link between endometriosis and obesity-related traits.[142] Other studies have also linked the impact of maternal cholesterol metabolism to ovarian follicle development and fertility.[143] The role of the low-density lipoprotein receptor in cellular metabolism in inhibiting human reproduction has likewise been established.[144] Others have linked metabolic syndrome, which is a compilation of symptoms such as a high BMI (obesity), type 2 diabetes, dyslipidemia, and hypertension with an increased prevalence of infertility in men.[145]

A wide body of research links reproductive history to Type 2 Diabetes. Early studies found that nulliparity and multiparity or grand parity (5 or more children) was associated with higher levels of fasting glucose and insulin levels among nondiabetic women.[146–148] Multiparity has been associated with higher risks of cardiovascular disease in both women and men[27,149,150] and higher insulin resistance and type 2 diabetes.[149,151] Other research found that high parity was associated with insulin resistance and type 2 diabetes, which even after adjusting for confounders (socioeconomic, higher obesity, inflammatory markers) grand parity is associated with a 27% increased risk for diabetes (95% CI, 1.02-1.57).[151]

It is essential to note, however, that early cross-sectional and retrospective studies did not control for age, body size or socioeconomic status. Later cross-sectional studies that controlled for the abovementioned factors, continue to produce highly mixed results (for a review see ref [152]). A key limitation is that many of the previous studies lack universal GDM (gestational diabetes mellitus) screening and did therefore not measure preconception glycaemia or glucose intolerance during pregnancy. A systematic review and meta-analysis demonstrated that women who had gestational diabetes had a seven-fold greater risk of developing Type 2 Diabetes.[152] This suggests that once GDM status is accounted for, the direct parity effect will be very small or null. One the other hand, unobserved conditions such as PCOS, obesity or insulin resistance could in fact cause infertility (nulliparity) which would in turn lead to an underestimation of the association.

Gunderson et al. (2007)[153] examined whether childbearing increased the incidence of Type 2 Diabetes after preconception glycaemia and gestational glucose intolerance were controlled

for. They concluded that childbearing did not elevate the incidence of diabetes among those without GDM (i.e., normal glucose tolerance during pregnancy). It was GDM rather that was associated with the highest risk of developing diabetes, which remained even after controlling for family history of diabetes, preconception glycaemia and obesity. Another study using GDM screening found that a woman's age remained a strong predictor even after adjusting for prior GDM history, mirroring the general historical increase in GDM (and related levels of obesity) across time in certain groups. A logistic regression analysis also showed that mother's age at birth (OR 95% CI per 5 years 1.6–1.8) was significantly associated with GDM. Parity was not significantly associated with GDM and had no effect on the GDM increase over time.[154]

7.4.3 Anthropometric traits

A considerable body of literature links anthropometric traits (such adult height, BMI and increasingly waist-hip ratio) with fertility timing and success.[133,155] Anthropological research argue that shorter women may have more reproductive success because of the trade-off between investing in energy in growth or reproduction.[156] Moreover, taller women appear to become fertile at a later age (e.g., age at menarche) than shorter women, and women who have children at an early age reach a shorter adult height, which may result in a negative relationship between women's height and reproductive success.[155,157] The relationship between men's height and fertility is more complex. One paper revealed a curvilinear association between men's height and number of children in a nationally representative sample of US respondents.[158] Men of average height appear to have a higher reproductive success than either short or tall men. The relationship between height and number of children in advanced societies is not always negative. A recent paper showed that in the Netherlands – the country with the highest average population height – the relationship is the opposite.[155] A possible mechanism through which height may affect fertility is sexual selection and assortative mating. There is a certain degree of homogamy in anthropometric traits among spouses, even after controlling for a variety of socio-economic traits.[159,160]

BMI and waist-hip ratio (WHR) is another area of research often linked with fertility success, particularly in couples seeking ART treatment.[161] Both a very low and a very high BMI have been found to delay both the timing and number of children in both men and women.[162]
Waist-hip ratio measures body fat distribution and serves as a more accurate predictor of metabolic consequences independent of overall adiposity. A study locating new loci for WHR also found that seven of the loci exhibited marked sexual dimorphism, or in other words, that the genetic loci that modulate fat distribution have a stronger effect on WHR for women than men, suggesting strong gene-by-sex interactions.[163]

7.4.4 AFB and educational attainment

As described already in detail in Supplementary Note Section 1.5, the strong relationship between AFB and years of education is not surprising, since educational attainment is associated with higher AFB and a lower NEB in most advanced societies.[54,164] As discussed previously, the study of the relationship between higher educational attainment and reproduction has been a central focus within demography and related social sciences.[7,10,58,114,165] The majority of the research demonstrates that achieving higher education (particularly of women) operates to postpone AFB. Other studies have shown that fertility postponement may be related to higher cognitive ability,[166] but additional research is required to separate cognitive scores from social environment (e.g., family environment, social class). Others have found that after controlling for age, physical maturity and mother's education, there is a significant curvilinear relationship with intelligence and early sexual intercourse with both very low and very high intelligence operating as a protective factor against early sexual activity.[167] Further careful research in this area would be necessary to understand the relationship.

7.4.5 AFB, personality and neuropsychiatric disorders

The results of the LD score regression did not find any significant association with neuroticism, openness, schizophrenia, bipolar disorder, well-being, Alzheimer's disease or autism, so we will only touch upon this topic briefly. Personality has been demonstrated to be predictive of fertility intentions[20,168] and the timing of childbearing.[169,170] The finding that AFB is negatively correlated with neuroticism has also been found in previous non-genetic

studies linking AFB to personality traits.[171,172] A bidirectional effect between fertility and psychological development has likewise been documented.[168,173] This may suggest that the interaction between genetic and environment factors could be interpreted as genetic influences on fertility that have an effects on both fertility behavior and psychological outcomes. Since personality, educational attainment and cognitive ability are largely formed before individuals enter into their childbearing years, it is plausible that personality and cognitive traits are likely causal and precede fertility variables.[174] A recent study also demonstrated a genetic overlap between schizophrenia and AFB, showing a U-shaped relationship. The study confirmed that the schizophrenia risk profile score significantly predicted the relationship between maternal age and risk of schizophrenia in offspring.[16]

7.4.6 Smoking behavior

The strong negative correlation of a lower genetic risk of smoking (less cigarettes per day, lower chance to have ever smoked and later age of onset smoking) with a later AFB could operate via two mechanisms. First, it is well established that cigarette smoking has a detrimental biological effect on ovarian function and spermatozoa. There is an established link of a longer time to conception and decreased fertility with the increasing number of cigarettes smoked per day.[175] Other studies have linked cigarette smoking to infertility such as problems with preimplantation[176], shrinking the size and quality of oocytes[177], and abnormal spermatozoa by decreasing sperm motility in smokers.[178] A second potential mechanism is that the earlier onset of smoking and higher number of cigarettes smoked per day is also highly stratified by socioeconomic status. Smoking and low socioeconomic status are often linked to other environmental risk factors and a higher co-morbidity for other diseases.[179] Smoking is thus often a marker for structural, health and material disadvantage in addition to being concentrated in groups with the lowest levels of education.[180]

7.4.7 Limitations of LD score regression genetic correlations

Although LD score regression is a powerful tool to identify possible relationships between traits, we acknowledge that it does not allow us to establish causal directions or relationships or to adjust for potential mediating factors. The relationship between many of the traits discussed in this section is highly complex with potential bi-directional mechanisms. Further

studies are required to explore these relationships and establish whether the genetic risk related to AFB and NEB are either partially or fully mediated by other factors.

URLs.

The LDSC software is available at the website: http://www.github.com/bulik/ldsc;
GWAS summary statistics are available at the following websites: PGC (psychiatric) summary statistics, http://www.med.unc.edu/pgc/downloads; GIANT (anthropometric) summary statistics, http://www.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files; data on birth length, birth weight, Tanner stages have been contributed by EGG Consortium and has been downloaded from www.egg-consortium.org.; data on glycaemic traits have been contributed by MAGIC investigators and have been downloaded from www.magicinvestigators.org; DIAGRAM (type 2 diabetes) summary statistics, http://www.diagram-consortium.org/; SSGAC (educational attainment) summary statistics, http://www.thessgac.org/.

8. LOOK-UP OF LEAD SNPS IN AFB GWAS FOR AGE AT MENOPAUSE AND AGE AT MENARCHE

Following the results on genetic overlap with other phenotypes we tested – in a quasi-phenotype replication setting – whether the SNPs strongly associated with AFB in women are empirically plausible candidates SNPs for age at menarche and age at menopause. Our results reported in the previous section (Supplementary Note, Section 7) indicate a strong genetic correlation between these traits, suggesting a common genetic basis of reproductive behavior and reproductive life span.

Here we use a two-stage approach that has been applied in other contexts.[80,181] Since we are only looking at phenotypes measured among women with menarche and menopause, we started our analysis from the meta-analysis results from the AFB sample of women. In the first stage, we conduct a meta-analysis of age AFB excluding the cohorts that were part of the meta-analysis of the phenotype we intend to replicate. This step reduces the risk of overlap between the AFB sample from which the candidate SNPs are drawn and the sample used for testing the other phenotypes. We merged these SNPs with the publically available association results on the most recent GWAS on age at menarche[116] and age at menopause[117] from the Reprogen consortium website[b]. We first merged the two association files and dropped SNPs that are not present in both the files. We aligned the alleles and the effects direction using the software package EasyStrata.[182] We then selected the independent SNPs with a p-value$<1x10^{-5}$, using the clump procedure in PLINK[83], using the same settings described in section SI.2 (1000Kb window and LD threshold of $R^2>0.1$) to identify the most significant SNPs in associated regions included in both files. We define "prioritized SNP associations" as those that passed the Bonferroni correction for the number of SNPs tested ($P=0.05/122=4.10x10^{-4}$, both in age at menarche and age at menopause).

Supplementary Figure 36 shows the QQplots of the leading SNPs for AFB on age at Menarche (panel a) and age at menopause (panel b). Our results identified three SNPs after

[b] Data downloaded in November 2015 from http://www.reprogen.org/data_download.html

Bonferroni-correction that can be used as good candidates for age at menarche. We do not isolate any clear “candidate SNP” for age at menopause. The three SNPs that we identified (rs9589; rs6803222; rs9858889) are all in Chromosome 3. Two of them (rs9589; rs6803222) lie in proximity (<500Kb) of rs2777888, which has been identified as the strongest signal in our AFB GWAS.

9. BIOLOGICAL ANNOTATION

9.1. Identifying potentially causal variants

We followed the post-GWAS pipeline reported by Vaez et al[183] to shed light on the genomic context of the 12 independent genome-wide significant SNPs (Table 1 of the main text).

***In silico* sequencing:** For *in silico* sequencing, we used the data of the 1000 Genomes Project phase3 release of variant calls. This data set is based on the 20130502 sequence freeze and alignments. We used version v5a (Feb. 20th, 2015), and included only the 503 subjects of European ancestry (accessed April 5, 2016)[184]. The Variant Call Format (VCF)[185] files for regions of 1 Mb at either side of each lead SNP were downloaded using the Tabix software package.[186] Then, the r^2 between the lead SNPs and all other bi-allelic SNPs within the corresponding 2 Mb area was calculated as a metric of linkage disequilibrium (LD) using the Plink software package (v1.07).[83] All SNPs in LD with any of the lead SNPs were then annotated by ANNOVAR software[187] (version 1 Feb 2016, accessed April 9, 2016). We also used Sorting Intolerant From Tolerant (SIFT)[188] and Polymorphism Phenotyping (PolyPhen)[189] prediction scores to characterize the damaging impact of the nonsynonymous SNPs on the corresponding proteins. These scores were obtained from Ensembl release 83 (accessed April 11, 2016).[190]

In silico pleiotropy analysis

To identify any other trait or outcome associated with these 12 independent loci, we used the publicly available data of the National Human Genome Research Institute (NHGRI) GWAS Catalog (Catalog of Published Genome-Wide Association Studies).[191] We checked for pleiotropic effects of all lead SNPs as well as their linked variants (revealed in the previous phase of *in silico* sequencing) on other complex traits or diseases identified in previous GWAS studies and listed in the GWAS Catalog using ANNOVAR software[187] (version 1 Feb 2016, accessed April 9, 2016).

9.2. Gene-based GWAS analysis

We performed gene-based testing with the full GWAS set (~2.5M HapMap-imputed SNPs) of both phenotypes using VEGAS.[192,193] This software has the advantage of accounting for LD structure and the possibility to define a range beyond the gene bounds to include intergenic regions in the analysis. We defined a 50kb extra window surrounding the genes and considered every SNP for the gene-based analysis, ran the analyses per chromosome with up to 10^6 permutations and considered $P<2.5x10^{-06}$ as the threshold for significance (0.05/~20.000 genes).

9.3. eQTL and mQTL analyses

eQTL[194] and mQTL[195] analyses performed by the BIOS consortium have been described previously. The methods described in these papers are summarized below.

Genotype data

The BIOS consortium used samples from five Dutch cohorts; genotype QC and generation was described previously for each cohort: The Leiden Longevity Study,[196] The Rotterdam Study,[197] The LifeLines-DEEP cohort,[198] The Cohort on Diabetes and Atherosclerosis Maastricht (CODAM)[199] and The Netherlands Twin Register.[200] Genotype data were harmonized towards the Genome of the Netherlands (Genome of the Netherlands Consortium, 2014) (GoNL) using Genotype Hamonizer and subsequently imputed per cohort using Impute2 using the GoNL reference panel (v5). We removed SNPs with an imputation info-score below 0.5, a HWE *P*-value smaller than 10^{-4}, a call rate below 95% or a minor allele frequency smaller than 0.05.

