Genetic variants determining survival and fertility in an adverse African environment: a population-based large-scale candidate gene association study

Jacob J.E. Koopman[1], Jeroen Pijpe[2,3], Stefan Böhringer[4], David van Bodegom[1,5], Ulrika K. Eriksson[1], Hernando Sanchez-Faddeev[2], Juventus B. Ziem[6], Bas Zwaan[7], P. Eline Slagboom[8], Peter de Knijff[2], and Rudi G.J. Westendorp[1,9]

[1]Section of Gerontology and Geriatrics, Department of Internal Medicine, Leiden University Medical Center, Leiden, the Netherlands
[2]Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands
[3]University of Applied Sciences Leiden, Leiden, the Netherlands
[4]Section of Medical Statistics, Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, the Netherlands
[5]Leyden Academy on Vitality and Ageing, Leiden, the Netherlands
[6]Department of Clinical Laboratory Sciences, School of Medicine and Health Sciences, University for Development Studies, Tamale, Ghana
[7]Laboratory of Genetics, Wageningen University, Wageningen, the Netherlands
[8]Section of Molecular Epidemiology, Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, the Netherlands
[9]Department of Public Health and Center of Healthy Aging, University of Copenhagen, Copenhagen, Denmark

Abstract: Human survival probability and fertility decline strongly with age. These life history traits have been shaped by evolution. However, research has failed to uncover a consistent genetic determination of variation in survival and fertility. As an explanation, such genetic determinants have been selected in adverse environments, in which humans have lived during most of their history, but are almost exclusively studied in populations in modern affluent environments. Here, we present a large-scale candidate gene association study in a rural African population living in an adverse environment. In 4387 individuals, we studied 4052 SNPs in 148 genes that have previously been identified as possible determinants of survival or fertility in animals or humans. We studied their associations with survival comparing newborns, middle-age adults, and old individuals. In women, we assessed their associations with reported and observed numbers of children. We found no statistically significant associations of these SNPs with survival between the three age groups nor with women's reported and observed fertility. Population stratification was unlikely to explain these results. Apart from a lack of power, we hypothesise that genetic heterogeneity of complex phenotypes and gene-environment interactions prevent the identification of genetic variants explaining variation in survival and fertility in humans.

INTRODUCTION

Age patterns of survival and fertility vary widely across species [1]. During evolution, natural selection has shaped these age patterns, referred to as life histories, so to optimise the fitness of each species by maximising reproduction [2]. Reproduction is increased if survival or fertility are enhanced or if a decline in survival or

fertility with age is resisted. The diversity in life history across species indicates that it has a strong genetic basis. Several genetic pathways have been found in animals that regulate survival and fertility, including the signalling cascade of growth hormone (GH), insulin-like growth factor 1 (IGF1), and insulin, signalling by target of rapamycin (TOR), DNA repair mechanisms, immune regulation, and telomere maintenance [3, 4]. These pathways have been discovered mostly in studies on mutant animal models, but likely contribute to variation in survival and fertility in wild-type animals as well [5, 6]. Genetic variation in these pathways is thought to likewise determine the age patterns of survival and fertility in humans [7], but it remains disappointingly inconclusive as an explanation of observed variation in human survival and fertility. Candidate-gene studies, linkage studies, and genome-wide association studies have yielded consistent evidence for only a handful of genetic variants to determine variation in human survival to old age, of which most notably *APOE* [8-11]. Genetic variants have been described as determinants of human infertility [12, 13], but have rarely been studied for variation in human fertility [14].

Research on the genetic determinants of human life history has almost exclusively been conducted in populations living in modern affluent environments. These modern affluent environments, however, are radically different from the environments in which over many generations humans have been subjected to evolutionary pressures. During most of human history, survival and fertility were compromised by infectious diseases [15, 16], malnutrition [17, 18], climatic hardships, predation, and violence [19]. Fitness, which includes survival and fertility, has long been shaped by natural selection enforced through these environmental stressors. It is, therefore, likely that natural selection has enhanced survival and fertility by promoting genetic variants that shape inflammatory processes to improve resistance against infections, metabolic processes to facilitate consumption and storage of nutrients, and psychological strategies to cope with environmental stressors. In modern affluent environments, however, where survival and fertility depend less on these adverse environmental stressors, such genetic variants are of less influence on life history.

If we aim to identify the genetic determinants of life history that have enhanced fitness in adverse environments, we should search for them in such environments. In this study, we investigate genetic variants that determine life history through variation in survival and fertility in a traditional rural African population that lives in such an adverse environment.

Compared with modern affluent environments, this population's mortality rates and fertility rates are high, various infectious diseases are endemic, periodic food shortage and malnutrition are common, and a sedentary lifestyle is absent [20-29].

RESULTS

Table 1 shows the general characteristics of the Ghanaian study population. Genetic variants as determinants of variation in survival were investigated in men and women together, grouped as newborns, middle-aged adults of fertile ages from 20 through 44 years, and old individuals aged 60 years or over. Genetic variants as determinants of variation in fertility were investigated in women only. Observed fertility was registered in middle-aged women of fertile ages from 20 through 44 years during follow-up. Reported fertility was registered in postmenopausal women aged 45 years or over at the beginning of follow-up.

As shown in Figure 1, after the quality control 4052 SNPs, encompassing 148 genes, were included in the analyses. The median (interquartile range) number of kb between SNPs included in the analyses was 2.3 kb (1.8-3.4 kb). An overview of the included genes is given in Supplementary Table 1.

Figure 2 reports on the investigation of the genetic variants as determinants of variation in survival. We assessed the association of each SNP with the chance of being an old individual as compared with a newborn, a middle-aged adult as compared with a newborn, and an old individual as compared with a middle-aged adult, reflecting the survival between each pair of age groups. None of the SNPs were statistically significantly associated with survival between these three age groups.

Although each was non-significant, we list the ten SNPs with the lowest *p* values for the association with survival between each pair of age groups in Supplementary Tables 2, 3, and 4. SNPs in many different genes appeared in these lists, among which two neutrally selected control SNPs in the list comparing newborns and old individuals. None of these SNPs appeared in more than one of these lists, except for rs2026816 in the insulin receptor substrate 2 gene (*IRS2*) and rs2069842 in the interleukin 6 gene (*IL6*), both of which were in the lists comparing newborns and old individuals and comparing newborns and middle-aged adults.

Figure 3 reports on the investigation of the genetic variants as determinants of variation in fertility. We assessed the associations of each SNP with observed

fertility in middle-aged women and with reported fertility in postmenopausal women. None of the SNPs were statistically significantly associated with either observed or reported fertility.

Table 1. General characteristics of the Ghanaian study population

	Men and women			Women	
	0 years	**20-44 years**	**≥ 60 years**	**20-44 years**	**≥ 45 years**
Individuals, *n*	1482	1589	1144	732	708
Females, *n* (%)	695 (46.9)	1394 (87.7)	608 (53.1)	732 (100.0)	708 (100.0)
Age, years	0 (0–0)	33 (26–40)	70 (65–77)	33 (27–37)	63 (56–71)
Tribe, *n* (%)					
Bimoba	1017 (68.6)	1124 (70.7)	696 (60.8)	549 (75.0)	446 (63.0)
Kusasi	367 (24.8)	365 (23.0)	357 (31.2)	142 (19.4)	205 (29.0)
Other	98 (6.6)	100 (6.3)	91 (8.0)	41 (5.6)	57 (8.1)
Observed fertility	NA	NA	NA	1 (1–2)	NA
Reported fertility	NA	NA	NA	NA	8 (6–9)

