Genome-wide analysis identifies genetic effects on reproductive success and ongoing natural selection at the FADS locus

Title: Genome-wide analysis identifies genetic effects on reproductive success and ongoing natural selection at the *FADS* locus.

Iain Mathieson[1,†,*], Felix R. Day[2,†], Nicola Barban[3,†], Felix C. Tropf[4,5,6,7,†], David M. Brazel[4,5,†], *eQTLGen Consortium*, *BIOS Consortium*, Ahmad Vaez[8,9], Natalie van Zuydam[10], Bárbara D. Bitarello[1], Eugene J. Gardner[2], Evelina T. Akimova[4,5], Ajuna Azad[11], Sven Bergmann[12,13,14], Lawrence F. Bielak[15], Dorret I. Boomsma[16], Kristina Bosak[17], Marco Brumat[18], Julie E. Buring[19,20], David Cesarini[21,22,23], Daniel I. Chasman[19,20], Jorge E. Chavarro[24,25,26], Massimiliano Cocca[27], Maria Pina Concas[27], George Davey-Smith[28], Gail Davies[29], Ian J. Deary[29], Tõnu Esko[30,31], Jessica D. Faul[32], FinnGen study, Oscar Franco[33], Andrea Ganna[34,35], Audrey J. Gaskins[36], Andrea Gelemanović[37], Eco J.C. de Geus[16], Christian Gieger[38], Giorgia Girotto[18,27], Bamini Gopinath[39], Hans Jörgen Grabe[40], Erica P. Gunderson[41], Caroline Hayward[42], Chunyan He[43,44], Diana van Heemst[45], W. David Hill[29], Eva R. Hoffmann[11], Georg Homuth[46], Jouke Jan Hottenga[16], Hongyang Huang[24], Elina Hyppönen[47,48], M. Arfan Ikram[33], Rick Jansen[49], Magnus Johannesson[50], Zoha Kamali[8,9], Sharon L.R. Kardia[15], Maryam Kavousi[33], Annette Kifley[39], Tuomo Kiiskinen[34,51], Peter Kraft[24,52], Brigitte Kühnel[38], Claudia Langenberg[2], Gerald Liew[39], Lifelines Cohort Study[8,53], Penelope A. Lind[54], Jian'an Luan[2], Reedik Mägi[30], Patrik K.E. Magnusson[55], Anubha Mahajan[56,57], Nicholas G. Martin[58], Hamdi Mbarek[16,59], Mark I. McCarthy[56,57], George McMahon[60], Sarah E. Medland[54], Thomas Meitinger[61], Andres Metspalu[30,62], Evelin Mihailov[30], Lili Milani[30], Stacey A. Missmer[24,63,64], Paul Mitchell[39], Stine Møllegaard[65], Dennis O. Mook-Kanamori[66,67], Anna Morgan[27], Peter J. van der Most[8], Renée de Mutsert[66], Matthias Nauck[68], Ilja M. Nolte[8], Raymond Noordam[45], Brenda W.J.H. Penninx[69], Annette Peters[70], Patricia A. Peyser[15], Ozren Polašek[37,71], Chris Power[72], Ajka Pribisalić[37], Paul Redmond[29], Janet W. Rich-Edwards[24,26,73], Paul M. Ridker[19,20], Cornelius A. Rietveld[74,75], Susan M. Ring[28], Lynda M. Rose[19], Rico Rueedi[12,13], Vallari Shukla[11], Jennifer A. Smith[15,32], Stasa Stankovic[2], Kári Stefánsson[76], Doris Stöckl[70], Konstantin Strauch[77,78,79], Morris A. Swertz[53], Alexander Teumer[80], Gudmar Thorleifsson[76], Unnur Thorsteinsdottir[76], A. Roy Thurik[74,75,81], Nicholas J. Timpson[28], Constance Turman[24], André G. Uitterlinden[74,82], Melanie Waldenberger[38,70], Nicholas J. Wareham[2], David R. Weir[32], Gonneke Willemsen[16], Jing Hau Zhao[2], Wei Zhao[15], Yajie Zhao[2], Harold Snieder[8], Marcel den Hoed[10], Ken K. Ong[2], Melinda C. Mills[4,5,†,*], and John R.B. Perry[2,†,*]

[1] Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, United States of America

[2] MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge, Cambridge, United Kingdom

[3] Alma Mater Studiorum University of Bologna, Italy

[4] Leverhulme Centre for Demographic Science, University of Oxford, Oxford, United Kingdom

[5] Nuffield College, University of Oxford, Oxford, United Kingdom

[6] École Nationale de la Statistique et de L'administration Économique (ENSAE), Paris, France

[7] Center for Research in Economics and Statistics (CREST), Paris, France

[8] Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

[9] Department of Bioinformatics, Isfahan University of Medical Sciences, Isfahan, Iran

[10] The Beijer Laboratory and Department of Immunology, Genetics and Pathology, Uppsala University and SciLifeLab, Uppsala, Sweden

[11] DNRF Center for Chromosome Stability, Department of Cellular and Molecular Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark

[12] Department of Computational Biology, University of Lausanne, Lausanne, Switzerland

[13] Swiss Institute of Bioinformatics, Lausanne, Switzerland

[14] Department of Integrative Biomedical Sciences, University of Cape Town, Cape Town, South Africa

[15] Department of Epidemiology, University of Michigan, Ann Arbor, MI, United States of America

[16] Department of Biological Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

[17] Psychiatric hospital "Sveti Ivan", Zagreb, Croatia

[18] Department of Medical, Surgical and Health Sciences, University of Trieste, Trieste, Italy

[19] Brigham and Women's Hospital, Boston, MA, United States of America

[20] Harvard Medical School, Boston, MA, United States of America

[21] Department of Economics, New York University, New York, NY, United States of America

[22] Research Institute for Industrial Economics, Stockholm, Sweden

[23] National Bureau of Economic Research, Cambridge, MA, United States of America

[24] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, United States of America

[25] Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, United States of America

[26] Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States of America

[27] Institute for Maternal and Child Health IRCCS "Burlo Garofolo", Trieste, Italy

[28] MRC Integrative Epidemiology Unit, University of Bristol, Bristol, United Kingdom

[29] Lothian Birth Cohorts, Department of Psychology, University of Edinburgh, Edinburgh, United Kingdom

[30] Estonian Genome Center, University of Tartu, Tartu, Estonia

[31] Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, MA, United States of America

[32] Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, United States of America

[33] Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands

[34] Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland

[35] Analytic and Translational Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, United States of America

[36] Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, United States of America

[37] University of Split School of Medicine, Split, Croatia

[38] Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[39] Centre for Vision Research, Westmead Institute for Medical Research and Department of Ophthalmology, University of Sydney, Sydney, Australia

[40] Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Germany

[41] Division of Research, Kaiser Permanente Northern California, Oakland, CA, United States of America

[42] Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom

[43] University of Kentucky Markey Cancer Center, Lexington, KY, United States of America

[44] Department of Internal Medicine, Division of Medical Oncology, University of Kentucky College of Medicine, Lexington, KY, United States of America

[45] Department of Internal Medicine, Section of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands

[46] Interfaculty Institute for Genetics and Functional Genomics, University of Greifswald, Greifswald, Germany

[47] Australian Centre for Precision Health, University of South Australia Cancer Research Institute, Adelaide, Australia

[48] South Australian Health and Medical Research Institute, Adelaide, Australia

[49] Department of Psychiatry, Amsterdam Public Health and Amsterdam Neuroscience, Amsterdam UMC, Vrije Universiteit, Amsterdam, The Netherlands

[50] Department of Economics, Stockholm School of Economics, Stockholm, Sweden

[51] National Institute for Health and Welfare, Helsinki, Finland

[52] Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, United States of America

[53] Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

[54] Psychiatric Genetics, QIMR Berghofer Medical Research Institute, Herston Brisbane, Queensland, Australia

[55] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

[56] Wellcome Centre for Human Genetics, University of Oxford, Oxford, United Kingdom

[57] Oxford Centre for Diabetes, Endocrinology and Metabolism, Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom

[58] Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Herston Brisbane, Queensland, Australia

[59] Qatar Genome Programme, Qatar Foundation Research, Development and Innovation, Qatar Foundation, Doha, Qatar

[60] School of Social and Community Medicine University of Bristol, Bristol, United Kingdom

[61] Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[62] Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia

[63] Division of Adolescent and Young Adult Medicine, Department of Medicine, Boston Children's Hospital and Harvard Medical School, Boston, MA, United States of America

[64] Department of Obstetrics, Gynecology, and Reproductive Biology, College of Human Medicine, Michigan State University, Grand Rapids, MI, United States of America

[65] Department of Sociology, University of Copenhagen, Copenhagen, Denmark

[66] Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

[67] Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, The Netherlands

[68] Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, Germany

[69] Department of Psychiatry, EMGO Institute for Health and Care Research and Neuroscience Campus Amsterdam, VU University Medical Center/GGZ inGeest, Amsterdam, The Netherlands

[70] Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[71] Gen Info, LLC, Zagreb, Croatia

[72] Population, Policy and Practice Research and Teaching Department, UCL Great Ormond Street Institute of Child Health, London, United Kingdom

[73] Division of Women's Health, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States of America

[74] Erasmus University Rotterdam Institute for Behavior and Biology, Rotterdam, The Netherlands

[75] Department of Applied Economics, Erasmus School of Economics, Rotterdam, The Netherlands

[76] deCODE Genetics/Amgen Inc., Reykjavik, Iceland

[77] Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI), University Medical Center, Johannes Gutenberg University, Mainz, Germany

[78] Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[79] Chair of Genetic Epidemiology, IBE, Faculty of Medicine, LMU Munich, Germany

[80] Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany

[81] Montpellier Business School, Montpellier, France

[82] Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands

[†] Denotes equal contribution

* Correspondence to Iain Mathieson, mathi@pennmedicine.upenn.edu, Melinda C. Mills melinda.mills@nuffield.ox.ac.uk, and John R.B. Perry, john.perry@mrc-epid.cam.ac.uk