9.3.2 RNA data preparation, sequencing and quantification

Total RNA from whole blood was deprived of globin using Ambions GLOBINclear kit and subsequently processed for sequencing using Illumina's Truseq version 2 library preparation kit. Paired-end sequencing of 2x50bp was performed using Illumina's Hiseq2000, pooling samples at 10 per lane, and aiming for >15M read pairs per sample. Finally, read sets per sample were generated using CASAVA, retaining only reads passing Illumina's Chastity Filter for further processing. The quality of the raw reads was checked using FastQC. The adaptors identified by FastQC (v0.10.1) were clipped using cutadapt (v1.1) applying default settings (min overlap 3, min length). Sickle (v1.200) was used to trim low quality ends of the

reads (min length 25, min quality 20). Read alignment was performed using STAR 2.3.0e. To avoid reference mapping bias all GoNL SNPs with MAF > 0.01 in the reference genome were masked. Read pairs with at most 8 mismatches, mapping to at most 5 positions were used. Mapping statistics from the BAM files were acquired through Samtools flagstat (v0.1.19-44428cd). The 5' and 3' coverage bias, duplication rate and insert sizes were assessed using Picard tools (v1.86). We estimated expression on the gene, exon, exon ratio and polyA ratio levels using Ensembl v.71 annotation (which corresponds to Gencode v.16). Overlapping exons (on either of the two strands) were merged into meta-exons and expression was quantified for the whole meta-exon. To this end, custom scripts were developed which uses coverage per base from coverageBed and intersectBed from the Bedtools suite (v2.17.0) and R (v2.15.1). This resulted in base counts per exon or meta-exon. Expression data was first normalized using Trimmed Mean of M-values (TMM). Then expression values were log2 transformed, probe and sample means were centered to zero. To correct for batch effects, principal component analysis (PCA) was run on the sample correlation matrix and the first 25 PCs were removed. We saw that removing these PCs resulted in highest number of eQTLs detected. To ascertain that none of these 25 PCs are under genetic control, we ran separate QTL mapping on each principal component and ensured that there were no SNPs associated with them. After QC[194] data was available from 2,116 samples.

9.3.3 Methylation data generation, mapping and normalization.

For the generation of genome-wide DNA methylation data, 500 ng of genomic DNA was bisulfite modified using the EZ DNA Methylation kit (Zymo Research, Irvine, California, USA) and hybridized on Illumina 450k arrays according to the manufacturer's protocols. The original IDAT files were extracted from the HiScanSQ scanner. We remapped the 450K probes to the human genome reference (HG19) to correct for inaccurate mappings of probes and identify probes that mapped to multiple locations on the genome. Next, we removed probes with a known SNP (GoNL, MAF > 0.01) at the single base extension (SBE) site or CpG site. Lastly, we removed all probes on the sex chromosomes, leaving 405.709 high quality methylation probes for the analyses. Methylation data was directly processed from IDAT files resulting from the Illumina 450k array analysis. After QC,[195], data was available from 3,841 samples.

9.3.4 eQTL and mQTL analysis

For each of the 12 SNPs identified in the GWAS, local (cis, exons/methylation sites < 1 MB from the SNP) and genome-wide (trans, exons/methylation sites > 5 MB from the SNP) effects were identified by computing Spearman rank correlations between SNPs and local or global exons/methylation sites. Bonferroni multiple testing correction was performed for the 12 SNPs tested ($P<2.5x10^{-06}$ for cis mQTL analysis, $P<1x10^{-08}$ for trans mQTL analysis, $P<1.2\ x10^{-06}$ for cis eQTL analysis, $P<1.3x10^{-08}$ for trans eQTL analysis). For each of the significant associations, the exons/methylation sites were selected, the strongest eQTLs were identified for these exons/methylation sites, and LD between these strongest eQTLs and the corresponding SNP identified in the GWAS were computed. LD was computed using BIOS genotypes (the genotypes used for eQTL and mQTL mapping).

9.4. Functional variant analysis using RegulomeDB

We used RegulomeDB[201] to identify variants amongst the 322 SNPs that reached $P<5x10^{-08}$ for association with AFB and/or NEB in the meta-analysis of GWAS that likely influence regulation of gene expression. RegulomeDB integrates results from RoadMap Epigenomics[202] and the ENCODE project.[203] SNPs that showed most evidence of being functional – defined as a RegulomeDB score <4 – were subsequently examined in more detail in terms of effects on gene expression (eQTLs) and their protein-binding capacity (Supplementary Supplementary Table 6).

9.4.1 Gene prioritization using four bioinformatics approaches

Potentially causal genes for the associations identified by GWAS were identified using four previously described bioinformatics tools: ToppGene,[204] Endeavour,[205] MetaRanker,[206] and DEPICT.[207] To this end, we first retrieved positional coordinates for all lead SNPs according to GRCh37/hg19 using Ensembl's BioMart. These coordinates were used to subsequently extract all genes located within ±40kb of lead SNPs using the UCSC Supplementary Notebrowser. The identified genes then served as input for ToppGene and Endeavour. Genes with established roles in fertility served as training genes in this procedure, i.e. *BRCA1*, *EGFR*, *ERBB2-4*, *HSD17B1*, *RBM5*, *ESR1*, *ESR2* and *FSHB*. All 10 genes were used in the

pooled and sex-specific analyses, while *ESR1*, *ESR2* and *FSHB* were not used in the analyses in data from men only, for biological reasons. For MetaRanker we provided SNPs that reached $P<5x10^{-04}$ and their chromosomal position as input, together with the previously mentioned set of training genes. Since ToppGene, Endeavour and MetaRanker are biased towards larger and well-described genes, we additionally performed a gene prioritization procedure using DEPICT.[207] All SNPs that reached $P<5x10^{-04}$ in the meta-analysis served as input, and information on prioritized genes, gene set enrichment, and tissue/cell type enrichment were extracted. Genes were subsequently prioritized that reached: 1) $P<0.05$ in DEPICT; or 2) $P<0.05$ in ToppGene, Endeavour and MetaRanker (Supplementary Tables 11, 12).

9.5. Functional network and enrichment analyses

DEPICT was additionally used to identify gene set, cell type and tissue enrichment analyses, using the GWAS-identified SNPs with $P<5x10^{-04}$ as input.[c] Due to the relatively small number of identified loci, DEPICT was only able to perform these analyses for AFB and NEB pooled, and AFB women.

To construct a functional association network, we combined five prioritized candidate gene sets into a single query gene set: closest genes to the lead SNPs, closest genes to the nonsynonymous SNPs in high LD ($r^2>0.50$) with the corresponding lead SNP, closest genes to other types of SNPs in very high LD ($r^2>0.80$) with the corresponding lead SNP, and expression probe gene names of cis, and trans eQTLs. The single query gene set was then used as input for the functional network analysis.[183] We applied the GeneMANIA algorithm together with its large set of accompanying functional association data.[208] We used the Cytoscape software platform,[209] extended by the GeneMANIA plugin (Data Version:

[c] We initially used a threshold of P<1E-5 for association with the respective outcomes in the meta-analyses of GWAS for SNPs to serve as input for the gene and tissue set enrichment analyses, as per the developers' recommendations.[206] We contacted the 1st author when this did not yield gene and tissue sets that were significantly enriched, and were advised to apply the slightly more lenient inclusion criterion of P<5E-4.

8/12/2014, accessed April 24, 2016).[210] All the genes in the composite network, either from the query or the resulting gene sets, were then used for functional enrichment analysis against Gene Ontology terms (GO terms)[211] to identify the most relevant GO terms using the same plugin.[210]

10. GENE-ENVIRONMENT INTERACTIONS

Previous research based on twin studies shows differential heritability of fertility behavior across birth cohorts.[212,213] With the exception of one recent mega-analysis[112] and a recent related study,[214] we are not aware of any study that examines variation at the molecular level to understand whether the genetic effect of AFB and NEB changes across birth cohort, level of education or other environmental factors. There is an implicit assumption that the genes associated with phenotypes are often constant across different historical, geographic or socio-economic groups.[39] In this section, we therefore examine gene-environment interaction by birth cohort and educational attainment.

As elaborated upon already in detail in Section 1.5, there has been considerable environmental variation over time and among cohorts in different historical periods that has undoubtedly influenced AFB and NEB. It is plausible, therefore, that there are differences across birth cohorts (time) since individuals born in different periods face diverse environmental conditions, such as the introduction and availability of effective contraception, sexual norms and diversity in factors that 'compete' with fertility, such as the expansion of educational attainment and labor force participation of women.[7]

This builds upon research that has examined changes across cohorts on the genetics of smoking. An early study adopted a twin design to demonstrate that genetic factors underlying smoking desistance were more important after the introduction of a restrictive legislation on smoking.[215] A related study also showed strong genetic influences on smoking of cohorts born in the 1920s, 1930s and 1950s, but not for those born in the 1940s and 1960s. They link these differences to changes in legislation prohibiting smoking in public places.[216] Using GREML methods and a modified DeFries-Fulker approach, a recent study likewise

demonstrated that there were cohort differences in the genetic influence on smoking, which increased over time.[113]

It may also be the case that the PGS for AFB and NEB is moderated by educational attainment. If the genetic association operates differently by the level of educational attainment, it would provide additional insight into understanding how fertility preferences and education are transmitted across generations. A recent study using the HRS in the US suggested that natural selection has taken place in contemporary societies and that there has been slow selection of lower educational attainment for both sexes.[214] In other words, the study argues that individuals endowed with genes predisposing them to more years of education are having fewer children and that natural selection (of those born from the 1930s to 1953) favors variants associated with less education. A commentary on this article[39] emphasizes four main reasons to be tentative about the conclusions that can be drawn. First, selection on education is weak and evolutionary changes are slow. Second, the PGS for educational attainment is likely associated with many other (non)cognitive traits. Third, socio-environmental, cultural and economic factors often override genetic factors for this phenotype. Fourth, 'years of education' is not a precise measurement and finally, that there may be mortality selection in the HRS sample of genotyped individuals, who have a higher socioeconomic status.[217]

10.1 Cohort analysis

We used the Swedish Twin Register (STR) to examine if the effect of a polygenic score (PGS) of AFB and NEB varies across birth cohort. We followed the analysis presented in the recent GWAS of education[218] and divide the sample into six groups based on their year of birth. Each group spans five birth years, with the oldest ranging from 1929-1933 and the youngest born between 1954- 1958. We then estimated the following regressions:

$$AFB_i = \beta_0 + \beta_1 PGS^{AFB}{}_i + \beta_2 Sex_i + \sum_{c=1}^{6} \gamma_1^c cohort_{ci} + \sum_{c=1}^{6} \gamma_2^c PGS^{AFB}{}_i \times cohort_{ci} + \sum_{k=1}^{10} \beta_k^{pc} PC^k{}_i + \varepsilon_i$$

$$NEB_i = \beta_0 + \beta_1 PGS^{NEB}{}_i + \beta_2 Sex_i + \sum_{c=1}^{6} \gamma_1^c cohort_{ci} + \sum_{c=1}^{6} \gamma_2^c PGS^{NEB}{}_i \times cohort_{ci} + \sum_{k=1}^{10} \beta_k^{pc} PC^k{}_i + \varepsilon_i$$

where i indicate individuals and k indexes principal components ( ) of the genetic data. We used a PGS standardized to have mean 0 and standard deviation 1 based on the GWAS meta-analysis results excluding the STR (details on how we constructed the PGS are available in Section 7 of the SI). The coefficients γ_2^c estimate whether there is an interaction between the PGS and an individual's birth cohort.

Supplementary Table 38 reports the estimated coefficient from these regressions. The results indicate a U-shaped trend in AFB and a linear decline in NEB, but do not provide any clear evidence of interaction effects between the PGS's and birth cohort. The only interaction coefficient that is significantly different from zero is the interaction between the PGS for NEB in the most recent birth cohort (those born 1954-1958). This analysis is a first descriptive attempt to examine GxE effects with birth cohorts. However, the PGSs are weighted by association coefficients of a GWAS where each cohort consists of individuals born in different years. Moreover, individual cohorts controlled for linear, quadratic and cubic trends in fertility behavior in their analysis. It would be informative to extend these analyses to more recent cohorts and contexts and refine the approach.

10.2 Educational attainment

We tested the interaction effects between educational level and the PGS of AFB and NEB in three different samples (LifeLines, STR and HRS). To ensure out of sample prediction, the PGS excluded each respective sample as required.

For each cohort, we estimated the following regressions[d]:

$$AFB_i = \beta_0 + \beta_1 PGS^{AFB}{}_i + \beta_2 Sex_i + \beta_3 education_i + \beta_4 PGS^{AFB}{}_i \times education_i + \sum_{k=1}^{10} \beta_k^{pc} PC^k{}_i + \varepsilon_i$$

$$NEB_i = \beta_0 + \beta_1 PGS^{NEB}{}_i + \beta_2 Sex_i + \beta_3 education_i + \beta_4 PGS^{NEB}{}_i \times education_i + \sum_{k=1}^{10} \beta_k^{pc} PC^k{}_i + \varepsilon_i$$

[d] For HRS, we estimated only a PGS for NEB, since AFB is not collected in that data.

Where $education_i$ is measured as years of education. Supplementary Table 39 reports the estimated coefficient from these regressions. The results indicate that years of education are positively associated with AFB in both LifeLines and STR, and negatively associated with NEB in the HRS. With the exception of NEB in the HRS, we found no evidence of GxE effects with education. We can therefore conclude that it appears that education does not appear to moderate the effect of the PGS for AFB and NEB.