Data are given as numbers with percentages or as medians with interquartile ranges. Observed fertility is expressed as the number of children that a woman gave birth to during the period of follow-up. Reported fertility is expressed as the number of children that a woman had given birth to during life. NA: not applicable.
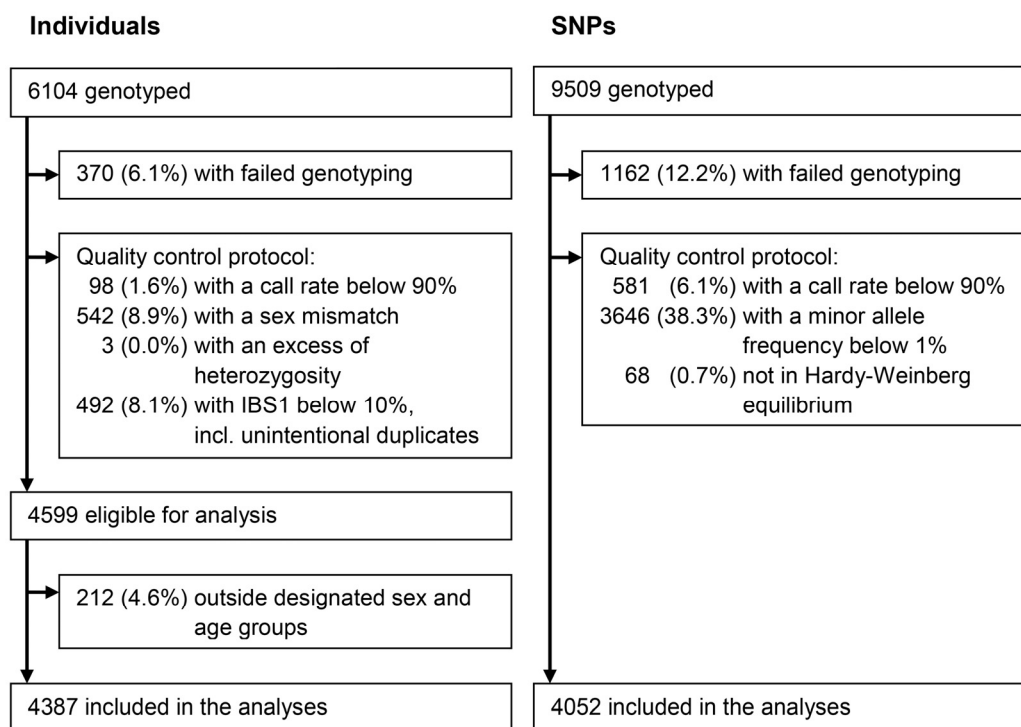

Figure 1. Summary of the exclusions and inclusions of individuals and SNPs.
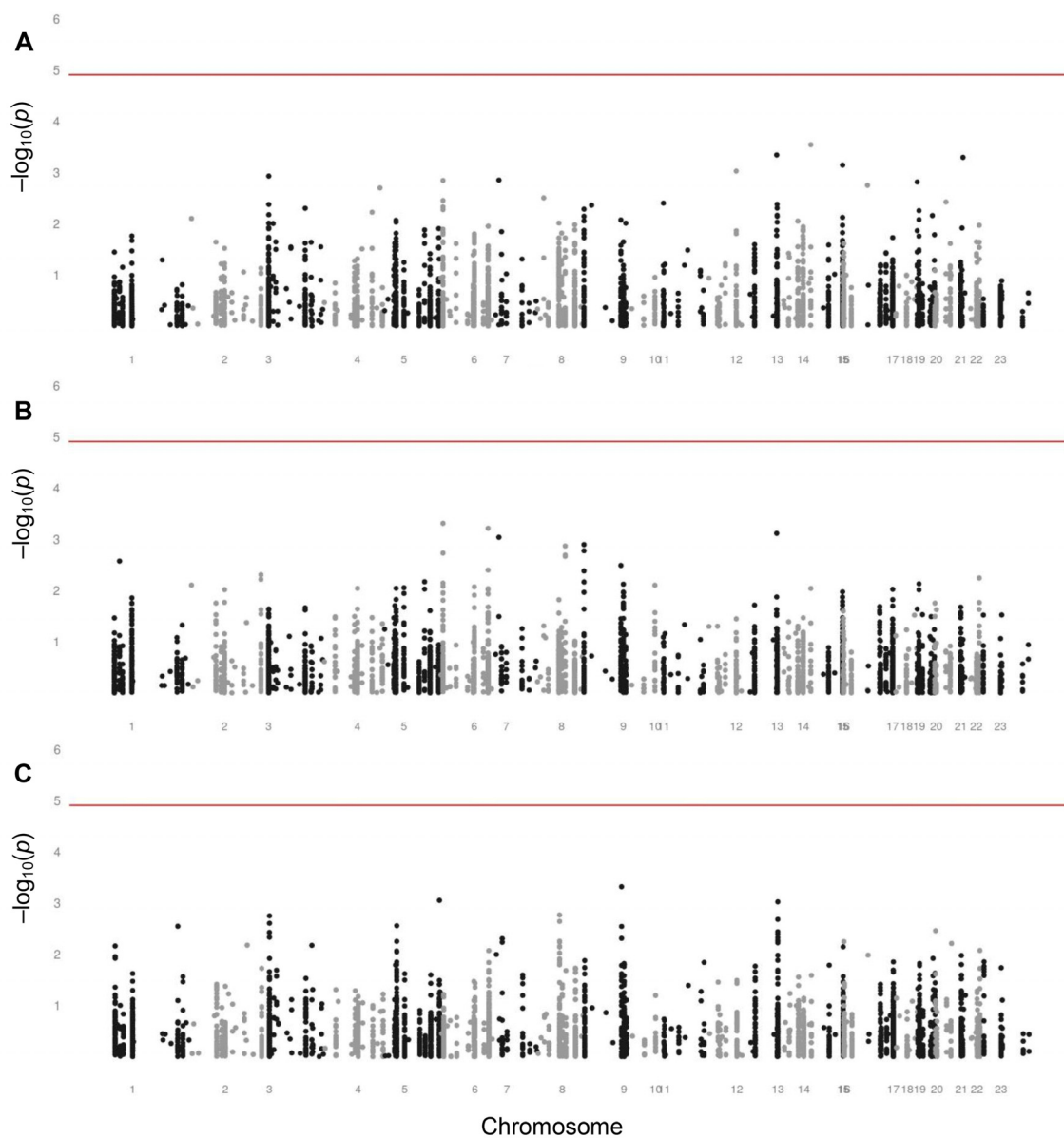
Figure 2. Manhattan plots assessing the associations of SNPs with survival. (**A**) Manhattan plot assessing the associations of SNPs with survival between newborns and old individuals aged 60 years or over. (**B**) Manhattan plot assessing the associations of SNPs with survival between newborns and middle-aged adults of fertile ages from 20 through 44 years. (**C**) Manhattan plot assessing the associations of SNPs with survival between middle-aged adults of fertile ages from 20 through 44 years and old individuals aged 60 years or over. The analyses were adjusted for sex. The level of significance is 1.23×10^{-5}, indicated by the red lines.

Although each was non-significant, we list the ten SNPs with the lowest p values for the associations with observed and reported fertility in Supplementary Tables 5 and 6. SNPs in many different genes appeared in these lists. Among them was a neutrally selected control SNP in the list for the association with observed fertility in middle-aged women. None of the SNPs appeared in both lists.
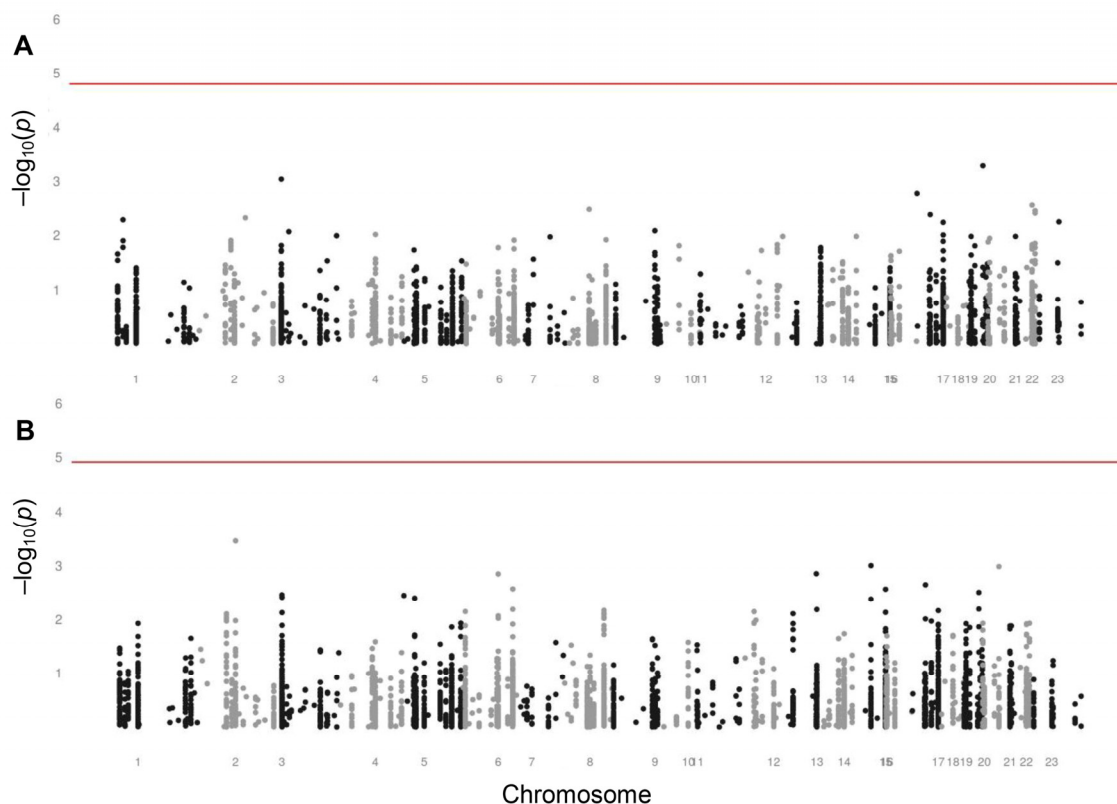
Figure 3. Manhattan plots assessing the associations of SNPs with fertility in women. (**A**) Manhattan plot assessing the associations of SNPs with observed fertility in middle-aged women of fertile ages from 20 through 44 years. The level of significance is 1.61×10^{-5}, indicated by the red line. (**B**) Manhattan plot assessing the associations of SNPs with reported fertility in postmenopausal women aged 45 years and older. The level of significance is 1.23×10^{-5}, indicated by the red line.