11. ROBUSTNESS CHECKS

To estimate the robustness of our results for AFB, we conducted two additional analyses. First, we estimated how the coefficients change if we control for Educational Attainment (EA). Using data from deCODE, we ran an additional association analysis using the 10 loci that were genome-wide significant in the meta-analysis (p-value<$5x10^{-08}$). The analysis has been restricted to individuals born between 1910 and 1975, who also had data available on completed education. The total sample size is 42,187 (17,996 males and 24,191 females). The analysis is adjusted for sex, year of birth (linear, squared and cubic), interaction between sex and year of birth and the first 10 PCAs. Education is measured by years of education, ranging between 10 and 20 years. Supplementary Table 40 reports the association results before and after adjusting for educational attainment. Our analysis shows that the effect sizes shrink after including educational attainment as a covariate, with an average reduction of around 15%. We also estimated the effect of a polygenic risk score of AFB calculated from meta-analysis data excluding the deCODE cohort. The polygenic score remains highly significant. The effect of 1SD of the AFB score decreases from 0.19 years (69 days) without controlling for education to 0.16 years (59 days) when we control for years of education. To summarize, this analysis shows that the coefficients are robust to the inclusion of educational attainment in the model.

Second, we estimated how the coefficients change after controlling for Education Attainment (EA) and Age at First Sex using the UKBiobank (N=50,954). We ran two association models: the first follows the GWAS analysis plan with no additional covariates and the second added years of education and age at first sexual intercourse as covariates. The results are presented in Supplementary Table 41 and Supplementary Figure 37. Our analysis shows

that the effect sizes of our top hits are highly concordant (R^2=0.94). The inclusion of EA and AFS as covariates weakens the effect sizes on average by 40% and increases the p-value of the estimated coefficients. However, both EA and AFS have a significant genetic basis and are highly genetically correlated with AFB. Therefore, possible genetic pleiotropy may affect the results and capture a considerable proportion of the genetic effect. Nevertheless, 7 SNPs out of 10 tested, have a p-value<0.05 in the model that controls for EA and AFS. Overall, we interpret this additional analysis as a robustness test that confirm that the top hits from our meta-analysis are robust to the inclusion of the confounding factors of EA and AFS

12. POSITIVE SELECTION

We performed a Haploplotter analysis[219] to examine if lead SNPs and/or functional variants identified using RegulomeDB show evidence of positive selection. Three variants showed standardized integrated haplotype scores <-2 or >2, indicating that these variants represent the top 5% of signals in the population. These SNPs are: 1) rs7628058 on chromosome 3 for AFB, an eQTLs for *RBM6* in monocytes; 2) rs2247510 on chromosome 3 for AFB, an eQTL for *RBM6* and *HYAL3* in monocytes and binding site for a range of transcription factors; 3) rs2415984, the lead SNP in the chromosome 14 locus for NEB. Results are presented in Supplementary Table 42.

13. ADDITIONAL ACKNOWLEDGMENTS

ERC, ESRC/NCRM, NWO

The research leading to these results has received funding from the following awards to PI M.C. Mills, European Research Council (ERC) Consolidator Grant SOCIOGENOME (615603, www.sociogenome.com), Economic & Social Research Council (ESRC) UK, National Centre for Research Methods (NCRM) grant SOCGEN (www.ncrm.ac.uk/research/SoCGEN/) and NWO (Dutch National Science Organization) (VIDI grant 452-10-012).

SSGAC

This research was carried out under the auspices of the Social Science Genetic Association Consortium (SSGAC). The SSGAC seeks to facilitate studies that investigate the influence of

genes on human behavior, well-being, and social-scientific outcomes using large genome-wide association study meta-analyses. The SSGAC also provides opportunities for replication and promotes the collection of accurately measured, harmonized phenotypes across cohorts. The SSGAC operates as a working group within the CHARGE consortium. The SSGAC was supported by funding from the US National Science Foundation (EAGER: 'Workshop for the Formation of a Social Science Genetic Association Consortium'), a supplementary grant from the National Institute of Health Office of Behavioral and Social Science Research, the Ragnar Söderberg Foundation (E9/11), the Swedish Research Council (421-2013-1061), and the NIA/NIH through grants P01-AG005842, P01-AG005842-20S2, P30-AG012810, and T32-AG000186-23 to NBER and R01-AG042568-02 to the University of Southern California. Philipp Koellinger (co-PI of the SSGAC) gratefully acknowledges funding from the European Research Council (ERC consolidator grant 647648 EdGe). For further information and data access, see http://www.thessgac.org/."

1958BC-T1DGC and 1958BC-WTCCC2

DNA collection was funded by MRC grant G0000934 and cell-line creation by Wellcome Trust grant 068545/Z/02. This research used resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases, National Human Genome Research Institute, National Institute of Child Health and Human Development, and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. This study makes use of data generated by the Wellcome Trust Case-Control Consortium. A full list of investigators who contributed to generation of the data is available from the Wellcome Trust Case-Control Consortium website. Funding for the project was provided by the Wellcome Trust under the award 076113. The 1958 birth cohort data can be accessed via the UK Data Service (http://ukdataservice.ac.uk/).

23andMe

We would like to thank the research participants and employees of 23andMe for making this work possible. This work was supported by the National Human Genome Research Institute of the National Institutes of Health (grant number R44HG006981).

ABCFS

The ABCFS was supported by the National Health and Medical Research Council (NHMRC) of Australia, the New South Wales Cancer Council, the Victorian Health Promotion Foundation (Australia), the Inkster-Ross Memorial Fund of the University of Otago, and the US National Cancer Institute, National Institutes of Health, under Request for Application CA-95-003 as part of the Breast Cancer Family Registries (CFRs).

ALSPAC

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The Centre National de Génotypage (CNG) carried out DNA genotyping on the Illumina Human660W-Quad array, and genotypes were called with Illumina GenomeStudio supported by the Wellcome Trust (WT088806). The UK Medical Research Council and the Wellcome Trust (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. This work was also supported by the Medical Research Council Integrative Epidemiology Unit (MC_UU_12013/1-9). This publication is the work of the authors and they will serve as guarantors for the contents of this paper. ALSPAC summary data will be published on the data repository at data.bris.ac.uk. Please note that the study website contains details of all the data that is available through a fully searchable data dictionary. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees.

Amish

The Amish study was funded by the National Institutes of Health [U01 HL72515, U01 GM074518, R01 HL088119], with additional funding for CardioChip analysis provided by an American Heart Association Scientist Development grant [0830146N]. Genotyping of CardioChip was carried out in the Genomics Core at the University of Maryland, Baltimore with support from the Mid-Atlantic Nutrition and Obesity Research Center (National Institutes of Health [P30 DK072488]).

ASPS

The authors thank the staff and the participants of the ASPS for their valuable contributions. The authors thank Birgit Reinhart for her long-term administrative commitment and Ing Johann Semmler for the technical assistance at creating the DNA bank. The research reported in this article was funded by the Austrian Science Fond (FWF) grant number P20545-P05 and P13180. The Medical University of Graz supports the databank of the ASPS. ASPS data can be accessed upon request.

BASEII

BASE-II was funded by the German Federal Ministry of Education and Research (BMBF) and has been formally divided into four subprojects: “Psychology & Project Coordination and Database” (Max Planck Institute for Human Development [MPIB], grant number 16SV5837), “Survey Methods and Social Science” (German Institute for Economic Research and Socioeconomic Panel [SOEP/DIW], grant number 16 SV5537), Medicine and Geriatrics (Charité – Universitätsmedizin, Berlin [Charité], grant number 16SV5536K), and “Molecular Genetics” (Max Planck Institute for Molecular Genetics, now University of Lübeck [MPIMG-ULBC], grant number 16SV5538).

External scientists can apply to the Steering Committee of BASE-II for data access. Although the data are available for other parties are scientific data and not personal contact data, the scientific data are subject to a security level as if they were personal data to ensure that the BASE-II Steering Committee sufficiently protects the large volume of data collected from each BASE-II participant. All existing variables are documented in a handbook. Contact: Katrin Schaar, scientific coordinator, schaar@mpib-berlin.mpg.de.

BIOS

The Biobank-based Integrative Omics Study (BIOS) consortium is funded by the Biobanking and Biomolecular Research Infrastructure (BBMRI-NL, NWO project 184.021.007).

BMES (Blue Mountains Eye Study) cohort

The Blue Mountains Eye Study (BMES) was supported by the Australian National Health & Medical Research Council (NHMRC), Canberra Australia (NHMRC project grant IDs 974159, 211069, 302068, and Centre for Clinical Research Excellence in Translational

Clinical Research in Eye Diseases, CCRE in TCR-Eye, grant ID 529923). The BMES GWAS and genotyping costs was supported by Australian NHMRC, Canberra Australia (NHMRC project grant IDs 512423, 475604 and 529912), and the Wellcome Trust, UK as part of Wellcome Trust Case Control Consortium 2 (A Viswanathan, P McGuffin, P Mitchell, F Topouzis, P Foster, grant IDs 085475/B/08/Z and 085475/08/Z)

Bruneck Study

This Bruneck Study is part of the excellence initiative (Competence Centers for Excellent Technologies - COMET) of the Austrian Research Promotion Agency FFG: Research Center of Excellence in Vascular Ageing - Tyrol, VASCage (K-Project No. 843536) funded by the BMVIT, BMWFW, Wirtschaftsagentur Wien and Standortagentur Tirol. There is no public access to the Bruneck data base.

CARLA

This study was funded by a grant from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) as part of the Collaborative Research Center 598 ‘‘Heart failure in the elderly–cellular mechanisms and therapy’’ at the Medical Faculty of the Martin-Luther-University Halle-Wittenberg; by a grant of the Wilhelm-Roux Programme of the Martin-Luther-University Halle-Wittenberg; by the Federal Employment Office; and by the Ministry of Education and Cultural Affairs of Saxony-Anhalt.
The study was in accordance with the declaration of Helsinki. All participants gave their written informed consent. The study was approved by the local ethic commission of the Medical Faculty of the Martin-Luther-University Halle-Wittenberg.

CHAP

We acknowledge NIH grants R01AG11101 and R01AG030146 for our cohort.

CHS

Cardiovascular Health Study: This CHS research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, and R01HL120393

with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org.
The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center.
The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Cilento Study
The Cilento study was supported by the Italian Ministry of Education Universities and Research (Interomics Flagship Project, PON03PE_00060_7), FP6 (Vasoplus-037254), the Assessorato Ricerca Regione Campania, the Fondazione con il SUD (2011-PDR-13), and the Istituto Banco di Napoli - Fondazione to MC. We address special thanks to the populations of Cilento for their participation in the study.

CoLaus
The CoLaus/PsyCoLaus study was supported by four grants of the Swiss National Science Foundation (#105993, 118308, 139468 and 122661), two unrestricted grants from GlaxoSmithKline as well as by the Faculty of Biology and Medicine of the University of Lausanne.

COPSAC2000
We greatly acknowledge the private and public research funding allocated to COPSAC and listed on www.copsac.com, with special thanks to The Lundbeck Foundation; Danish State Budget; Danish Council for Strategic Research; The Danish Council for Independent Research and The Capital Region Research Foundation as core supporters. The funding agencies did not have any influence on study design, data collection and analysis, decision to publish or preparation of the manuscript. No pharmaceutical company was involved in the study. We gratefully express our gratitude to the children and families of the COPSAC2000

cohort study for all their support and commitment. We also acknowledge and appreciate the unique efforts of the COPSAC research team. We greatly acknowledge the private and public research funding allocated to COPSAC and listed on www.copsac.com, with special thanks to The Lundbeck Foundation; Danish State Budget; Danish Council for Strategic Research; The Danish Council for Independent Research and The Capital Region Research Foundation as core supporters. The funding agencies did not have any influence on study design, data collection and analysis, decision to publish or preparation of the manuscript. No pharmaceutical company was involved in the study. We gratefully express our gratitude to the children and families of the COPSAC2000 cohort study for all their support and commitment. We also acknowledge and appreciate the unique efforts of the COPSAC research team, the families participating in the COPSAC cohort for their effort and commitment. The authors also thank the COPSAC study team. COPSAC is funded by private and public research funds, all of which are listed on the COPSAC website (www.copsac.com; see URLs). The Lundbeck Foundation, the Pharmacy Foundation of 1991, the Augustinus Foundation, the Danish MRC and The Danish Pediatric Asthma Centre provided core support for COPSAC. The funding agencies did not have any role in study design, data collection and analysis, the decision to publish or preparation of the manuscript.

CROATIA cohorts

We would like to acknowledge the invaluable contributions of the recruitment team in Korcula, the administrative teams in Croatia and Edinburgh and the people of Korcula. The CROATIA-Korcula study was funded by grants from the Medical Research Council (UK), European Commission Framework 6 project EUROSPAN (Contract No. LSHG-CT-2006-018947), FP7 project BBMRI-LPC (grant 313010), Ministry of Science, Education and Sports of the Republic of Croatia (grant 108-1080315-0302) and the Croatian Science Foundation (grant 8875). External researchers who wish to obtain access to CROATIA-Korcula's data or EA2 results may contact Ozren Polasek, ozren.polasek@mefst.hr.

deCODE

All deCODE collaborators in this study are employees of deCODE Genetics/Amgen, Inc. External researchers who wish to obtain access to data or EA2 results may contact Gudmar Thorleifsson gudmar.thorleifsson@decode.is.

DESIR

The D.E.S.I.R. Study Group. INSERM CESP U1018: B. Balkau, M-A. Charles, P. Ducimetière, E. Eschwège; INSERM U367: F. Alhenc-Gelas; CHU D'Angers: Y. Gallois, A. Girault; Bichat Hospital: F. Fumeron, M. Marre; CHU de Rennes: F. Bonnet; UMR8090, LILLE: P. Froguel; Centres d'Examens de Santé: Alençon, Angers, Caen, Chateauroux, Cholet, Le Mans, Tours; Institute de Recherche Médecine Générale: J. Cogneau; General practitioners of the region; Institute inter-Regional pour la Santé: C. Born, E. Caces, M. Cailleau, J.G. Moreau, F. Rakotozafy, J. Tichet, S. Vol.

The D.E.S.I.R. study has been financed by INSERM contracts with CNAMTS, Lilly, Novartis Pharma and Sanofi-Aventis; by INSERM (Réseaux en Santé Publique, Interactions entre les déterminants de la santé, Cohortes Santé TGIR 2008), the Association Diabète Risque Vasculaire, the Fédération Française de Cardiologie, La Fondation de France, ALFEDIAM, ONIVINS, Ardix Medical, Bayer Diagnostics, Becton Dickinson, Cardionics, Merck Santé, Novo Nordisk, Pierre Fabre, Roche, Topcon.