DISCUSSION

The aim of this study was to identify the genetic variants that determine life history through variation in survival and fertility in a traditional rural African population that lives in an adverse environment without a western lifestyle. We conducted a large-scale candidate gene study using a high density of SNPs. We found no statistically significant association of any genetic variant with either survival or fertility.

Studies on the genetic determinants of variation in human survival and fertility in adverse environments have never been executed with numbers of individuals and SNPs comparable to this study. Due to a lack of previous genetic studies in this region, we could not rely on standard genome-wide association analyses given the higher degree of population stratification and the lesser degree of linkage disequilibrium in African populations [30]. Instead, we used a custom-made array for geno-

typing a high density of SNPs in the genes of interest. A relatively large proportion of SNPs failed genotyping or had an insufficient minor allele frequency. The exceptional circumstances that are inherent to adverse environments – such as an inadequate infrastructure, a missing civil registry, and language and culture barriers – compromised the study's execution. The proportion of individuals with failed genotyping was comparable with studies in western affluent environments, but we rigidly excluded a relatively large proportion that appeared to have a mismatch between their genetic and registered sexes or to be unintentionally duplicated. To enlarge the contrast in the analysis of the SNPs' associations with survival the selection of old individuals could have been restricted to higher ages, but this would have diminished the number of included old individuals and the power of the analysis. A post hoc calculation of the present study's power, based on the numbers of individuals and SNPs included in the analyses after the quality control using an additive logistic regression

model, reveals that we could find statistically significant associations with an odds ratio of at least 1.4 with a power of 80% and a minimal minor allele frequency ranging from 0.08 to 0.29 (Supplementary Figure 1). As we did not find such statistically significant associations, the associations are possibly present, but less strong.

Population stratification was unlikely to explain our results. As a result of the polygynous and patrilocal culture in the research area, men preferentially marry women from outside the region. Previous analyses of this study population have confirmed that the female-mediated gene flow is nearly fully random and prevents population stratification of autosomal genes [31]. To account for possible population stratification in this study, we rigidly excluded individuals that had a different allele at less than 10% of the loci indicated by an IBS1 below 10% and SNPs that were not in Hardy-Weinberg equilibrium. In addition, we adjusted our analyses for tribe, which represents a mainly cultural population stratification, and for the main dimensions of the multidimensional scaling (MDS) analysis, which represents a mainly genetic population stratification. The MDS analysis did not reveal any population stratification. The adjustments did not affect our results.

Here we did not find genetic variants that were statistically significantly associated with variation in survival or fertility. This finding corresponds with the difficulty to identify such variants in modern affluent environments. In such environments, variation in human lifespan is genetically determined for only less than 30%. Many and various studies have yielded consistent evidence for only a handful of genetic variants to determine variation in survival, most notably *APOE,* and most of their associations with survival are not strong [8-10]. Genetic variants that determine infertility have been identified, but genetic variants that determine variation in fertility have rarely been reported [14]. These disappointing results have been explained by the proposition that survival and fertility are determined by rare genetic variants with a strong effect or by a complex of interacting genetic variants with small effects. Both remain undetected in genome-wide or large-scale candidate gene association studies, such as our study [10, 11].

Contrary to many studies in affluent populations, the present study did not identify *APOE* as a determinant of variation in survival. Three SNPs in *APOE* were included in the present analyses: rs1081101, rs877973, and rs769450. The *p* values for their associations with survival ranged between 0.36 and 1.00. In West Africans, only the latter of these SNPs is in linkage

disequilibrium (r^2=0.21) with rs429358, the SNP that constitutes the *APOEε4* allele and most consistently determines variation in survival in modern affluent populations, but contrary to rs429358, it does not influence blood lipid levels. The other two SNPs are not in linkage disequilibrium with rs429358 and influence blood triglyceride, but not cholesterol levels [32].

Our finding that genetic variants known to influence survival in modern affluent populations, such as *APOE,* did not influence survival in adverse environments may likely be explained, apart from chance, by gene-environment interaction. The effects of genes vary depending on environmental conditions, which thus determine the effects of genes on survival and fertility [33]. Our study was conducted in an environment characterised by endemic infectious diseases, shortages of food, necessary physical activity, and a scarcity of cardiovascular disease and diabetes up to high ages [20-29]. These characteristics differ radically from those of modern affluent environments. Variants of *APOE* affect blood lipid levels, the risk of cardiovascular disease, and survival in affluent populations [34]. In this study population, lipid levels as well as the risk of cardiovascular disease are far lower than in affluent populations [25-28], which may explain why *APOE* does not affect survival here. The higher levels of physical activity may provide an additional explanation for the absence of such an effect [35]. Likewise, we have previously shown that variants of *IL10,* associated with the inflammatory strength of the immune response, enhance survival in those exposed to contaminated drinking water, but diminish survival in those exposed to clean drinking water [36]. These examples of different effects of genetic variants in different environments indicate that such effects can only be identified if, firstly, a specific hypothesis is formulated, secondly, the environmental conditions are measured with the same rigour as the genetic variants, and, thirdly, corresponding appropriate statistical methods are applied.

Life histories have evolved as natural selection has optimised fitness by increasing the frequencies of genetic variants that enhance survival and fertility, while decreasing the frequencies of genetic variants that diminish survival and fertility. The effects of genetic variants on survival and fertility, and thereby the evolution of life history, are largely dependent on environmental factors [2]. The diversity in life history across species indicates that it has a genetic basis, but the diversity in the genetic variants that determine variation in survival and fertility across populations and environ-ments within species indicates that these genetic variants differ across species, populations, and environments.

Evolutionary theory predicts that most of the genetic variants that enhance survival diminish fertility or vice versa, because a trade-off exists between investments in survival and investments in fertility [37, 38]. Indeed, genetic variants have been described that influence both survival and fertility [39, 40]. The close relation between survival and fertility during evolution has encouraged us to study them jointly. Furthermore, several evolutionary theories explain why survival and fertility decline with age in humans and many other species. According to one theory, random damage accumulates in the genetic determinants of survival and fertility as natural selection loses its strength [37, 38]. Such genetic damage as determinants of life history that differ across individuals cannot be identified by a study like the present. According to another theory, some genetic determinants that enhance survival and fertility at early ages have an antagonistic pleiotropic effects that diminish survival and fertility at later ages [37, 38]. The present study would have been able to identify such genetic determinants.

It is critical to acknowledge that natural selection works to maximize fitness as a phenotype rather than a genotype, with survival and fertility being the most important components of this phenotype. The genotype only matters to natural selection as a determinant of the phenotype that is under selection. Meanwhile, the phenotype may be the result of a complex interaction between various genetic variants. Natural selection exists by virtue of such variation in the genetic determinants. Genetic variation is conserved, since genes are prone to mutations and new genetic variants are introduced by sexual reproduction. The resulting genetic heterogeneity of populations may explain why so few unique genetic variants have been found to determine variation in the complex phenotypes of survival and fertility. If different genetic variants determine a similar phenotype, each of these genetic variants is shared by only a proportion of the individuals and a patchwork of shared and unshared variants is established throughout the population. Moreover, if different genetic variants interact with each other to express a phenotype, the penetrance of one of these genetic variants relative to another may vary throughout the population. When studied in a population as a whole, these patterns lead to annulation of the effect of a single genetic variant. As an example of genetic heterogeneity underlying a single phenotype, the ability to digest milk after childhood due to lactase persistence has independently evolved multiple times in similar environments of animal domestication. Across populations, various SNPs at different positions in the lactase gene bring about this trait. Still, these specific variants are insufficient to explain differences in the

frequencies of lactase persistence across populations [41, 42]. Methods to analyse multiple interacting genetic variants are possible, but beyond the aim of this study. More fundamentally, the complexity of the genotype to give rise to a phenotype suggests that, next to explaining genetic variation in survival and fertility, it may be more worthwhile to search for the causal biological mechanisms that determine survival and fertility.