EGCUT

This study was supported by the Estonian Research Council (grant IUT20-60), the Development Fund of the University of Tartu (grant SP1GVARENG), EU structural support through Archimedes Foundation, grant no: 3.2.1001.11-0033, EU 7FP grant 278913, and H2020 grants 633589, 676550, 654248. Reference: [PMID: 24518929] Leitsalu et al, Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. Int J Epidemiol. 2014 Feb 11

EPIC-Norfolk

EPIC-Norfolk is supported by programme grants from the Medical Research Council (MRC) [G9502233; G0401527] and Cancer Research UK [C864/A8257]. JHZ, KKO and NJW are supported by MRC Unit programme grants [MC_UU_12015/1 and MC_UU_12015/2]. Further information about the cohort can be found at http://www.epic-norfolk.org.uk/

ERF

The ERF study as a part of EUROSPAN (European Special Populations Research Network) was supported by European Commission FP6 STRP grant number 018947 (LSHG-CT-2006-01947) and also received funding from the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F4-2007-201413 by the European Commission under the programme "Quality of Life and Management of the Living Resources" of 5th Framework Programme (no. QLG2-CT-2002-01254). High-throughput analysis of the ERF data was supported by joint grant from Netherlands Organization for Scientific Research and the Russian Foundation for Basic Research (NWO-RFBR 047.017.043). Exome sequencing analysis in ERF was supported by the ZonMw grant (project 91111025). We are grateful to all study participants and their relatives, general practitioners and neurologists for their contributions and to P. Veraart for her help in genealogy, J. Vergeer for the supervision of the laboratory work and P. Snijders for his help in data collection. Najaf Amin is supported by the Netherlands Brain Foundation (project number F2013(1)-28). The ERF study genome-wide array data and phenotype data (age and gender) is archived in European Genome-Phenome Database (EGA). The study is archived in the DAC named Erasmus Rucphen Family Study with the accession code: EGAS00001001134. Researchers who wish to use other phenotypic data of the Erasmus Rucphen Family Study must seek approval from the management team of the Erasmus Rucphen Family study. They are advised to contact the study PI, professor Cornelia van Duijn (c.vanduijn@erasmusmc.nl).

Dortmund Health Study DHS

DHS (Dortmund Health Study) - The collection of sociodemographic and clinical data in the Dortmund Health Study was supported by the German Migraine & Headache Society (DMKG) and by unrestricted grants of equal share from Almirall, Astra Zeneca, Berlin Chemie, Boehringer, Boots Health Care, Glaxo-Smith-Kline, Janssen Cilag, McNeil Pharma, MSD Sharp & Dohme and Pfizer to the University of Muenster. Blood collection in the Dortmund Health Study was done through funds from the Institute of Epidemiology and Social Medicine University of Muenster. Genotyping for the Human Omni Chip was supported by the German Ministry of Education and Research (BMBF, grant no. 01ER0816). Researchers interested in using DHS data are required to sign and follow the terms of a Cooperation Agreement that includes a number of clauses designed to ensure protection of

privacy and compliance with relevant laws. For further information, contact Klaus Berger (bergerk@uni-muenster.de).

Finnish Twin Cohort

Phenotype data collection and genotyping in the twin cohort have been supported by the Wellcome Trust Sanger Institute, ENGAGE – European Network for Genetic and Genomic Epidemiology, FP7-HEALTH-F4-2007, grant agreement number 201413, Academy of Finland (grants 265240, 263278 to JKaprio), and Global Research Awards for Nicotine Dependence (GRAND) to JK.

FINRISK

This study was supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (grant numbers 213506, 129680), the Academy of Finland (grant numbers 139635, 129494, 136895, 263836 and 141054), the Sigrid Juselius Foundation , the Paulo foundation, the Finnish Medical Foundationand ENGAGE – European Network for Genetic and Genomic Epidemiology, FP7-HEALTH-F4-2007, grant agreement number 201413 and The Finnish Foundation for Cardiovascular Research.

Generation R

The Generation R Study is conducted by the Erasmus Medical Center in close collaboration with the School of Law and Faculty of Social Sciences of the Erasmus University Rotterdam, the Municipal Health Service Rotterdam area, Rotterdam, the Rotterdam Homecare Foundation, Rotterdam and the Stichting Trombosedienst & Artsenlaboratorium Rijnmond (STAR-MDC), Rotterdam. We gratefully acknowledge the contribution of children and parents, general practitioners, hospitals, midwives and pharmacies in Rotterdam. The study protocol was approved by the Medical Ethical Committee of the Erasmus Medical Centre, Rotterdam. Written informed consent was obtained from all participants. The general design of Generation R Study is made possible by financial support from the Erasmus Medical Center, Rotterdam, the Erasmus University Rotterdam, the Netherlands Organization for Health Research and Development (ZonMw), the Netherlands Organisation for Scientific Research (NWO), the Ministry of Health, Welfare and Sport and the Ministry of Youth and Families. Vincent W. Jaddoe received an additional grant from the Netherlands Organization

for Health Research and Development (VIDI 016.136.361) and a European Research Council Consolidator Grant (ERC-2014-CoG-648916). The generation and management of GWAS genotype data for the Generation R Study were done at the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Netherlands. We would like to thank Karol Estrada, Dr. Tobias A. Knoch, Anis Abuseiris, Luc V. de Zeeuw, and Rob de Graaf, for their help in creating GRIMP, BigGRID, MediGRID, and Services@MediGRID/D-Grid, (funded by the German Bundesministerium fuer Forschung und Technology; grants 01 AK 803 A-H, 01 IG 07015 G) for access to their grid computing resources. We thank Mila Jhamai, Manoushka Ganesh, Pascal Arp, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters for his help in creating, managing and QC of the GWAS database. Also, we thank Karol Estrada for their support in creation and analysis of imputed data. J.F.F. has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 633595 (DynaHEALTH).

GENOA

GENOA (Genetic Epidemiology Network of Arteriopathy): Support for GENOA was provided by the National Heart, Lung and Blood Institute (HL119443, HL118305, HL054464, HL054457, HL054481, HL071917 and HL87660) of the National Institutes of Health. Genotyping was performed at the Mayo Clinic (Stephen T. Turner, MD, Mariza de Andrade PhD, Julie Cunningham, PhD). We thank Eric Boerwinkle, PhD and Megan L. Grove from the Human Genetics Center and Institute of Molecular Medicine and Division of Epidemiology, University of Texas Health Science Center, Houston, Texas, USA for their help with genotyping. We would also like to thank the families that participated in the GENOA study. Data Access: GENOA (Genetic Epidemiology Network of Arteriopathy): In accordance with the informed consents of the GENOA study, we provide individual-level genotype and phenotype data to GENOA investigators and collaborators. To collaborate with GENOA investigators, please contact Sharon L.R. Kardia (skardia@umich.edu). We fully welcome collaboration with researchers that would like to include the GENOA sample in their analyses. We can allow transfer of individual-level data with an appropriate Data Transfer Agreement.

GOYA

The Danish National Research Foundation established the Danish Epidemiology Science Centre, which initiated and created the Danish National Birth Cohort. The cohort is a result of a major grant from this Foundation. Additional support for the Danish National Birth Cohort was obtained from the Pharmacy Foundation, the Egmont Foundation, the March of Dimes Birth Defects Foundation, and the Augustinus Foundation. Genotyping for the GOYA Study within the Danish National Birth Cohort was funded by the Wellcome Trust (Grant ref: 084762MA).

HBCS

We thank all study participants as well as everybody involved in the Helsinki Birth Cohort Study. Helsinki Birth Cohort Study has been supported by grants from the Academy of Finland, the Finnish Diabetes Research Society, Folkhälsan Research Foundation, Novo Nordisk Foundation, Finska Läkaresällskapet, Signe and Ane Gyllenberg Foundation,University of Helsinki, Ministry of Education, Ahokas Foundation, Emil Aaltonen Foundation.

Health 2000

The Health 2000 Study was mainly funded from the budget of the National Institute for Health and Welfare (THL). Additional funding was received from the Finnish Centre for Pensions, the Social Insurance Institution of Finland, the Local Government Pensions Institution, the National Research and Development Centre for Welfare and Health, the Finnish Dental Association, the Finnish Dental Society, Statistics Finland, the Finnish Institute for Occupational Health, The Finnish Work Environment Fund, the UKK Institute for Health Promotion Research and the Occupational Safety and Health Fund of the State Sector. The data used for this study can be made available on request to the Health 2000/2011 scientific committee according to the ethical and research guidelines (www.terveys2011.info/aineisto) as well as Finnish legislation.

Health ABC

The Health ABC Study was supported by NIA contracts N01AG62101, N01AG62103, and N01AG62106 and, in part, by the NIA Intramural Research Program. The genome-wide association study was funded by NIA grant 1R01AG032098-01A1 to Wake Forest University

Health Sciences and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C. This study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, Md. (http://biowulf.nih.gov).

HRS

HRS (Health and Retirement Study): HRS is supported by the National Institute on Aging (NIA U01AG009740). The genotyping was funded separately by the National Institute on Aging (RC2 AG036495, RC4 AG039029). Our genotyping was conducted by the NIH Center for Inherited Disease Research (CIDR) at Johns Hopkins University. Genotyping quality control and final preparation of the data were performed by the Genetics Coordinating Center at the University of Washington. Data Access:
HRS (Health and Retirement Study): Genotype data can be accessed via the database of Genotypes and Phenotypes (dbGaP, http://www.ncbi.nlm.nih.gov/gap, accession number phs000428.v1.p1). Researchers who wish to link genetic data with other HRS measures that are not in dbGaP, such as fertility data, must apply for access from HRS. See the HRS website (http://hrsonline.isr.umich.edu/gwas) for details.

HTO

We thank all the families who contributed to this study. Phenotyping and genotyping of the HTO cohort was funded by the Wellcome Trust, the UK Medical Research Council and the British Heart Foundation. Data are available upon request from the Principal Investigator, Bernard Keavney (bernard.keavney@manchester.ac.uk).

INGI-CARL

We thank Martina La Bianca and Angela D'Eustacchio for technical support. We are very grateful to the municipal administrators for their collaboration on the project and for logistic support. We would like to thank all participants to this study.

INGI-Val Borbera

We thank all the participants to the project, the San Raffaele Hospital MDs who contributed to clinical data collection, prof. Clara Camaschella who coordinated the data collection, Corrado Masciullo and Massimiliano Cocca for the database informatics.
The research was supported by funds from Compagnia di San Paolo, Torino, Italy; Fondazione Cariplo, Italy; Telethon Italy; Ministry of Health, Ricerca Finalizzata 2008 and 2011-2012 and Public Health Genomics Project 2010.

KORA F3
The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank all the study participants, all members of staff of the Institute of Epidemiology II and the field staff in Augsburg who planned and conducted the study.

LBC1921 and LBC1936
We thank the cohort participants and team members who contributed to these studies. Phenotype collection in the Lothian Birth Cohort 1921 was supported by the UK's Biotechnology and Biological Sciences Research Council (BBSRC), The Royal Society, and The Chief Scientist Office of the Scottish Government. Phenotype collection in the Lothian Birth Cohort 1936 was supported by Age UK (The Disconnected Mind project). Genotyping of the cohorts was funded by the BBSRC. The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the BBSRC and Medical Research Council (MRC) is gratefully acknowledged.

LIFELINES
Lifelines is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviors of 167,729 persons

living in the North of The Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioral, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multi-morbidity and complex genetics.[220,221]

The Lifelines Cohort Study, and generation and management of GWAS genotype data for the Lifelines Cohort Study is supported by the Netherlands Organization of Scientific Research NWO (grant 175.010.2007.006), the Ministry of Economic Affairs, the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the Northern Netherlands Collaboration of Provinces (SNN), the Province of Groningen, University Medical Center Groningen, the University of Groningen, Dutch Kidney Foundation and Dutch Diabetes Research Foundation. The authors wish to acknowledge the services of the Lifelines Cohort Study, the contributing research centers delivering data to Lifelines, and all the study participant. Data availability: Lifelines is a facility that is open for all researchers. Information on application and data access procedure is summarized on www.lifelines.net.

Longevity

National Institutes of Health (AG028872, CA164468 and DA033788 to A.B., AG042188 to G.A., AG021654-01 and AG-18728-02A1 to N.B.) and the Glenn Center for the Biology of Human Aging.

MCTFR

MCTFR acknowledges support by the National Institutes of Health under award numbers R37DA005147, R01AA009367, R01AA011886, R01DA013240, and R01MH066140.

MESA

MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-001079, UL1-TR-000040, and DK063491. Funding for SHARe genotyping was provided by NHLBI Contract N02-HL-64278. Genotyping was performed at Affymetrix (Santa Clara, California, USA)

and the Broad Institute of Harvard and MIT (Boston, Massachusetts, USA) using the Affymetrix Genome-Wide Human SNP Array 6.0.

MoBa

MoBa (The Norwegian Mother and Child Cohort Study of NIPH) – the genotyping and analyses were supported by the grants from: Jane and Dan Olsson Foundations (Gothenburg, Sweden), Swedish Medical Research Council (2015-02559), Norwegian Research Council/FUGE (grant no. 151918/S10; FRI-MEDBIO 249779) and Swedish Medical Society (SLS 2008-21198), Swedish government grants to researchers in the public health service (ALFGBG-507701).

The Norwegian Mother and Child Cohort Study is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research, NIH/NIEHS (contract no N01-ES-75558), NIH/NINDS (grant no.1 UO1 NS 047537-01 and grant no.2 UO1 NS 047537-06A1). We are grateful to all the participating families in Norway who take part in this on-going cohort study.