In conclusion, we aimed to identify the genetic variants that determine life history through variation in survival and fertility in an adverse environment in rural Africa, which resembles the environments during most of recent human evolution. In this large-scale candidate gene study, we did not find statistically significant associations of genetic variants with survival or fertility. Apart from a lack of power, we hypothesise that genetic heterogeneity of complex phenotypes and gene-environment interaction prevent the identification of such unique genetic variants that humans have been selected for.

METHODS

Study population. This study was conducted in the Garu-Tempane District in the Upper East Region in Ghana. The region is rural, remote, and one of the least developed in the country. The vast majority of the inhabitants are involved in subsistence agriculture performed by manual labour without proper means of transportation or mechanized farming. The mean annual per capita income and expenditure in the region are one third of those in Ghana nationally and one fifth of those in the capital Accra [43]. Of the adult inhabitants, 31% has attended school as compared with 69% in Ghana nationally and 89% in the capital Accra [43]. Hospital care is absent. Various infectious diseases – including malaria, measles, meningitis, tuberculosis, typhoid fever, trachoma, and intestinal helminths – are highly endemic and constitute the main causes of death both in childhood and adulthood, although the prevalence of human immunodeficiency virus (HIV) is low (<4%) compared with other African regions [44].

From 2002 through 2011, we kept a demographic registry of the population in a research area of 375 km^2 comprising 32 villages. During annual visits we registered the name, age, sex, tribe, and location of living of each inhabitant. If an inhabitant's age was unknown, it was estimated by oral methods, as described previously [20, 21]. Households were occupied by extended families with 48% of the married men having multiple wives [45]. Annual migration relative to the study population's size was 2% into and

1% out of the research area. The average property of the households included small numbers of cattle and bicycles with a value of circa 1,000 US$ and 15% of the households had access to electricity [22, 45]. Drinking water was drawn from boreholes, open wells, and rivers [36]. Of apparently healthy adults, 86% were infected by the malaria species *Plasmodium falciparum*, 44% by the protozoan *Giardia lamblia*, and 31% by the helminth *Necator americanus* [24]. During the nine years of follow-up, 46 to 53% of the population was aged less than 15 years and 6 to 7% of the population was aged 60 years and more [28].

Ethical approval was given by the Committee Medical Ethics of the Leiden University Medical Center, the Ethical Review Committee of Ghana Health Services, and the local chiefs and elders. Because of illiteracy, informed consent was obtained orally from the participants after explanation of the purpose and conduction of this research project. The data were analysed anonymously. This study was conducted in accordance with the Declaration of Helsinki.

Survival and fertility. Survival of all inhabitants was registered during the annual follow-up from 2002 through 2011 [28]. Fertility of women was measured through their reproduction in two manners. Firstly, observed fertility was registered prospectively for women of all ages during the annual follow-up from 2002 through 2011. Observed fertility was expressed as the number of children that a woman gave birth to during the period of follow-up [28, 46]. Secondly, reported fertility was retrospectively determined in 2003 by interviewing women who were available and willing to participate. Data on reported fertility was restricted to women aged 45 years and older, who were considered to be postmenopausal and for whom reported fertility represented lifelong reproduction. Reported fertility was expressed as the number of children that a woman had given birth to during life [21, 47].

DNA collection, isolation, processing, and genotyping. To identify genetic variants associated with survival, we aimed to contrast the genotypes of newborns, of middle-aged individuals, and of individuals who had survived to old age. To identify genetic variants associated with fertility, we aimed to associate the genotypes of middle-aged and postmenopausal women with their reported or observed fertility. Since 2003, we took buccal samples from all newborns who were present during our visits, older than one week, and born in the same year. We took buccal samples from men and women of the fertile ages from 20 through 44 years, from women aged 45 years and older, and from men and women aged 60 years and older. The buccal samples of middle-aged and

older individuals were collected together with measurements of phenotypic characteristics, including fertility [21, 47], infectious diseases [24], inflammatory and metabolic markers in blood [23, 25, 26, 36], cardiovascular health [25-27], muscle strength [29, 48]. These characteristics had been measured in randomly selected individuals. In addition, we collected buccal samples in 2010 from individuals randomly selected from our demographic registry to obtain balanced numbers in the three age groups. With an eye to the analyses of fertility, we oversampled middle-aged women over men.

The buccal samples were stored in 2.5 ml STE buffer (100 mM NaCl, 10 mM Tris/HCl, 10 mM EDTA, pH 8.0) with 0.05 mg/ml proteinase K, 0.1 mg/ml pronase, and 0.5% sodium dodecylsulphate. DNA was isolated and processed from samples collected in 2002 through 2006 by BaseClear (Leiden, the Netherlands) using a Chemagic bead-extraction method, from samples collected in 2006 through 2008 by the Department of Molecular Epidemiology of Leiden University Medical Center (Leiden, the Netherlands) using Qiagen silica spin-columns, and from samples collected in 2007 through 2011 by LGC Genomics, formerly KBioscience (Middlesex, UK), using a proprietary silica column method. DNA was genotyped using a custom-made Illumina Infinium iSelect High-Density Custom Genotyping BeadChip (Illumina, San Diego, CA) at the Department of Human Genetics of Leiden University Medical Center following the manufacturer's instructions.

Candidate gene selection. For the genotyping and analyses, we selected 153 candidate genes that were considered relevant for regulation of life history based on literature, discussion with experts, and general inference. These candidate genes included genes associated with survival or fertility in genome-wide association studies in humans or model organisms [49], human genes that have been under positive or balancing selection pressure during the last 100,000 years and that are associated with survival or fertility [15, 50, 51], and homologues of genes associated with survival in model organisms [52].

SNP selection. We aimed for a dense coverage of SNPs in the candidate genes. We selected SNPs that are known to be causally associated with survival or fertility in humans, had a minor allele frequency equal to or higher than 1% and a correlation equal to or higher than 0.8 in the Yoruba population [53], and/or had a genotype score higher than 0.8 according to the Illumina Design Tool (Illumina, San Diego, CA). If necessary, we chose tag SNPs using Tagger [54] and/or

OpenHelix Genome Variation Server [55] based on the Yoruba population [53]. We additionally selected 170 control SNPs that are presumed to be selectively neutral for analysis of population stratification [56, 57]. The median (interquartile range) number of 1000 base pairs (kb) between genotyped SNPs was 0.8 kb (0.8-1.0 kb).

Quality control. After genotyping, buccal samples and SNPs were subjected to a stringent quality control protocol. Of the 6104 individuals from whom buccal samples had been taken, we excluded 370 (6.1%) because of genotyping failure. Furthermore, we excluded 98 (1.6%) individuals with a call rate below 90%, 542 (8.9%) individuals with a sex mismatch, 3 (< 0.1%) individuals with an excess of heterozygosity indicated by an inbreeding coefficient below -0.3 or above 0.3, and 492 (8.1%) individuals with a different allele at less than 10% of the loci indicated by an IBS1 below 10%, which included unintentional duplicates. Of the 4599 individuals eligible for analysis, 212 (4.6%) individuals outside the designated sex and age groups were excluded. Of the 9509 genotyped SNPs, we excluded 1162 (12.2%) because of genotyping failure. Furthermore, we excluded 581 (6.1%) SNPs with a call rate below 90%, 3646 (38.3%) SNPs with a minor allele frequency below 1%, and 68 (0.7%) SNPs that were not in Hardy-Weinberg equilibrium. As a result, 4387 individuals and 4052 SNPs were included in the analyses. Of these SNPs, 98 (2.4%) were evolutionarily neutrally selected controls. The exclusions and inclusions are described in Figure 1.

As previously reported for this study population, population stratification is unlikely to influence any associations with genetic variation in autosomal genes. As a result of the polygynous and patrilocal culture in the research area, the female-mediated gene flow is nearly fully random [31]. Multidimensional scaling (MDS) analysis did not reveal any population stratification. Potential residual population stratification was addressed in the analyses.