MrOS Sweden

MrOS Sweden was funded by the Swedish Research Council, the Swedish Foundation for Strategic Research, the ALF/LUA research grant in Gothenburg, the Lundberg Foundation, the Torsten and Ragnar Söderberg's Foundation and the Novo Nordisk Foundation.

NEO

The authors of the NEO study thank all individuals who participated in the Netherlands Epidemiology in Obesity study, all participating general practitioners for inviting eligible participants and all research nurses for collection of the data. We thank the NEO study group, Pat van Beelen, Petra Noordijk and Ingeborg de Jonge for the coordination, lab and data management of the NEO study. The genotyping in the NEO study was supported by the Centre National de Génotypage (Paris, France), headed by Jean-Francois Deleuze. The NEO study is supported by the participating Departments, the Division and the Board of Directors of the Leiden University Medical Center, and by the Leiden University, Research Profile Area Vascular and Regenerative Medicine. Dennis Mook-Kanamori is supported by Dutch Science Organization (ZonMW-VENI Grant 916.14.023).

NESDA

The infrastructure for the NESDA study is funded through the Geestkracht programme of the Dutch Scientific Organization (ZON-MW, grant number 10-000-1002) and matching funds from participating universities and mental health care organizations. Genotyping in NESDA was funded by the Genetic Association Information Network (GAIN) of the Foundation for the US National Institutes of Health. Statistical analyses were carried out on the Genetic Cluster Computer (http://www.geneticcluster.org), which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003) along with a supplement from the Dutch Brain Foundation.

Data availability

Data are available upon request from the NESDA data management bureau.

Nurses' Health Study (NHS) and Health Professionals Follow-up Study (HPFS)

Supported by grants UM1 CA186107, UM1 CA167552, DK091718, HL071981, HL073168, CA87969, CA49449, CA055075, HL34594, HL088521, U01HG004399, DK080140, 5P30DK46200, U54CA155626, DK58845, U01HG004728-02, EY015473, DK70756 and DK46200 from the National Institutes of Health, with additional support for genotyping from Merck Research Laboratories, North Wales, PA.

NTR (Netherlands Twin Register)

Netherland Twin Register: Funding was obtained from the Netherlands Organization for Scientific Research (NWO) and The Netherlands Organisation for Health Research and Development (ZonMW) grants 904-61-090, 985-10-002, 904-61-193,480-04-004, 400-05-717, Addiction-31160008, Middelgroot-911-09-032, Spinozapremie 56-464-14192, Biobanking and Biomolecular Resources Research Infrastructure (BBMRI –NL, 184.021.007). VU Institute for Health and Care Research (EMGO+); the European Community's Seventh Framework Program (FP7/2007-2013), ENGAGE (HEALTH-F4-2007-201413); the European Research Council (ERC Advanced, 230374, ERC Starting grant 284167), Rutgers University Cell and DNA Repository (NIMH U24 MH068457-06), the Avera Institute, Sioux Falls, South Dakota (USA) and the National Institutes of Health (NIH, R01D0042157-01A, MH081802; R01 DK092127-04, Grand Opportunity grants 1RC2

MH089951). Part of the genotyping and analyses were funded by the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health. Computing was supported by BiG Grid, the Dutch e-Science Grid, which is financially supported by NWO.

Nurses' Health Study and Health Professionals Follow-up Study
We need to acknowledge support of the following grants from the National Institutes of Health: UM1 CA186107; R01 CA49449; UM1 CA167552; Nurses' Health Study and Health Professionals Follow-up Study

OGP Ogliastra Genetic Park

Funding: Grant from the Italian Ministry of Education, University and Research (MIUR) n°: 5571/DSPAR/2002

ORCADES
The Orkney Complex Disease Study (ORCADES) was supported by the Chief Scientist Office of the Scottish Government, the Royal Society, the MRC Human Genetics Unit, Arthritis Research UK and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. We would like to acknowledge the invaluable contributions of Lorraine Anderson and the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney. Details regarding data access are available at the ORCADES website (http://www.orcades.ed.ac.uk/orcades/orcades2.html).

QIMR
Funding was provided by the Australian National Health and Medical Research Council (241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, 552498), the Australian Research Council (A7960034, A79906588, A79801419, DP0770096, DP0212016, DP0343921), the FP-5 GenomEUtwin Project (QLG2-CT-2002-01254), and the U.S. National Institutes of Health (NIH grants AA07535, AA10248, AA13320, AA13321, AA13326, AA14041, DA12854, MH66206). A portion of the genotyping on which the QIMR study was based (Illumina 370K scans) was carried out at the Center for Inherited Disease Research, Baltimore (CIDR), through an access award to the authors' late colleague Dr. Richard Todd (Psychiatry, Washington University School of Medicine, St Louis). Imputation was carried out on the Genetic Cluster Computer, which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003).

S.E.M., is supported by the Australian Research Council (ARC) Fellowship Scheme. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Researchers interested in using QIMR data can contact Nick Martin (Nick.Martin@qimrberghofer.edu.au).

Rotterdam Study

The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters for their help in creating the GWAS database, and Karol Estrada and Maksim V. Struchalin for their support in creation and analysis of imputed data. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. Some of the statistical analyses were carried out on the Genetic Cluster Computer (http://www.geneticcluster.org) which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003 PI: Posthuma) along with a supplement from the Dutch Brain Foundation and the VU University Amsterdam. Cornelius A. Rietveld gratefully acknowledges funding from the Netherlands Organization for Scientific Research (NWO Veni grant 016.165.004). Researchers who wish to use data of the Rotterdam Study must obtain approval from the Rotterdam Study Management Team. They are advised to contact the PI of the Rotterdam Study, Dr Albert Hofman (a.hofman@erasmusmc.nl).

Kaiser Permanente Research Program on Genes, Environment, and Health (RPGEH)

Data used in this study were provided by the Kaiser Permanente Research Program on Genes, Environment, and Health (RPGEH): Genetic Epidemiology Research on Adult Health and

Aging (GERA), funded by the National Institutes of Health [RC2 AG036607 (Schaefer and Risch)], the Robert Wood Johnson Foundation, the Wayne and Gladys Valley Foundation, The Ellison Medical Foundation, and the Kaiser Permanente Community Benefits Program. Access to RPGEH data used in this study may be obtained by application via the RPGEH Research portal: https://rpgehportal.kaiser.org. A subset of the GERA cohort consented for public use can be found at NIH/dbGaP: phs000674.v1.p1

SardiNIA

The SardiNIA (ProgeNIA) team was supported by Contract NO1-AG-1-2109 from the NIA, and in part by the Intramural Research Program of the National Institute on Aging (NIA), National Institutes of Health (NIH). The authors are grateful to all of the volunteers who participated in this study, Monsignore Piseddu, Bishop of Ogliastra, the mayors and citizens of the Sardinian towns (Lanusei, Ilbono, Arzana, and Elini), the head of the Public Health Unit ASL4 for their volunteerism and cooperation, and team of physicians, nurses, biologists and the recruitment personnel.

SHIP

SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network 'Greifswald Approach to Individualized Medicine (GANI_MED)' funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg- West Pomerania. The University of Greifswald is a member of the Caché Campus program of the InterSystems GmbH. External data access: Researchers may apply for access on the SHIP data by filling in a data application and sending it to the SHIP steering committee. The data application form can be accessed online at https://fvcm.med.uni-greifswald.de/

Sorbs

This project was supported by grants from the Collaborative Research Center funded by the German Research Foundation (CRC 1052; C01, B01, B03, SPP 1629 TO 718/2), from the German Diabetes Association, from the DHFD (Diabetes Hilfs- und Forschungsfonds Deutschland) and from Boehringer Ingelheim Foundation . We thank all those who participated in the study. Sincere thanks are given to Knut Krohn (Microarray Core Facility of the Interdisciplinary Centre for Clinical Research, University of Leipzig) for the genotyping support. Inga Prokopenko and Vasiliki Lagou were partial funded through the European Community's Seventh Framework Programme (FP7/2007-2013), ENGAGE project, grant agreement HEALTH-F4-2007-201413.

THISEAS

THISEAS (The Hellenic study of Interactions between SNPs & Eating in Atherosclerosis Susceptibility) - Recruitment for THISEAS was partially funded by a research grant (PENED 2003) from the Greek General Secretary of Research and Technology; we thank all the dieticians and clinicians for their contribution to the project. The genotyping was funded by the Wellcome Trust. We like to thank the members of the WTSI GenotypingFacility in particular Sarah Edkins and Cordelia Langford. Researchers interested in using the THISEAS data must obtain approval from the THISEAS study group. Researchers using the data are required to follow the terms of a research agreement between them and the THISEAS investigators. Note that individual level data cannot be released to external investigators, only summary GWAS results. For further information contact George Dedoussis (dedousi@hua.gr)

TwinGene (STR)

STR (Swedish Twin Registry) – The Jan Wallander and Tom Hedelius Foundation (P2012-0002:1), the Ragnar Söderberg Foundation (E9/11), The Swedish Research Council (421-2013-1061), the Ministry for Higher Education, The Swedish Research Council (M-2205-1112), GenomEUtwin (EU/QLRT-2001-01254; QLG2-CT-2002-01254), NIH DK U01-066134, The Swedish Foundation for Strategic Research (SSF). Researchers interested in using STR data must obtain approval from the Swedish Ethical Review Board and from the Steering Committee of the Swedish Twin Registry. Rietveld gratefully acknowledges funding

from the Netherlands Organization for Scientific Research (NWO Veni grant 016.165.004). Researchers using the data are required to follow the terms of an Assistance Agreement containing a number of clauses designed to ensure protection of privacy and compliance with relevant laws. For Further information, contact Patrik Magnusson (Patrik.magnusson@ki.se).

TwinsUK

The study was funded by the Wellcome Trust; European Community's Seventh Framework Programme (FP7/2007-2013). The study also receives support from the National Institute for Health Research (NIHR)- funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. SNP Genotyping was performed by The Wellcome Trust Sanger Institute and National Eye Institute via NIH/CIDR.

Statistical analyses were carried out on the Genetic Cluster Computer (http://www.geneticcluster.org), which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003) along with a supplement from the Dutch Brain Foundation.

Data availability: Data are available upon request from the TwinsUK data management bureau.

UKBiobank

This research has been conducted using the UK Biobank Resource

WGHS

The WGHS is supported by HL043851, HL080467 and CA047988 from the National Institutes of Health, with collaborative scientific support and funding for genotyping provided by Amgen.

WHICAP

WHICAP is supported by a grant (R01AG0372) from the National Institute on Aging of the National Institutes of Health.

WHITEHALL

The Whitehall II study has been supported by grants from the Medical Research Council (K013351); British Heart Foundation; Health and Safety Executive; Department of Health; National Heart Lung and Blood Institute (NHLBI: HL36310) and National Institute on Aging (AG13196), US, NIH; Agency for Health Care Policy Research (HS06516); and the John D and Catherine T MacArthur Foundation Research Networks on Successful Midlife Development and Socio-economic Status and Health. MeKu is partially supported by the Economic and Social Research Council International Centre for Life Course Studies in Society and Health (RES-596-28-0001). MK is partially supported by the Medical Research Council and the Economic and Social Research Council.

YFS

The Young Finns Study has been financially supported by the Academy of Finland: grants 286284 (T.L.), 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Kuopio, Tampere and Turku University Hospital Medical Funds (grant X51001 for T.L.); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation of Cardiovascular Research (T.L.); Finnish Cultural Foundation; Tampere Tuberculosis Foundation (T.L.); Emil Aaltonen Foundation (T.L.); and Yrjö Jahnsson Foundation (T.L.). We gratefully acknowledge the THL DNA laboratory for its skillful work to produce the DNA samples used in this study, and Ville Aalto and Irina Lisinen for the expert technical assistance in the statistical analyses. External researchers can get access to the YFS data in collaboration with the study group (contact information: Raitakari. Department of Clinical Physiology, University of Turku, PO Box 52, Turku FIN-20521, Finland. E-mail: olli.raitakari@utu.fi.)

14. LIST OF SUPPLEMENTARY TABLES

Supplementary Table 1 Study design, numbers of individuals and sample quality control for GWAS cohorts

Supplementary Table 2 Study-specific NEB and AFB measure and phenotype descriptive statistics

Supplementary Table 3 Results of gene-based analysis for AFB

Supplementary Table 4 Results of gene-based analysis for NEB

Supplementary Table 5 The results of in silico sequencing and in silico pleiotropy analysis. AF_EUR indicates the allele frequency of the alternative allele (A2) in the European population.

Supplementary Table 6 RegulomeDB functional variant analysis

Supplementary Table 7 cis eQTLs

Supplementary Table 8 cis mQTLs

Supplementary Table 9 trans eQTLs

Supplementary Table 10 trans mQTLs

Supplementary Table 11 Gene prioritisation for age at first birth for SNPs with p<5E-4 in the meta-analysis using Depict, MetaRanker, ToppGene and Endeavour

Supplementary Table 12 Gene prioritisation for number of children ever born for SNPs with p<5E-4 in the meta-analysis using Depict, MetaRanker, ToppGene and Endeavour

Supplementary Table 13 Depict's gene set enrichment analysis

Supplementary Table 14 Depict's tissue set enrichment analysis

Supplementary Table 15 Depict's gene set enrichment analysis for age at first birth in pooled data from men and women

Supplementary Table 16 Depict's tissue set enrichment analysis for age at first birth in pooled data from men and women

Supplementary Table 17 Depict's gene set enrichment analysis for age at first birth in data from women

Supplementary Table 18 Depict's tissue set enrichment analysis for age at first birth in data from women

Supplementary Table 19 Depict's gene set enrichment analysis for number of children ever born in pooled data from men and women

Supplementary Table 20 Depict's tissue set enrichment analysis for number of children ever born in pooled data from men and women

Supplementary Table 21 Logit regression childlessness, (Age 45 for women, age 55 for men) on NEB polygenic score (using all SNPs)

Supplementary Table 22 Cox regression model of AFB polygenic score (all SNPs) on Age at first birth

Supplementary Table 23 Linear prediction of age at menarche and age at menopause using AFB linear score

Supplementary Table 24 Lookup from related GWAS

Supplementary Table 25 LD Score estimates of genetic correlations

Supplementary Table 26 Number of SNPs (prior and post-QC) and inflation statistic λGC

Supplementary Table 27 Cohort-specific filters applied in deviation of generic filters to main excel SI file

Supplementary Table 28 SNP-based additive and dominant genetic effects on the age at first birth and the number of children ever born of women in the Netherlands and the UK

Supplementary Table 29 Bivariate and conditional analysis of the two fertility-related traits

Supplementary Table 30 Individual level and within-family (WF) regression models of AFB on polygenic scores, for scores constructed with sets of SNPs meeting different P-value thresholds.

Supplementary Table 31 Individual level and within-family (WF) regression models of NEB on polygenic scores, for scores constructed with sets of SNPs meeting different P-value thresholds.

Supplementary Table 32 Sex-specific effects and differences between sample and effect sizes

Supplementary Table 33 Summary statistics for women and men for all datasets separately. Sex-genotype interaction model using GREML

Supplementary Table 34 Bivariate analysis to estimate genetic correlations across sexes, GREML analysis

Supplementary Table 35 OLS regressions and Heckman two-stage regression models of AFB on polygenic score for AFB (using all SNPs). First stage selection models based on NEB polygenic scores (all SNPs)

Supplementary Table 36 Linear prediction of AFB and NEB using ANM PGS linear score

Supplementary Table 37 Overview of the 24 genes that were identified as likely being causal for multiple fertility-related traits and/or strata

Supplementary Table 38 Results from regression of AFB and NEB on the polygenic score (PGS) in a set of unrelated individuals of the STR cohorts (weights for the score are constructed from the GWAS meta-analysis excluding STR)

Supplementary Table 39 Linear OLS of AFB on polygenic score X education for men and women in STR and Lifelines, separate and pooled

Supplementary Table 40 Association results from deCODE before and after controlling for educational attainment (EA)

Supplementary Table 41 Association results from UKBiobank before and after controlling for educational attainment (EA) and age at first sex (AFS).

Supplementary Table 42 Haplotter results for evidence of positive selection in GWAS lead SNPs and RegulomeDB-identified functional variants

Supplementary Table 43 Author and Cohort Contributions

15. LIST OF SUPPLEMENTARY FIGURES

Supplementary Figure 1 Summary of fertility heritability estimates by birth cohort and country by fertility trait: (AFB) age at first birth, (NEB) number of children ever born.

Supplementary Figure 2 Variance explained by AFB and NEB polygenic scores calculated with the inclusion of SNPs at different levels of significance.

Supplementary Figure 3 Trans eQTL effect of rs2777888 is stronger in females as compared to males.

Supplementary Figure 4 Forest plot for *rs10908557* (chr1:153927052), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 5. Regional association plot of *rs10908557* (chr1:153927052), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 6 Forest plot for *rs1160544* (chr2:100832218), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 7 Regional association plot of *rs1160544* (chr2:100832218), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 8 Forest plot for *rs2777888* (chr3:4989000), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 9 Regional association plot of *rs2777888* (chr3:4989000), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 10 Forest plot for *rs6885307* (chr5:45094503), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 11 Regional association plot of *rs6885307* (chr5:45094503), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 12 Forest plot for *rs10056247* (chr5:133898136), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 13 Regional association plot of *rs10056247* (chr5:133898136), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 14 Forest plot for *rs2347867* (chr6:152229850), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 15 Regional association plot of *rs2347867* (chr6:152229850), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 16 Forest plot for *rs10953776* (chr7:114313218), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 17 Regional association plot of *rs10953776* (chr7:114313218), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 18 Forest plot for *rs2721195* (chr8:145677011), a genome-wide significant SNP for AFB women.

Supplementary Figure 19 Regional association plot of *rs2721195* (chr8:145677011), a genome-wide significant SNP for AFB women.

Supplementary Figure 20 Forest plot for *rs293566* (chr20:31097877), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 21 Regional association plot of *rs293566* (chr20:31097877), a genome-wide significant SNP for AFB women.

Supplementary Figure 22 Forest plot for *rs242997* (chr22:34503059), a genome-wide significant SNP for AFB pooled.

Supplementary Figure 23 Regional association plot of *rs242997* (chr22:34503059), a genome-wide significant SNP for AFB women.

Supplementary Figure 24 Forest plot for *rs10908474* (chr1:153753725), a genome-wide significant SNP for NEB pooled.

Supplementary Figure 25 Regional association plot of for *rs10908474* (chr1:153753725), a genome-wide significant SNP for AFB women.

Supplementary Figure 26 Forest plot for *rs13161115* (chr5:107050002), a genome-wide significant SNP for NEB men.

Supplementary Figure 27 Regional association plot of for *rs13161115* (chr5:107050002), a genome-wide significant SNP for NEB men.

Supplementary Figure 28 Forest plot for *rs2415984* (chr14:46873776), a genome-wide significant SNP for NEB pooled.

Supplementary Figure 29 Regional association plot of for *rs2415984* (chr14:46873776), a genome-wide significant SNP for NEB pooled.

Supplementary Figure 30 Bivariate analysis of the two fertility-related traits, comparing to each of the single trait analysis.

Supplementary Figure 31 Conditional analysis of the two fertility-related traits, comparing to each of the single trait analysis.

Supplementary Figure 32 Assessing the extent to which population stratification affects the estimates from the GWAS of Age at first birth.

Supplementary Figure 33 Assessing the extent to which population stratification affects the estimates from the GWAS of number of children ever born.

Supplementary Figure 34 Miami plots for AFB and NEB sex-specific single genomic control meta-analysis.

Supplementary Figure 35 Quantile-quantile plots of SNPs for AFB (panel a) and NEB (panel b) in single genomic control, meta-analysis.

Supplementary Figure 36 Look-up of female AFB SNPs with $p<1x10^{-04}$ for association with age at Menarche and Age at Menopause. Quantile-quantile plots.

Supplementary Figure 37 Comparison of effect size of 10 SNPs associated with AFB in the meta-analysis before and after controlling for educational attainment and age at first sexual intercourse.

References

1. Mills, M. C. & Tropf, F. C. The Biodemography of Fertility: A Review and Future Research Frontiers. *Kolner Z. Soz. Sozpsychol.* **55,** 397–424 (2016).
2. Byars, S. G., Ewbank, D., Govindaraju, D. R. & Stearns, S. C. Natural selection in a contemporary human population. *Proc. Natl. Acad. Sci.* **107,** 1787–1792 (2010).
3. Zietsch, B. P., Kuja-Halkola, R., Walum, H. & Verweij, K. J. Perfect genetic correlation between number of offspring and grandoffspring in an industrialized human population. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 1032–1036 (2014).
4. Kirk, K. M. *et al.* Natural selection and quantitative genetics of life-history traits in Western women: a twin study. *Evolution (N. Y).* **55,** 423–435 (2001).
5. Stearns, S. C., Byars, S. G., Govindaraju, D. R. & Ewbank, D. Measuring selection in contemporary human populations. *Nat. Rev. Genet.* **11,** 611–622 (2010).
6. Tanturri, M.L., Mencarini, L. Childless or Childfree? Paths to Voluntary Childlessness in Italy. *Popul. Dev. Rev.* **34,** 51–77 (2008).
7. Mills, M. C., Rindfuss, R. R., McDonald, P. & te Velde, E. Why do people postpone parenthood? Reasons and social policy incentives. *Hum. Reprod. Update* **17,** 848–860 (2011).
8. OECD. *Doing Better for Families*. (2011).
9. Leridon, H. A new estimate of permanent sterility by age: sterility defined as the inability to conceive. *Popul. Stud. (NY).* **62,** 15–24 (2008).
10. Balbo, N., Billari, F. C. & Mills, M. C. Fertility in advanced societies: A review of research. *Eur. J. Popul. Eur. Démographie* **29,** 1–38 (2013).
11. Mascarenhas, M. N., Flaxman, S. R., Boerma, T., Vanderpoel, S. & Stevens, G. A. National, regional, and global trends in infertility prevalence since 1990: a systematic analysis of 277 health surveys. *PLoS Med.* **9,** e1001356 (2012).
12. Boivin, J., Bunting, L., Collins, J. A. & Nygren, K. G. International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. *Hum. Reprod.* **22,** 1506–12 (2007).
13. Messerlian, C., Maclagan, L. & Basso, O. Infertility and the risk of adverse pregnancy outcomes: a systematic review and meta-analysis. *Hum. Reprod.* **28,** 125–37 (2013).
14. Jolly, M. The risks associated with pregnancy in women aged 35 years or older. *Hum.*

Reprod. **15,** 2433–2437 (2000).

15. Tarin, J. J., Brines, J. & Cano, A. Long-term effects of delayed parenthood. *Hum. Reprod.* **13,** 2371–2376 (1998).
16. Mehta, D. *et al.* Evidence for genetic overlap between schizophrenia and age at first birth in women. *JAMA Psychiatry* (2016).
17. Berrington, A. Perpetual postponers? Women's, men's and couple's fertility intentions and subsequent fertility behaviour. *Popul. Trends* **117,** 9–19 (2004).
18. Mills, M. C. *et al.* Why do people postpone parenthood? Reasons and social policy incentives. *Hum. Reprod. Update* **17,** 848–860 (2011).
19. Hakim, C. A new approach to explaining fertility patterns: preference theory. *Popul. Dev. Rev.* **29,** 349–374 (2003).
20. Avison, M., Furnham, A. Personality and voluntary childlessness. *J. Popul. Res.* **32,** 45–67 (2015).
21. Jeffries, S., Konnert, C. Regret and Psychological Well-Being among Voluntarily and Involuntarily Childless Women and Mothers. *Int. J. Aging Hum. Dev.* **54,** 89–106 (2002).
22. Koropeckyj-Cox, T., Call, V. R. A. Characteristics of older childless persons and parents. *J. Fam. Issues* **28,** 1362–1414 (2007).
23. Day, F. R. *et al.* Physical and neurobehavioral determinants of reproductive onset and success. *Nat. Genet.* doi:10.1038/ng.3551 (2016). doi:10.1038/ng.3551
24. Montgomery, G W, K.T. Zondervan, et. al. The future for genetic studies in reproduction. *Mol. Hum. Reprod.* **20,** 1–14 (2014).
25. Sulem, P. *et al.* Genome-wide association study identifies sequence variants on 6q21 associated with age at menarche. *Nat. Genet.* **41,** 734–738 (2009).
26. Elks, C. *et al.* Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. *Nat. Genet.* **42,** 1077–1085 (2010).
27. Day, Felix R., C.E. Elks, A. Murray, K.K. Ong, J. R. B. P. Puberty timing associated with diabetes, cardiovascular disease and also diverse health outcomes in men and women: the UK Biobank study. *Sci. Rep.* **5,** 11208 (2015).
28. Snieder, H., MacGregor, A. J. & Spector, T. D. Genes control the cessation of a woman's reproductive life: a twin study of hysterectomy and age at menopause. *J. Clin. Endocrinol. Metab.* **83,** 1875–1880 (1998).

29. Stolk, L. *et al.* Loci at chromosomes 13, 19 and 20 influence age at natural menopause. *Nat. Genet.* **41,** 645–647 (2009).
30. Stolk, L. *et al.* Meta-analyses identify 13 loci associated with age at menopause and highlight DNA repair and immune pathways. *Nat. Genet.* **44,** 260–268 (2012).
31. Perry, J., Corre, T. & Esko, T. A genome-wide association study of early menopause and the combined impact of identified variants. *Hum. Mol. Genet.* 1465–1472 (2013).
32. He, C. *et al.* Genome-wide association studies identify loci associated with age at menarche and age at natural menopause. *Nat. Genet.* **41,** 724–728 (2009).
33. Painter, J. N. *et al.* Genome-wide association study identifies a locus at 7p15. 2 associated with endometriosis. *Nat. Genet.* **43,** 51–54 (2011).
34. Rahmioglu, N. et al. Genetic variants underlying risk of endometriosis: insights from meta-analysis of eight genome-wide association and replication datasets. *Hum. Reprod. Update* **20,** 702–716 (2014).
35. Albertsen, H. M., Chettier, R., Farrington, P. & Ward, K. Genome-Wide Association Study Link Novel Loci to Endometriosis. *PLoS One* **8,** e58257 (2013).
36. Dhawan, V., Brookes, Z. L. S. & Kaufman, S. Long-term effects of repeated pregnancies (multiparity) on blood pressure regulation. *Cardiovasc. Res.* **64,** 179–86 (2004).
37. Day, F. R. *et al.* Causal mechanisms and balancing selection inferred from genetic associations with polycystic ovary syndrome. *Nat. Commun.* **6,** 8464 (2015).
38. Tropf, F. C. *et al.* Human Fertility, Molecular Genetics, and Natural Selection in Modern Societies. *PLoS One* **10,** e0126821 (2015).
39. Courtiol, A, Tropf, F.C., Mills, M. C. When genes and environment disagree: Making sense of trends in recent human evolution. *Proc Natl Acad Sci US A* **113,** 7693–7695 (2016).
40. Fisher, R. A. *The genetical theory of natural selection.* (Oxford University Press, 1930).
41. Kohler, H.-P., Rodgers, J. L. & Christensen, K. Is fertility behavior in our genes? Findings from a Danish twin study. *Popul. Dev. Rev.* **25,** 253–288 (1999).
42. Visscher, P. M., Hill, W. G. & Wray, N. R. Heritability in the genomics era—concepts and misconceptions. *Nat. Rev. Genet.* **9,** 255–266 (2008).
43. Mousseau, T. A. & Roff, D. A. Natural selection and the heritability of fitness

components. *Heredity (Edinb).* **59,** 181–197 (1987).