Analyses. For the investigation of the genetic determinants of survival, we assessed the association of each SNP with the chance of being in one of three age groups: newborns, middle-aged adults of fertile ages from 20 through 44 years, and old individuals aged 60 years or over. We compared the chances between pairs of these groups using an additive logistic regression model adjusted for sex. For the investigation of the genetic determinants of fertility, we assessed the associations of each SNP with observed fertility in middle-aged women of fertile ages from 20 through 44 years and with reported fertility in postmenopausal women aged 45 years and older. We assessed the

association with reported fertility using an additive linear regression model. We assessed the association with observed fertility using an Andersen-Gill model, which is an extension to the Cox regression model for analysis of recurrent events [58]. The model was adjusted for calendar year in order to account for the decline in fertility observed during the period of follow-up [28]. To avoid convergence of this model it was necessary to exclude 954 SNPs with a minor allele frequency below 5% from this assessment. All models were repeated with additional adjustment for tribe as a categorical covariate, with additional adjustment for the first two dimensions of the MDS analysis, with additional clustering per household, or with a combination of these in order to account for potential population stratification. These additional adjustments did not alter the results. After a Bonferroni correction for the number of SNPs included in the analyses the threshold for significant results was set at 1.23×10^{-5} for the logistic and linear regression models and at 1.61×10^{-5} for the Andersen-Gill model. The analyses were performed using R (R Foundation, Vienna, Austria).

Author contributions

Conception and design of the study: PdK and RGJW. Collection of the data: JJEK, JP, UKE, and HSF. Analysis of the data: JJEK, JP, and SB. Interpretation of the results: all authors. Drafting of the manuscript: JJEK and JP. All authors provided intellectual contributions and read and approved the manuscript.

Conflict of interest statement

The authors have no conflict of interests to declare.

REFERENCES

1. Jones OR, Scheuerlein A, Salguero-Gómez R, Camarda CG, Schaible R, Casper BB, Dahlgren JP, Ehrlén J, García MB, Menges ES, Quintana-Ascencio PF, Caswell H, Baudisch A, and Vaupel JW. Diversity of ageing across the tree of life. Nature. 2014; 505:169-173.

2. Stearns SC. The Evolution of Life Histories. Oxford, UK: Oxford University Press, 1992.

3. Kenyon CJ. The genetics of ageing. Nature. 2010; 464:504-12.

4. Soerensen M, Dato S, Tan Q, Thinggaard M, Kleindorp R, Beekman M, Jacobsen R, Suchiman HED, De Craen AJM, Westendorp RGJ, Schreiber S, Stevnsner T, Bohr VA, et al. Human longevity and variation in GH/IGF-1/insulin signaling, DNA damage signaling and repair and pro/antioxidant pathway genes: cross sectional and longitudinal studies. Exp Gerontol. 2012; 47:379-87.

5. Vijg J and Suh Y. Genetics of longevity and aging. Annu Rev Med. 2005; 56:193-212.

6. Nussey DH, Froy H, Lemaitre JF, Gaillard JM, and Austad SN. Senescence in natural populations of animals: widespread evidence and its implications for bio-gerontology. Ageing Res Rev. 2013; 12:214-25.

7. Van Heemst D, Beekman M, Mooijaart SP, Heijmans BT, Brandt BW, Zwaan BJ, Slagboom PE, Westendorp RGJ. Reduced insulin/IGF-1 signalling and human longevity. Aging Cell. 2005; 4:79-85.

8. Deelen J, Beekman M, Capri M, Franceschi C, and Slagboom PE. Identifying the genomic determinants of aging and longevity in human population studies: progress and challenges. BioEssays. 2013; 35:386-96.

9. Sebastiani P, Bae H, Sun FX, Andersen SL, Daw EW, Malovini A, Kojima T, Hirose N, Schupf N, Puca A, and Perls TT. Meta-analysis of genetic variants associated with human exceptional longevity. Aging (Albany NY). 2013; 5:653-61. doi: 10.18632/aging.100594

10. Broer L and Van Duijn CM. GWAS and meta-analysis in aging/longevity. Adv Exp Med Biol. 2015; 847:107-125.

11. Pilling LC, Atkins JL, Bowman K, Jones SE, Tyrrell J, Beaumont RN, Ruth KS, Tuke MA, Yaghootkar H, Wood AR, Freathy RM, Murray A, Weedon MN, et al. Human longevity is influenced by many genetic variants: evidence from 75,000 UK Biobank participants. Aging (Albany NY). 2016; 8:597-613. doi: 10.18632/aging.100930

12. Layman LC. Human gene mutations causing infertility. J Med Genet. 2002; 39:153-61.

13. Matzuk MM and Lamb DJ. Genetic dissection of mammalian fertility pathways. Nat Cell Biol. 2002; 4:41-49.

14. Kosova G, Scott NM, Niederberger C, Prins GS, and Ober C. Genome-wide association study identifies candidate genes for male fertility traits in humans. Am J Hum Genet. 2012; 90:950-61.

15. Barreiro LB and Quintana-Murci L. From evolutionary genetics to human immunology: how selection shapes host defence genes. Nat Rev Genet. 2010; 11:17-30.

16. Karlsson EK, Kwiatkowski DP, and Sabeti PC. Natural selection and infectious disease in human populations. Nat Rev Genet. 2014; 15:379-93.

17. O'Keefe JH and Cordain L. Cardiovascular disease resulting from a diet and lifestyle at odds with our paleolithic genome: how to become a 21st-century hunter-gatherer. Mayo Clin Proc. 2004; 79:101-08.

18. Pijl H. Obesity: evolution of a symptom of affluence. Neth J Med. 2011; 69:159-66.

19. Pinker S. The Better Angels of our Nature: Why Violence Has Declined. New York, NY: Viking, 2011.

20. Meij JJ, De Craen AJM, Agana J, Plug D, and Westendorp RGJ. Low-cost interventions accelerate epidemiological transition in Upper East Ghana. Trans R Soc Trop Med Hyg. 2009; 103:173-78.

21. Meij JJ, Van Bodegom D, Ziem JB, Amankwa J, Polderman AM, Kirkwood TBL, De Craen AJM, Zwaan BJ, and Westendorp RGJ. Quality-quantity trade-off of human offspring under adverse environmental conditions. J Evol Biol. 2009; 22:1014-23.

22. Van Bodegom D, May L, Kuningas M, Kaptijn R, Thomése F, Meij HJ, Amankwa J, and Westendorp RGJ. Socio-economic status by rapid appraisal is highly correlated with mortality risks in rural Africa. Trans R Soc Trop Med Hyg. 2009; 103:795-800.

23. May L, Van den Biggelaar AHJ, Van Bodegom D, Meij JJ, De Craen JJ, Amankwa J, Frölich M, Kuningas M, and Westendorp RGJ. Adverse environmental conditions influence age-related innate immune responsiveness. Immun Ageing. 2009; 6:7.

24. Boef AGC, May L, Van Bodegom D, Van Lieshout L, Verweij JJ, Maier AB, Westendorp RGJ, and Eriksson UK. Parasitic infections and immune function: effect of helminth infections in a malaria-endemic area. Immunobiology. 2013; 218:706-11.

25. Eriksson UK, Van Bodegom D, May L, Boef AGC, and Westendorp RGJ. Low C-reactive protein levels in a traditional West-African population living in a malaria-endemic area. PLoS One. 2013; 8:70076.

26. Koopman JJE, Van Bodegom D, Jukema JW, and Westendorp RGJ. Risk of cardiovascular disease in a traditional African population with a high infectious load: a population-based study. PLoS One. 2012; 7:46855.

27. Koopman JJE, Van Bodegom D, Westendorp RGJ, and Jukema JW. Scarcity of atrial fibrillation in a traditional African population: a community-based study. BMC Cardiovasc Disord. 2014; 14:87.

28. Engelaer FM, Koopman JJE, Van Bodegom D, Eriksson UK, and Westendorp RGJ. Determinants of the epidemiologic transition in rural Africa: the role of socioeconomic status and drinking water source. Trans R Soc Trop Med Hyg. 2014; 108:372-79.

29. Koopman JJE, Van Bodegom D, Van Heemst D, and Westendorp RGJ. Handgrip strength, ageing and mortality in rural Africa. Age Ageing. 2015; 44:465-70.

30. Campbell MC and Tishkoff SA. African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. Annu Rev Genomics Hum Genet. 2008; 9:403-33.

31. Sanchez-Faddeev H, Pijpe J, Van der Hulle T, Meij HJ, Van der Gaag KJ, Slagboom PE, Westendorp RGJ, and De Knijff P. The influence of clan structure on the genetic variation in a single Ghanaian village. Eur J Hum Genet. 2013; 21:1134-39.

32. Radwan ZH, Wang X, Waqar F, Pirim D, Niemsiri V, Hokanson JE, Hamman RF, Bunker CH, Barmada MM, Demirci FY, and Kamboh MI. Comprehensive evaluation of the association of APOE genetic variation with plasma lipoprotein traits in US whites and African blacks. PLoS One. 2014; 9:114618.

33. Hughes KA and Burleson MH. Evolutionary causes of genetic variation in fertility and other fitness components. In: Rodgers JL, Rowe DC, and Miller WB (eds). Genetic Influences on Human Fertility and Sexuality. New York, NY: Springer, 2000: 7-33.

34. Ang LS, Cruz RP, Hendel A, and Granville DJ. Apolipoprotein E, an important player in longevity and age-related diseases. Exp Gerontol. 2008; 43:615-22.

35. Raichlen DA and Alexander GE. Exercise, APOE genotype, and the evolution of the human lifespan. Trends Neurosci. 2014; 37:247-55.

36. Kuningas M, May L, Tamm R, Van Bodegom D, Van den Biggelaar AHJ, Meij JJ, Frölich M, Ziem JB, Suchiman HED, Metspalu A, Slagboom PE, and Westendorp RGJ. Selection for genetic variation inducing pro-inflammatory immune responses under adverse environmental conditions in a Ghanaian population. PLoS One. 2009; 4:7795.

37. Westendorp RGJ and Kirkwood TBL. Human longevity at the cost of reproductive success. Nature. 1998; 396:743-46.

38. Kirkwood TBL and Austad SN. Why do we age? Nature. 2000; 408:233-38.

39. Corbo RM, Ulizzi L, Piombo L, and Scacchi R. Study on a possible effect of four longevity candidate genes (ACE, PON1, PPAR-γ, and APOE) on human fertility. Biogerontology. 2008; 9:317-23.

40. Wainer-Katsir K, Zou JY, and Linial M. Extended fertility and longevity: the genetic and epigenetic link. Fertil Steril. 2015; 103:1117-24.

41. Tishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, Silverman JS, Powell K, Mortensen HM, Hirbo JB, Osman M, Ibrahim M, Omar SA, Lema G, et al. Convergent adaptation of human lactase persistence in Africa and Europe. Nat Genet. 2007; 39:31-40.

42. Itan Y, Jones BL, Ingram CJE, Swallow DM, and Thomas MG. A worldwide correlation of lactase persistence phenotype and genotypes. BMC Evol Biol. 2010; 10:36.

43. Ghana Living Standards Survey: Report of the Fifth Round (GLSS 5). Accra, Ghana: Ghana Statistical Service, 2008.

44. Ghana Health Service (GHS). Ghana Health Service 2012 Annual Report Upper East Region. Accra, Ghana, 2012.

45. Van Bodegom D, Eriksson UK, Houwing-Duistermaat JJ, and Westendorp RGJ. Clustering of child mortality in a contemporary polygamous population in Africa. Biodemography Soc Biol. 2012; 58:162-72.

46. Van Bodegom D, Rozing MP, May L, Meij HJ, Thomése F, Zwaan BJ, and Westendorp RGJ. Socioeconomic status determines sex-dependent survival of human offspring. Evol Med Public Health. 2013; 2013:37-45.

47. May L, Van Bodegom D, Frölich M, Van Lieshout L, Slagboom PE, Westendorp RGJ, and Kuningas M. Polymorphisms in TLR4 and TLR2 genes, cytokine production and survival in rural Ghana. Eur J Hum Genet. 2010: 18:490-95.

48. Beenakker KGM, Koopman JJE, Van Bodegom D, Kuningas M, Slagboom PE, Meij JJ, Maier AB, and Westendorp RGJ. Variants of the IL-10 gene associate with muscle strength in elderly from rural Africa: a candidate gene study. Aging Cell. 2014; 13:862-68.

49. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorff L, and Parkinson H. The NHGRI GWAS Catalog: a curated resource of SNP-trait associations. Nucl. Acids Res. 2014; 42:1001-06.

50. Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, Schaffner SF, Lander ES, and International HapMap Consortium. Genome-wide detection and characterization of positive selection in human populations. Nature. 2007; 449:913-18.

51. Gene Ontology Consortium. Gene Ontology Consortium: going forward. Nucl Acids Res. 2015; 43:1049-56.

52. Tacutu R, Craig T, Budovsky A, Wuttke D, Lehmann G, Taranukha D, Costa J, Fraifeld VE, and De Magalhães JP. Human Ageing Genomic Resources: integrated databases and tools for the biology and genetics of ageing. Nucl Acids Res. 2013; 41:1027-33.

53. International HapMap Consortium. The International HapMap Project. Nature. 2003; 426:789-796. Available at http://hapmap.org/.

54. De Bakker PIW, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, and Altshuler D. Efficiency and power in genetic association studies. Nat Genet. 2005; 37:1217-23.

55. Williams JM, Mangan ME, Perreault-Micale C, Lathe S, Sirohi N, and Lathe WC. OpenHelix: bioinformatics education outside of a different box. Brief Bioinform. 2010; 11:598-609.

56. Kosoy R, Nassir R, Tian C, White PA, Butler LM, Silva G, Kittles R, Alarcon-Riquelme ME, Gregersen PK, Belmont JW, De La Vega FM, and Seldin FM. Ancestry Informative marker sets for determining continental origin and admixture proportions in common populations in America. Hum Mutat. 2009; 30:69-78.

57. Pakstis AJ, Speed WC, Fang R, Hyland FC, Furtado MR, Kidd JR, and Kidd KK. SNPs for a universal individual identification panel. Hum Genet. 2010; 127:315-24.

58. Andersen PK and Gill RD. Cox's regression model for counting processes: a large sample study. Ann Statist. 1982; 10:1100-20.

Supplementary Table 1. Candidate genes included in the analyses

Chromo-some	Gene symbol	Gene name	Number of SNPs	Average kb per SNP
14	*AKT1*	V-akt murine thymoma viral oncogene homolog 1	10	3.1
19	*AKT1S1*	AKT1 substrate 1	7	1.9
22	*APOBEC3G*	Apolipoprotein B mRNA editing enzyme catalytic polypeptide-like 3G	4	3.9
11	*APOC3*	Apolipoprotein C-III	5	1.6
19	*APOE*	Apolipoprotein E	3	2.9
X	*AR*	Androgen receptor	36	5.2
13	*ARHGEF7*	Rho guanine nucleotide exchange factor 7	98	2.0
1	*ARID1A*	AT rich interactive domain 1A	21	4.3
21	*BACH1*	BTB and CNC homology 1, basic leucine zipper transcription factor 1	73	4.5
19	*BRSK1*	BR serine/threonine kinase 1	11	3.0
16	*CASKIN1*	Calcium/calmodulin-dependent serine protein kinase (CASK) interacting protein 1	2	4.5
3	*CCR5*	Chemokine (C-C motif) receptor 5	5	2.2
X	*CD40LG*	CD40 ligand	6	2.9
1	*CD55*	CD55	11	4.0
7	*CFTR*	Cystic fibrosis transmembrane conductance regulator	3	64.6
6	*CGA*	Glycoprotein hormones α polypeptide	4	2.4
11	*CRACR2B*	Calcium release activated channel regulator 2B	4	2.7
16	*CREBBP*	cAMP response element-binding protein (CREB) binding protein	60	2.7
1	*DARC*	Duffy blood group	1	6.8
6	*DAXX*	Death-domain associated protein	2	4.7
8	*DEPDC6*	DEP domain containing MTOR-interacting protein	112	1.6
14	*DIO2*	Deiodinase iodothyronine type II	11	1.8

14	DIO3	Deiodinase iodothyronine type III	0	NA
15	DYX1C1	Dyslexia susceptibility 1 candidate 1	38	2.5
4	EIF4E	Eukaryotic translation initiation factor 4E	33	1.7
8	EIF4EBP1	Eukaryotic translation initiation factor 4E binding protein 1	9	3.9
3	EPHA3	Ephrin receptor A3	4	93.7
19	EPOR	Erythropoietin receptor	3	3.8
6	ESR1	Estrogen receptor α	164	2.5
14	ESR2	Estrogen receptor β	51	2.3
20	FKBP1A	Tacrolimus binding protein 1A	20	1.5
14	FOXA1	Forkhead box A1	3	3.4
13	FOXO1	Forkhead box O1	44	2.6
6	FOXO3	Forkhead box O3	68	1.9
X	FOXO4	Forkhead box O4	6	2.1
19	FUT2	Fucosyltransferase 2	6	2.5
X	G6PD	Glucose-6-phosphate dehydrogenase	2	8.1
17	GH1	Growth hormone 1	1	6.6
5	GHR	Growth hormone receptor	100	3.0
20	GHRH	Growth hormone releasing hormone	6	1.8
7	GHRHR	Growth hormone releasing hormone receptor	14	1.5
11	HBB	Haemoglobin β	0	NA
5	HK3	Hexokinase 3	13	1.4
11	HRAS	Harvey rat sarcoma viral oncogene homolog	5	1.7
12	IFNg	Interferon γ	6	1.7
6	IFNGR1	Interferon γ receptor 1	7	3.8
21	IFNGR2	Interferon γ receptor 2	19	2.1
12	IGF1	Insulin-like growth factor 1	32	2.8
15	IGF1R	Insulin-like growth factor 1 receptor	166	1.9
7	IGFBP3	Insulin-like growth factor binding protein 3	6	2.3
2	IL1A	Interleukin 1α	8	2.1
2	IL1B	Interleukin 1β	6	2.0
2	IL1F10	Interleukin 1 family member 10	10	1.3
2	IL1RN	Interleukin 1 receptor antagonist	13	1.6
4	IL2	Interleukin 2	3	3.3
X	IL2RG	Interleukin 2 receptor γ	2	4.6
5	IL4	Interleukin 4	8	1.7