44. Polderman, T. *et al.* Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* doi: 10.1038/ng.3285 (2015).
45. Hughes, K. & Burleson, M. H. in (eds. Rodgers, J., Row, D. C. & Miller, W. B.) 7–33 (Kluwer Academic Publisher, 2000).
46. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88,** 76–82 (2011).
47. Visscher, P. M., Yang, J. & Goddard, M. E. A commentary on 'common SNPs explain a large proportion of the heritability for human height'by Yang et al.(2010). *Twin Res. Hum. Genet.* **13,** 517–524 (2010).
48. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42,** 565–569 (2010).
49. Zhu, Z. *et al.* Dominance genetic variation contributes little to the missing heritability for human complex traits. *Am. J. Hum. Genet.* **96,** 377–385. (2015).
50. Zaitlen, N. *et al.* Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. *PLoS Genet.* **9,** e1003520 (2013).
51. Goldin, C., Katz, L. F. The power of the pill: Oral contraceptives and women's career and marriage decisions. *J. Polit. Econ.* **110,** 730–770 (2002).
52. Balbo, N., Billari, F. C. & Mills, M. C. Fertility in Advanced Societies: A Review of Research. *Eur. J. Popul. / Rev. Eur. Démographie* **29,** 1–38 (2012).
53. Rindfuss, R. R., John, C. St. & Bumpass, L. L. Education and the timing of motherhood: Disentangling causation. *J. Marriage Fam.* 981–984 (1984).
54. Rindfuss, R. R., Morgan, S. P. & Offutt, K. Education and the changing age pattern of American fertility: 1963–1989. *Demography* **33,** 277–290 (1996).
55. Rindfuss, R. R. & St. John, C. Social Determinants of Age at First Birth. *J. Marriage Fam.* **45,** 553–565 (1983).
56. Oppenheimer, V. K. Women's Rising Employment and the Future of the Family in Industrial Societies. *Popul. Dev. Rev.* **20,** 293–342 (1994).
57. Becker, G. S. & Becker, G. S. *A Treatise on the Family*. (Harvard university press, 2009).
58. Begall, K. & Mills, M. The Influence of Educational Field, Occupation, and Occupational Sex Segregation on Fertility in the Netherlands. *Eur. Sociol. Rev.* **29,**

720–742 (2013).

59. Brewster, K. L. & Rindfuss, R. R. Fertility and Women's Employment in Industrialized Nations. *Annual Review of Sociology* **26,** 271–296 (2000).
60. Amuedo-Dorantes, C. & Kimmel, J. The motherhood wage gap for women in the United States: The importance of college and fertility delay. *Rev. Econ. Househ.* **3,** 17–48 (2005).
61. Miller, A. R. The effects of motherhood timing on career path. *J. Popul. Econ.* **24,** 1071–1100 (2011).
62. de Kaa, D. J. Van. Europe's second demographic transition. *Popul. Bull.* **42,** 1–59 (1987).
63. Lesthaeghe, R. The second demographic transition in Western countries: An interpretation. *Gend. Fam. Chang. Ind. Ctries.* 17–62 (1995).
64. Axinn, W. G., Clarkberg, M. E. & Thornton, A. Family influences on family size preferences. *Demography* **31,** 65–79 (1994).
65. Mills, M., Begall, K. Preferences for the sex-composition of children in Europe: A multilevel examination of its effect on progression to a third child. *Popul. Stud. (NY).* **64,** 77–95. (2010).
66. McDonald, P. Gender equity in theories of fertility transition. *Popul. Dev. Rev.* **26,** 427–439. (2002).
67. Mills, M. C. & Blossfeld, H.-P. Globalisation, Uncertainty and the Early Life Course: A Theoretical Framework: 1–24. *Glob. Uncertain. Youth Soc. ...* (2005).
68. Mulder, C. H. Home-Ownership and Family Formation. *J. Hous. Built Environ.* **21,** 281–298 (2006).
69. Balbo, N. & Barban, N. Does fertility behavior spread among friends? *Am. Sociol. Rev.* **79,** 412–431 (2014).
70. Balbo, N. & Mills, M. The effects of social capital and social pressure on the intention to have a second or third child in France, Germany, and Bulgaria, 2004–05. *Popul. Stud. (NY).* **65,** 335–351 (2011).
71. Poortman, A.R., Mills, M. Joint investments in marriage and cohabitation: The role of legal and symbolic factors. *J. Marriage Fam.* **74,** 357–376 (2012).
72. Wood, A. R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46,** 1173–1186 (2014).

73. van der Most, P. J. *et al.* QCGWAS: A flexible R package for automated quality control of genome-wide association results. *Bioinformatics* **30,** 1185–1186 (2014).
74. Winkler, T. W. *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nat. Protoc.* **9,** 1192–1212 (2014).
75. Locke, A. E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518,** 197–206 (2015).
76. A. Okbay, J.P. Beauchamp, M.A. Fontana, J.J. Lee, T.H. Pers, C.A. Rietveld, P. Turley,..., P.M. Visscher, T. Esko, P.D. Koellinger, D. Cesarini, D. J. B. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* (2016). doi:10.1038/nature17671
77. International, T. & Consortium, H. The International HapMap Project. *Nature* **426,** 789–796 (2003).
78. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38,** 904–909 (2006).
79. Rietveld, C. A. *et al.* GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment. *Science (80-.).* **340,** 1467–1471 (2013).
80. Okbay, A. *et al.* Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat. Genet.* **advance on,** (2016).
81. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26,** 2190–2191 (2010).
82. Evangelou, E. & Ioannidis, J. P. A. Meta-analysis methods for genome-wide association studies and beyond. *Nat. Rev. Genet.* **14,** 379–389 (2013).
83. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–575 (2007).
84. Pruim, RJ, Welch, RP, Sanna, S, Taslovich, TM, Chines, PS, Bliedt, TP, Boehnke, M., Abecasis, GR, Willer, C. LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics* **26,** 2336–7 (2010).
85. Shen, X. *et al. Simple multi-trait analysis identifies novel loci associated with growth and obesity measures*. *bioRxiv* (Cold Spring Harbor Labs Journals, 2015). at <http://biorxiv.org/content/early/2015/07/08/022269.abstract>
86. Bulik-Sullivan, B. K. et al. LD Score regression distinguishes confounding from

polygenicity in genome-wide association studies. *Nat. Genet.* **47,** 291–295 (2015).

87. Wood, A. R. et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* (2014).
88. Purcell, S. M. et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 748–752 (2009).
89. Rietveld, C. A. *et al.* Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. *Proc. Natl. Acad. Sci.* **111,** 13790–13794 (2014).
90. Bulik-Sullivan, B. K. & Al., E. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47,** 1236–41 (2015).
91. Finucane, H. K. et al. *Partionining heritability by functional category using GWAS summary statistics*. *bioRxiv*. (2015). at <https://github.com/bulik/ldsc/wiki/Genetic-Correlation>
92. Yang, J. et al. Genome-wide genetic homogeneity between sexes and populations for human height and body mass index. *Hum. Mol. Genet.* **24,** 7445–7449 (2015).
93. Ober, C., Loisel, D.A. & Gilard, Y. Sex-specific genetic architecture of human disease. *Nat. Rev. Genet.* **9,** 911–922 (2008).
94. Short, S.E., Yang, Y.C., Jenkins, T. M. Sex, Gender, Genetics, and Health. *Am. J. Public Health* **103,** S93–S101 (2013).
95. Perry, B. L. Gendering Genetics: Biological Contingencies in the Protective Effects of Social Integration for Men and Women. *Am. J. Sociol.* **121,** 1655–96 (2016).
96. Drefahl, S. How does the age gap between partners affect their survival? *Demography* **47,** 313–326 (2010).
97. Schousboe, K. et al. Sex Differences in Heritability of BMI: A Comparative Study of Results from Twin Studies in Eight Countries. *Twin Res.* **6,** 409–421 (2003).
98. Randall, J. C. et al. Sex-stratified Genome-wide Association Studies Including 270,000 Individuals Show Sexual Dimorphism in Genetic Loci for Anthropometric Traits. *PLoS Genet.* **9,** e1003500 (2013).
99. Hughes, K.A., Burleson, M. H. *Genetic Influences on Human Fertility and Sexuality*. (2000).
100. Rodgers, J. L. *et al.* Genetic influence helps explain variation in human fertility: Evidence from recent behavioral and molecular genetic studies. *Curr. Dir. Psychol.*

Sci. **10,** 184 (2001).

101. Gershoni, M., Pietrokovski, S. Reduced selection and accumulation of deleterious mutations in genes exclusively expressed in men. *Nat. Commun.* **5,** 1–10 (2014).
102. Visscher, P. M. *et al.* Statistical power to detect genetic (co) variance of complex traits using SNP data in unrelated samples. *PLoS Genet.* **10,** e1004269 (2014).
103. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42,** 565–569 (2010).
104. Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M. & Wray, N. R. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* **28,** 2540–2542 (2012).
105. Ge, Y.-Z. et al. Association of polymorphisms in estrogen receptors (ESR1 and ESR2) with male infertility: a meta-analysis and systematic review. *J. Assist. Reprod. Genet.* **31,** 601–611 (2014).
106. Mills, M. C. *Introducing survival and event history analysis*. (Sage Publications, 2011).
107. Mills, M. C., Blossfeld, H.-P. & Klijzing, E. Becoming an adult in uncertain times. *Glob. Uncertain. ...* (2005).
108. Day, F., Ruth, K. & Thompson, D. Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair. *Nat. Genet.* **47,** 1294–1303 (2015).
109. Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461,** 747–753 (2009).
110. Zuk, O. & Hechter, E. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc. Natl. Acad. Sci.* **109,** 1193–1198. (2012).
111. Yang, J. *et al.* Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat. Genet.* **47,** 1114–1120 (2015).
112. Tropf, F. C. *et al. Mega-analysis of 31,396 individuals from 6 countries uncovers strong gene-environment interaction for human fertility*. *bioRxiv* (Cold Spring Harbor Labs Journals, 2016). doi:10.1101/049163
113. Domingue, BW, Conley, D., Fletcher, J., Boardman, J. D. Cohort Effects in the

Genetic Influence on Smoking. *Behav. Genet.* **46,** 31–42 (2016).

114. Kravdal, O. & Rindsfuss, R. R. *Changing relationship between education and fertility: A study of women and men born 1940-64*. (2007).
115. Tropf, F. C., Barban, N., Mills, M. C., Snieder, H. & Mandemakers, J. J. Genetic influence on age at first birth of female twins born in the UK, 1919-68. *Popul. Stud. (NY).* 129–145. (2015).
116. Perry, J. R. B. *et al.* Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature* **514,** 92–97 (2014).
117. Day, F. R. *et al.* Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair. *Nat. Genet.* **47,** 1294–1303 (2015).
118. Cousminer, D. L. *et al.* 24770850. *Hum. Mol. Genet.* **23,** 4452–4464 (2014).
119. Day, F. R. et al. Shared genetic aetiology of puberty timing between sexes and with health-related outcomes. *Nat. Commun.* **6,** 8842 (2015).
120. Day, F. R. *et al.* Causal mechanisms and balancing selection inferred from genetic associations with polycystic ovary syndrome. *Nat. Commun.* **6,** 8464 (2015).
121. Mbarek, H. *et al.* Identification of Common Genetic Variants Influencing Spontaneous Dizygotic Twinning and Female Fertility. *Am. J. Hum. Genet.* **98,** 898–908 (2016).
122. van der Valk, R. J. P. *et al.* A novel common variant in DCST2 is associated with length in early life and height in adulthood. *Hum. Mol. Genet.* **24,** 1155–68 (2015).
123. Horikoshi, M. *et al.* New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism. *Nat. Genet.* **45,** 76–82 (2013).
124. Siedlinski, M. *et al.* Genome-wide association study of smoking behaviours in patients with COPD. *Thorax* **66,** 894–902 (2011).
125. de Moor, M. H. M. *et al.* Meta-analysis of Genome-wide Association Studies for Neuroticism, and the Polygenic Association With Major Depressive Disorder. *JAMA psychiatry* **72,** 642–50 (2015).
126. Ripke, S. *et al.* Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511,** 421–427 (2014).
127. Sklar, P. *et al.* Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat. Genet.* **43,** 977–983 (2011).

128. Lambert, J. C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45,** 1452–8 (2013).
129. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* **381,** 1371–9 (2013).
130. Willer, C. J. *et al.* Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* **45,** 1274–83 (2013).
131. Morris, A. P. *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* **44,** 981–90 (2012).
132. Manning, A. K. *et al.* A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat. Genet.* **44,** 659–669 (2012).
133. Locke, A. E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518,** 197–206 (2015).
134. Shungin, D. *et al.* New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518,** 187–196 (2015).
135. Guldbrandsen, K. et al. Age of menarche and time to pregnancy. *Hum. Reprod.* **29,** 2058–2064 (2014).
136. Weghofer, A. et al. Age at menarche: a predictor of diminished ovarian function? *Fertil. Steril.* **100,** 1039–1043 (2013).
137. Nnoaham, K. E. et al. Is early age at menarche a risk factor for endometriosis? A systematic review and meta-analysis of case-control studies. *Fertil. Steril.* **98,** 702–712 (2012).
138. Stolk, L. *et al.* Meta-analyses identify 13 loci associated with age at menopause and highlight DNA repair and immune pathways. *Nat. Genet.* **44,** 260–8 (2012).
139. Day, F. R. & Al., E. Shared genetic aetiology of puberty timing between sexes and with health-related outcomes. *Nat. Commun.* **6,** 8842 (2015).
140. Colditz, G. a *et al.* A prospective study of age at menarche, parity, age at first birth, and coronary heart disease in women. *Am. J. Epidemiol.* **126,** 861–70 (1987).
141. Bareh, G.M., Robinson, R. D. Endometriosis and Lipid Concentration. Do we have to screen patients for hyperlipidemia? *Obstet. Gynecol.* **125,** 110S
142. Rahmioglu, N. et al. Genome-wide enrichment analysis between endometriosis and

obesity-related traits reveals novel susceptibility loci. *Hum. Mol. Genet.* **24,** 1185–1199 (2015).