16	*IL4R*	Interleukin 4 receptor	27	2.1
5	*IL5*	Interleukin 5	6	1.2
7	*IL6*	Interleukin 6	2	4.9
8	*IL7*	Interleukin 7	29	2.7
5	*IL7R*	Interleukin 7 receptor	15	1.7
1	*IL10*	Interleukin 10	13	1.8
3	*IL12A*	Interleukin 12A	7	1.7
19	*IL12RB1*	Interleukin 12 receptor β1	14	2.3
5	*IL13*	Interleukin 13	0	NA
6	*IL17A*	Interleukin 17A	6	1.5
2	*IL18RAP*	Interleukin 18 receptor accessory protein	23	1.7
14	*IL25*	Interleukin 25	1	3.6
2	*IL36RN*	Interleukin 36 receptor antagonist	3	3.7
2	*IL37*	Interleukin 37	6	1.8
11	*INS-IGF2*	INS-IGF2 read-through gene	21	1.8
19	*INSR*	Insulin receptor	117	1.6
X	*IRAK1*	Interleukin-1 receptor-associated kinase 1	0	NA
12	*IRAK4*	Interleukin-1 receptor-associated kinase 4	16	2.2
2	*IRS1*	Insulin receptor substrate 1	35	2.1
13	*IRS2*	Insulin receptor substrate 2	23	1.6
9	*JAK2*	Janus kinase 2	53	2.8
2	*KDM3A*	Lysine (K)-specific demethylase 3A	33	1.7
19	*KIR3DL1*	Killer cell immunoglobulin-like receptor, three domains, long cytoplasmic tail 1	3	6.4
19	*KIR3DL3*	Killer cell immunoglobulin-like receptor, three domains, long cytoplasmic tail 3	5	3.4
12	*KMT2D*	Lysine (K)-specific methyltransferase 2D	14	2.4
4	*LEF1*	Lymphoid enhancer-binding factor 1	74	1.7
1	*LEPR*	Leptin receptor	101	2.2
19	*LILRA3*	Leukocyte immunoglobulin-like receptor subfamily A member 3	5	1.9
6	*LIN28B*	Lin-28 homolog B	37	3.5
9	*MAPKAP1*	Mitogen-activated protein kinase associated protein 1	108	2.5
10	*MBL2*	Mannose-binding lectin (protein C) 2	6	1.9
2	*MCM6*	Minichromosome maintenance complex component 6	3	12.3

20	*MCM8*	Minichromosome maintenance 8 homologous recombination repair factor	23	2.2
1	*MTOR*	Mechanistic target of rapamycin	81	2.0
3	*MYD88*	Myeloid differentiation primary response 88	5	1.9
11	*MYO7A*	Myosin VIIA	1	87.0
8	*NCOA2*	Nuclear receptor coactivator 2	123	2.4
14	*NFKBIA*	Nuclear factor of κ light polypeptide gene enhancer in B-cells inhibitor α	7	1.2
8	*NKX3-1*	Homeobox Nkx-3.1	3	3.1
1	*NR0B2*	Nuclear receptor subfamily 0 group B member 2	2	3.8
19	*NR1H2*	Nuclear receptor subfamily 1 group H member 2	6	1.9
11	*NR1H3*	Nuclear receptor subfamily 1 group H member 3	8	3.1
12	*PARPBP*	Poly (ADP-ribose) polymerase 1 binding protein	2	38.7
2	*PDK1*	Pyruvate dehydrogenase kinase isozyme 1	5	13.9
7	*PDK4*	Pyruvate dehydrogenase kinase isozyme 4	15	1.2
16	*PDPK1*	3-phosphoinositide dependent protein kinase 1	11	6.4
19	*PGPEP1*	Pyroglutamyl-peptidase I	6	4.9
3	*PIK3CB*	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit β	43	2.5
5	*PIK3R1*	Phosphoinositide-3-kinase regulatory subunit 1	40	2.0
12	*PMCH*	Pro-melanin-concentrating hormone	3	2.1
5	*PPARGC1B*	Peroxisome proliferator-activated receptor γ coactivator 1β	73	1.7
22	*PRR5*	Proline rich 5	44	1.7
10	*PTEN*	Phosphatase and tensin homolog	17	6.5
20	*PTPN1*	Protein tyrosine phosphatase non-receptor type 1	34	2.3
3	*PTX3*	Pentraxin 3	6	4.2
17	*RARA*	Retinoic acid receptor α	10	5.3
22	*RBFOX2*	RNA binding protein fox-1 homolog 2	92	3.1

5	*RICTOR*	RPTOR independent companion of MTOR complex 2	52	2.7
8	*RIPK2*	Receptor-interacting serine-threonine kinase 2	20	1.9
14	*RNASE3*	Ribonuclease A family 3	1	5.9
17	RPS6KB1	Ribosomal protein S6 kinase polypeptide 1	21	3.0
17	*RPTOR*	Regulatory associated protein of MTOR complex 1	235	1.8
18	*SLC14A1*	Urea transporter member 1 (Kidd blood group)	24	1.4
4	*SPP1*	Secreted phosphoprotein 1	2	3.9
3	*SST*	Somatostatin	3	2.2
22	*SSTR3*	Somatostatin receptor 3	6	1.9
17	*STAT5A*	Signal transducer and activator of transcription 5A	12	2.4
17	*STAT5B*	Signal transducer and activator of transcription 5B	31	2.7
19	*SUV420H2*	Suppressor of variegation 4-20 homolog 2	10	1.3
6	*SYCP2L*	Synaptonemal complex protein 2-like	66	1.4
19	*TCF3*	Transcription factor 3	18	2.4
19	*TGFB1*	Transforming growth factor β1	11	2.6
17	*THRA*	Thyroid hormone receptor α	16	2.3
1	*THRAP3*	Thyroid hormone receptor associated protein 3	26	3.3
3	*THRB*	Thyroid hormone receptor β	179	2.1
19	*TICAM1*	Toll-like receptor adaptor molecule 1	10	2.1
5	*TICAM2*	Toll-like receptor adaptor molecule 2	34	1.5
11	*TIRAP*	Toll-interleukin 1 receptor domain containing adaptor protein	8	2.1
4	*TLR1*	Toll-like receptor 1	4	3.4
4	*TLR2*	Toll-like receptor 2	19	2.0
4	*TLR3*	Toll-like receptor 3	17	1.2
9	*TLR4*	Toll-like receptor 4	33	0.9
1	TLR5	Toll-like receptor 5	20	1.9
4	*TLR6*	Toll-like receptor 6	4	1.9
X	*TLR7*	Toll-like receptor 7	11	2.6
X	*TLR8*	Toll-like receptor 8	9	2.4
3	*TLR9*	Toll-like receptor 9	2	5.0

4	*TLR10*	Toll-like receptor 10	11	1.4
6	*TNF*	Tumour necrosis factor	1	7.8
19	*TOMM40*	Translocase of outer mitochondrial membrane 40 homolog	11	1.6
9	*TSC1*	Tuberous sclerosis 1	22	2.6
16	*TSC2*	Tuberous sclerosis 2	18	2.5
14	*TSHR*	Thyroid stimulating hormone receptor	104	1.9
5	*UIMC1*	Ubiquitin interaction motif containing 1	51	2.1
22	*USP41*	Ubiquitin specific peptidase 41	1	4.0

Candidate genes included in the analyses are given with the number of SNPs associated to these genes included in the analyses. SNPs included as selectively neutral controls are not included in this table. An impression of the density of SNPs per gene is given by the average number of 1000 base pairs (kb) between the SNPs included in the analyses. The sizes of the gene regions were derived from Genome Reference Consortium Human Build 37 patch release 1 (GRCh37.p1).