143. van Montfoort, A. P. A. et al. Impact of maternal cholesterol metabolism on ovarian follicle development and fertility. *J. Reprod. Immunol.* **104–105,** 32–36 (2014).
144. DeAngelis, A.M, et al. Genetic Alterations Affecting Cholesterol Metabolism and Human Fertility. *Biol. Reprod.* **91,** 1–10 (2014).
145. Morrison, C.D., Brannigan, R. E. Metabolic syndrome and infertility in men. *Best Pract. Res. Clin. Obstet. Gynaecol.* **29,** 507–515 (2015).
146. Alderman, B. W. *et al.* Reproductive history, glucose tolerance, and NIDDM in Hispanic and non-Hispanic white women. The San Luis Valley Diabetes Study. *Diabetes Care* **16,** 1557–1564 (1993).
147. COWAN, L. D. *et al.* Parity, Postmenopausal Estrogen Use, and Cardiovascular Disease Risk Factors in American Indian Women: The Strong Heart Study. *J. Women's Heal.* **6,** 441–449 (1997).
148. Kritz-Silverstein, D., Barrett-Connor, E., Wingard, D. L. & Friedlander, N. J. Relation of pregnancy history to insulin levels in older, nondiabetic women. *Am. J. Epidemiol.* **140,** 375–382 (1994).
149. Lawlor, D. A. Is the Association Between Parity and Coronary Heart Disease Due to Biological Effects of Pregnancy or Adverse Lifestyle Risk Factors Associated With Child-Rearing?: Findings From the British Women's Heart and Health Study and the British Regional Heart S. *Circulation* **107,** 1260–1264 (2003).
150. Eisenberg, M. L. *et al.* Fatherhood and the risk of cardiovascular mortality in the NIH-AARP Diet and Health Study. *Hum. Reprod.* **26,** 3479–3485 (2011).
151. Nicholson, W. K. *et al.* Parity and risk of type 2 diabetes: the Atherosclerosis Risk in Communities Study. *Diabetes Care* **29,** 2349–2354 (2006).
152. Bellamy, L., Casas, J.-P., Hingorani, A. D. & Williams, D. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. *Lancet* **373,** 1773–1779 (2009).
153. Gunderson, E. P. *et al.* A 20-year prospective study of childbearing and incidence of diabetes in young women, controlling for glycemia before conception: the Coronary Artery Risk Development in Young Adults (CARDIA) Study. *Diabetes* **56,** 2990–2996 (2007).

154. Dabelea, D. & Snell-Bergeon, J. Increasing prevalence of gestational diabetes mellitus (GDM) over time and by birth Cohort Kaiser Permanente of Colorado GDM Screening Program. *Diabetes ...* **28,** 579–584 (2005).
155. Stulp, G., Barrett, L., Tropf, F. C. & Mills, M. C. Does natural selection favour taller stature among the tallest people on earth? *Proc. R. Soc. London B Biol. Sci.* **282,** 20150211 (2015).
156. Stearns, S. C. *The evolution of life histories*. **249,** (Oxford University Press Oxford, 1992).
157. Stulp, G., Verhulst, S., Pollet, T. V & Buunk, A. P. The effect of female height on reproductive success is negative in western populations, but more variable in non-western populations. *Am. J. Hum. Biol.* **24,** 486–494 (2012).
158. Stulp, G., Pollet, T. V., Verhulst, S. & Buunk, A. P. A curvilinear effect of height on reproductive success in human males. *Behav. Ecol. Sociobiol.* **66,** 375–384 (2012).
159. Oreffice, S. & Quintana-Domeque, C. Anthropometry and socioeconomics among couples: Evidence in the United States. *Econ. Hum. Biol.* **8,** 373–384 (2010).
160. Krzyżanowska, M., Mascie-Taylor, C. G. N. & Thalabard, J.-C. Is human mating for height associated with fertility? Results from a british national cohort study. *Am. J. Hum. Biol.* **0,** n/a-n/a (2015).
161. Velva, Z. et al. High and low BMI increase the risk of miscarriage after IVF/ICSI and FET. *Hum. Reprod.* **23,** 878–884. (2008).
162. Jokela, M. et al. Lower fertility associated with obesity and underweight: the US National Longitudinal Survey of Youth. *Am. J. Clin. Nutr.* **88,** 886–893 (2008).
163. Heid, I. et al. Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat. Genet.* **42,** 949–60. (2010).
164. Philip Morgan, S. Is low fertility a twenty-first-century demographic crisis? *Demography* **40,** 589–603 (2003).
165. Skirbekk, V. & Samir, K. C. Fertility-Reducing Dynamics of Women's Social Status and Educational Attainment. *Asian Popul. Stud.* **8,** 251–264 (2012).
166. Neiss, M., Rowe, D. C. & Rodgers, J. L. Does education mediate the relationship between IQ and age of first birth? A behavioural genetic analysis. *J. Biosoc. Sci.* **34,** 259–276 (2002).

167. Halpern, C. T. et al. Smart teens don't have sex (or kiss much either). *J. Adolesc. Heal.* **26,** 213–225 (2000).
168. Hutteman, R., Bleidorn, W., Penke, L. & Denissen, J. J. A. It takes two: A longitudinal dyadic study on predictors of fertility outcomes. *J. Pers.* **81,** 487–498 (2013).
169. Berg, V., Lummaa, V., Lahdenperä, M., Rotkirch, A. & Jokela, M. Personality and long-term reproductive success measured by the number of grandchildren. *Evol. Hum. Behav.* **35,** 533–539 (2014).
170. Alvergne, A., Jokela, M. & Lummaa, V. Personality and reproductive success in a high-fertility human population. *Proc Natl Acad Sci US A* **107,** 11745–11750 (2010).
171. Reis, O., Dörnte, M., & von der Lippe, H. Neuroticism, social support, and the timing of first parenthood: A prospective study. *Pers. Individ. Dif.* **50,** 381–386 (2011).
172. Courtiol, A., Pettay, J. E., Jokela, M., Rotkirch, A. & Lummaa, V. Natural and sexual selection in a monogamous historical human population. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 8044–8049 (2012).
173. Jokela, M., Kivimaki, M., Elovainio, M. & Keltikangas-Jarvinen, L. Personality and having children: A two-way relationship. *J. Pers. Soc. Psychol.* **96,** 208–230 (2009).
174. Briley, DA, Tucker-Drob, E. Genetic and environmental continuity in personality development: A meta-analysis. *Psychol. Bull.* **140,** 1303–1331 (2017).
175. Howe, G., C. Westhoff, M. Vessey, Yeates, D. Effects Of Age, Cigarette Smoking, And Other Factors On Fertility: Findings In A Large Prospective Study. *Br. Med. J.* **290,** 1697–1700 (1985).
176. Cooper, A.R., Moley, K. H. Maternal tobacco use and its preimplantation effects on fertility: more reasons to stop smoking. *Semin. Reprod. Med.* **26,** 204–212 (2008).
177. Mai, Z., Lei, M., Y, B., Du, H. & Lui, J. The effects of cigarette smoke extract on ovulation, oocyte morphology and ovarian gene expression in mice. *PLoS One* **9,** e95945 (2014).
178. Omurtag, K, et al. Modeling the effect of cigarette smoke on hexose utilization in spermatocytes. *Reprod. Sci.* **22,** 94–101 (2015).
179. Bergström, U. et al. Pulmonary dysfunction, smoking, socioeconomic status and the risk of developing rheumatoid arthritis. *Rheumatology* **50,** 2005–2013 (2011).
180. Laaksonen, M. et al. Socioeconomic status and smokingAnalysing inequalities with multiple indicators. *Eur. J. Public Health* **15,** 262–269 (2005).

181. Rietveld, C. A. *et al.* Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. *Proc Natl Acad Sci US A* **111,** 13790–13794 (2014).
182. Winkler, T. W. *et al.* EasyStrata: evaluation and visualization of stratified genome-wide association meta-analysis data. *Bioinformatics* **31,** 259–261 (2014).
183. Vaez, A. *et al.* In Silico Post Genome-Wide Association Studies Analysis of C-Reactive Protein Loci Suggests an Important Role for Interferons. *Circ. Cardiovasc. Genet.* **8,** 487–497 (2015).
184. Abecasis, G. R. *et al.* A map of human genome variation from population-scale sequencing. *Nature* **467,** 1061–1073 (2010).
185. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27,** 2156–2158 (2011).
186. Li, H. Tabix: fast retrieval of sequence features from generic TAB-delimited files. *Bioinformatics* **27,** 718–719 (2011).
187. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38,** e164–e164 (2010).
188. Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4,** 1073–1081 (2009).
189. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7,** 248–249 (2010).
190. Flicek, P. *et al.* Ensembl 2012. *Nucleic Acids Res.* **40,** D84–D90 (2011).
191. Hindorff, L. A. *et al.* Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci.* **106,** 9362–9367 (2009).
192. Liu, J. Z. *et al.* A versatile gene-based test for genome-wide association studies. *Am. J. Hum. Genet.* **87,** 139–45 (2010).
193. Mishra, A. & Macgregor, S. VEGAS2: Software for More Flexible Gene-Based Testing. *Twin Res. Hum. Genet.* 1–6 (2014). doi:10.1017/thg.2014.79
194. Zhernakova, D. *et al. Hypothesis-free identification of modulators of genetic risk factors*. *bioRxiv* (Cold Spring Harbor Labs Journals, 2015). at

<http://biorxiv.org/content/early/2015/11/30/033217.abstract>

195. Bonder, M. J. *et al. Disease variants alter transcription factor levels and methylation of their binding sites*. *bioRxiv* (Cold Spring Harbor Labs Journals, 2015). doi:10.1101/033084

196. Deelen, J. *et al.* Genome-wide association meta-analysis of human longevity identifies a novel locus conferring survival beyond 90 years of age. *Hum. Mol. Genet.* **23,** 4420–32 (2014).

197. Hofman, A. *et al.* The rotterdam study: 2010 objectives and design update. *Eur. J. Epidemiol.* **24,** 553–572 (2009).

198. Tigchelaar, E. F. *et al.* Cohort profile: LifeLines DEEP, a prospective, general population cohort study in the northern Netherlands: study design and baseline characteristics. *BMJ Open* **5,** e006772 (2015).

199. Van Dam, R. M., Boer, J. M. A., Feskens, E. J. M. & Seidell, J. C. Parental history off diabetes modifies the association between abdominal adiposity and hyperglycemia. *Diabetes Care* **24,** 1454–1459 (2001).

200. Willemsen, G. *et al.* The Adult Netherlands Twin Register: twenty-five years of survey and biological data collection. *Twin Res. Hum. Genet.* **16,** 271–81 (2013).

201. Boyle, A. P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22,** 1790–7 (2012).

202. Consortium, R. E. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518,** 317–330 (2015).

203. Encode Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489,** 57–74 (2013).

204. Chen, J., Xu, H., Aronow, B. J. & Jegga, A. G. Improved human disease candidate gene prioritization using mouse phenotype. *BMC Bioinformatics* **8,** 392 (2007).

205. Tranchevent, L.-C. *et al.* ENDEAVOUR update: a web resource for gene prioritization in multiple species. *Nucleic Acids Res.* **36,** W377-84 (2008).

206. Pers, T. H., Dworzyński, P., Thomas, C. E., Lage, K. & Brunak, S. MetaRanker 2.0: a web server for prioritization of genetic variation data. *Nucleic Acids Res.* **41,** 104–108 (2013).

207. Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6,** 5890 (2015).

208. Mostafavi, S., Ray, D., Warde-Farley, D., Grouios, C. & Morris, Q. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. *Genome Biol.* **9,** S4 (2008).
209. Saito, R. *et al.* A travel guide to Cytoscape plugins. *Nat. Methods* **9,** 1069–1076 (2012).
210. Montojo, J. *et al.* GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. *Bioinformatics* **26,** 2927–2928 (2010).
211. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25,** 25–29 (2000).
212. Kohler, H.-P., Rodgers, J. L. & Christensen, K. Is Fertility Behavior in Our Genes? Findings from a Danish Twin Study. *Popul. Dev. Rev.* **25,** 253–288 (1999).
213. Tropf, F. C., Barban, N., Mills, M. C., Snieder, H. & Mandemakers, J. J. Genetic influence on age at first birth of female twins born in the UK, 1919--68. *Popul. Stud. (NY).* **69,** 129–145 (2015).
214. Beauchamp, J. P. Genetic evidence for natural selection in humans in the contemporary United States. *Proc Natl Acad Sci US A* **113,** 7774–7779 (2016).
215. Boardman, J. *et al.* Population composition, public policy, and the genetics of smoking. *Demography* **48,** 1517–1533. (2011).
216. Boardman, J.D., Blalock, C.L., Pampel, F. C. Trends in the Genetic Influences on Smoking. *J. Health Soc. Behav.* **51,** 108–123 (2010).
217. Domingue, BW, Belksy, D.W., Harrati, A., Conley, D., Weir, D., Boardman, J. Mortality selection in a genetic sample and implications for association studies. *bioRxiv* (20016). doi:http://dx.doi.org/10.1101/049635
218. Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* 1467–1471 (2016). doi:10.1038/nature17671
219. Voight, B. F., Kudaravalli, S., Wen, X. & Pritchard, J. K. A Map of Recent Positive Selection in the Human Genome. *PLoS Biol.* **4,** e72 (2006).
220. Stolk, R. P. *et al.* Universal risk factors for multifactorial diseases. *Eur. J. Epidemiol.* **23,** 67–74 (2008).
221. Scholtens, S. *et al.* Cohort Profile: LifeLines, a three-generation cohort study and biobank. *Int. J. Epidemiol.* **44,** 1172–1180 (2015).