Supplementary Table 2. SNPs with the lowest *p* values for the association with survival between newborns and old individuals aged 60 years or over

Chromosome	Gene symbol	SNP	MAF	OR	*p* value
14	*AKT1*	rs61758466	0.14	0.73	2.88×10^{-4}
13	*IRS2*	rs2026816	0.32	0.81	4.55×10^{-4}
21	*CLDN14**	rs2835370	0.49	1.21	5.07×10^{-4}
15	*IGF1R*	rs2684796	0.41	0.83	7.17×10^{-4}
12	*PARPBP*	rs17032311	0.01	2.46	9.41×10^{-4}
3	*THRB*	rs75500315	0.02	2.08	1.18×10^{-3}
7	*IL6*	rs2069842	0.11	0.75	1.41×10^{-3}
6	*SYCP2L*	rs9368446	0.03	1.62	1.44×10^{-3}
19	*TCF3*	rs2353219	0.20	0.79	1.54×10^{-3}
16	*CLEC3A**	rs430046	0.27	0.82	1.78×10^{-3}

The analysis was adjusted for sex. MAF: minor allele frequency. OR: odds ratio. * These SNPs were included as selectively neutral controls. The level of significance is 1.23×10^{-5}.

Supplementary Table 3. SNPs with the lowest p values for the association with survival between newborns and middle-aged adults of fertile ages from 20 through 44 years

Chromosome	Gene symbol	SNP	MAF	OR	*p* value
6	*SYCP2L*	rs75788404	0.09	0.70	4.88×10^{-4}
6	*ESR1*	rs1293936	0.46	1.22	6.10×10^{-4}
13	*IRS2*	rs2026816	0.32	0.81	7.67×10^{-4}
7	*IL6*	rs2069842	0.11	0.74	9.15×10^{-4}
9	*JAK2*	rs60763646	0.10	0.74	1.27×10^{-3}
8	*RIPK2*	rs2158131	0.41	1.21	1.36×10^{-3}
9	*JAK2*	rs4372063	0.03	1.68	1.68×10^{-3}
6	*SYCP2L*	rs57579421	0.08	0.73	1.86×10^{-3}
8	*RIPK2*	rs16900627	0.42	1.20	2.05×10^{-3}
8	*RIPK2*	rs73291472	0.04	0.63	2.16×10^{-3}

The analysis was adjusted for sex. MAF: minor allele frequency. OR: odds ratio. The level of significance is 1.23×10^{-5}.

Supplementary Table 4. SNPs with the lowest *p* values for the association with survival between middle-aged adults of fertile ages from 20 through 44 years and old individuals aged 60 years or over

Chromosome	Gene symbol	SNP	MAF	OR	*p* value
9	*TLR4*	rs10818073	0.03	1.90	4.80×10^{-4}
5	*UIMC1*	rs75988289	0.03	1.70	8.90×10^{-4}
13	*ARHGEF7*	rs74691269	0.10	1.38	9.50×10^{-4}
8	*NCOA2*	rs6983366	0.02	0.52	1.71×10^{-3}
3	*THRB*	rs6785472	0.16	1.30	1.78×10^{-3}
13	*ARHGEF7*	rs79847747	0.09	1.36	2.09×10^{-3}
8	*NCOA2*	rs57293541	0.06	0.68	2.29×10^{-3}
3	*THRB*	rs6550862	0.15	1.28	2.49×10^{-3}
5	*GHR*	rs73085419	0.11	1.31	2.79×10^{-3}
1	*CD55*	rs75882193	0.02	1.92	2.85×10^{-3}

The analysis was adjusted for sex. MAF: minor allele frequency. OR: odds ratio. The level of significance is 1.23×10^{-5}.

Supplementary Table 5. SNPs with the lowest *p* values for the association with observed fertility in middle-aged women of fertile ages from 20 through 44 years

Chromosome	Gene symbol	SNP	MAF	OR	*p* value
19	*TGFB1*	rs4803455	0.38	0.78	5.23×10^{-4}
3	*THRB*	rs1505301	0.44	0.79	9.19×10^{-4}
17	*INPP5K**	rs1879488	0.23	1.27	1.68×10^{-3}
22	*RBFOX2*	rs916333	0.12	1.32	2.73×10^{-3}
8	*NCOA2*	rs16936902	0.12	1.32	3.25×10^{-3}
22	*PRR5*	rs2349642	0.40	1.22	3.47×10^{-3}
22	*PRR5*	rs2097535	0.40	1.22	3.81×10^{-3}
17	*STAT5B*	rs9897531	0.08	1.39	4.07×10^{-3}
2	*PABPCP2*	rs10496971	0.07	1.38	4.64×10^{-3}
1	*ARID1A*	rs4589135	0.42	1.22	5.05×10^{-3}

* This SNP was included as a selectively neutral control. MAF: minor allele frequency. OR: odds ratio. The level of significance is 1.61×10^{-5}.

Supplementary Table 6. SNPs with the lowest *p* values for the association with reported fertility in postmenopausal women aged 45 years and older

Chromosome	Gene symbol	SNP	MAF	OR	*p* value
2	*IL1RN*	rs4251987	0.02	4.95	3.50×10^{-4}
15	*DYX1C1*	rs79593205	0.04	2.77	9.97×10^{-4}
20	*PTPN1*	rs77833095	0.05	0.41	1.04×10^{-3}
13	*IRS2*	rs74886884	0.01	7.17	1.41×10^{-3}
6	*FOXO3*	rs75621079	0.10	1.90	1.42×10^{-3}
17	*STAT5B*	rs56938023	0.25	0.64	2.25×10^{-3}
6	*ESR1*	rs17082104	0.13	1.73	2.68×10^{-3}
15	*IGF1R*	rs75090574	0.04	2.61	2.71×10^{-3}
19	*APOE*	rs1081101	0.06	0.46	3.11×10^{-3}
3	*THRB*	rs62255856	0.15	1.63	3.45×10^{-3}

MAF: minor allele frequency. OR: odds ratio. The level of significance is 1.23×10^{-5}.
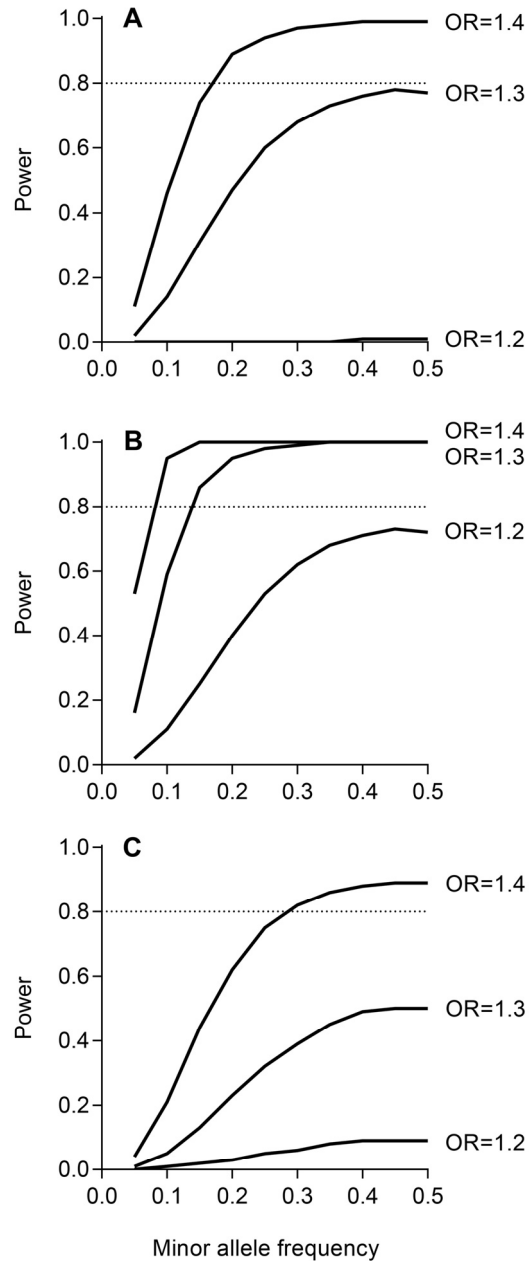
Supplementary Figure 1. A post hoc calculation of this study's power, depending on the strength of the SNP's effect and the minor allele frequency. The effect is expressed as an odds ratio (OR) determined using an additive logistic regression model assessing the associations of SNPs with survival between newborns and old individuals aged 60 and over (**A**), between newborns and middle-age adults of fertile ages from 20 through 44 years (**B**), and between middle-aged adults of fertile ages from 20 through 44 years and old individuals aged 60 years or over (**C**).