Abstract

Identifying genetic determinants of reproductive success may highlight mechanisms underlying fertility and also identify alleles under present-day selection. Using data in 785,604 individuals of European ancestry, we identify 43 genomic loci associated with either number of children ever born (NEB) or childlessness. Individual loci are associated with diverse aspects of reproductive biology across the life course, including puberty timing, age at first birth, sex hormone regulation, endometriosis and age at menopause. Missense variants in *ARHGAP27* were associated with higher NEB but shorter reproductive lifespan, suggesting a trade-off at this locus between reproductive ageing and intensity. Other genes implicated by coding variants include *PIK3IP1, ZFP82, LRP4*, and suggest a novel role for the melanocortin 1 receptor (*MC1R*) in reproductive biology. As NEB is one component of evolutionary fitness, our identified associations indicate loci under present-day natural selection. Integration with data from scans for natural selection identifies a unique example of an allele in the *FADS1/2* gene locus that has been under selection for thousands of years and remains so today. Collectively, our findings demonstrate that diverse biological mechanisms contribute to reproductive success, implicating both neuro-endocrine and behavioural factors.

Introduction

Variation in human reproductive behaviour and success is epidemiologically associated with disease risk and has profound psychological, clinical and societal implications. This is particularly true for infertility, where efforts to elucidate the underlying biological mechanisms have been limited by the lack of large, well-phenotyped cohorts with relevant outcomes. This situation is mirrored across many reproductive traits and diseases, such as polycystic ovary syndrome[1,2], where progress to identify genetic determinants and underlying mechanisms has lagged behind that of other complex diseases[3]. One reason for this is that natural selection limits the frequency of fertility-reducing alleles. The number of children ever born to an individual (NEB) has one of the highest degrees of polygenicity of any trait, consistent with a genetic architecture strongly influenced by negative selection[4,5]. Another reason is the fact that NEB is a behavioural phenotype influenced by multiple social, economic and environmental factors[6–8]. Nonetheless, studying the genetic basis of *fertility* may illuminate biological mechanisms underpinning infertility, with the advantage that relevant measures are more readily available. For example, recent studies have identified genetic determinants for NEB, age at first sexual intercourse and age at first birth[9–12]. These have provided several aetiological insights, such as highlighting a neuro-behavioural role for the estrogen receptor in men[9] and identifying biological mechanisms linking reproductive ageing to late-onset diseases[9,10,13,14].

Fertility-associated loci may act through a broad array of mechanisms. They may have direct effects on reproductive biology, or act through traits that contribute to partner selection or other aspects of behaviour and personality. For example, alleles associated with higher educational attainment are associated with lower fertility in some populations[15,16], reflecting the link between higher education and older age at childbearing[7]. Finally, fertility-associated loci might represent alleles under selection for some trait entirely disconnected from reproductive biology. By definition, any variant that is under natural selection affects fitness. Variants that affect fecundity would be detected by a genome-wide scan for NEB, although this scan would not capture all components of fitness.

Our present study substantially builds upon two earlier studies[9,10] to identify individual genetic determinants of NEB. In contrast to these previous studies, we double the sample size to 785,604 individuals and increase the number of genetic loci associated with NEB from 5 to 43. By linking our findings to scans for ancient selection, we isolate a unique example of an allele that has remained under selection for thousands of years and remains under selection today. We also highlight a number of novel biological mechanisms that contribute to reproductive success and uncover a previously unknown role for the melanocortin 1 receptor (*MC1R*) in reproductive biology.

Results

We identified genetic determinants of NEB by performing a genome-wide association study (GWAS) in 785,604 European ancestry individuals meta-analysed across 45 studies (**Table S1-S6**). Detailed methodology for how this discovery analysis was performed can be found in the methods section. SNP array data was imputed to at least 1000 Genomes Project reference panel density (phase 1 version 3) across all studies. The distribution of genome-wide test statistics for NEB showed substantial inflation (λ_{GC}=1.36), however LD score regression[17] indicated that this was attributable to polygenicity rather than population stratification (LD intercept 1.01; s.e. 0.008). In total, 5,283 variants reached genome-wide significance ($P<5x10^{-8}$) for association with NEB, which we resolved to 28 statistically independent signals (**Table S7**). These include all five signals previously reported for NEB in overlapping samples of up to 343,072 individuals[9,10].

The genetic architecture of NEB was only moderately correlated between men and women (r_g=0.74; 95% CI 0.66-0.82). Therefore, we ran separate GWAS meta-analyses in men (N=306,980) and women (N=478,624) and identified six additional statistically independent signals (two in men, four in women). We found evidence of heterogeneity ($P_{het}<0.05$) between sexes at 13/34 NEB loci (greater than expected by chance $P_{binomial}=4\times10^{-9}$) and an overall trend for larger effect sizes in women than men (24/34, $P_{binomial}$=0.02). Two notable examples were rs58117425 in testis expressed 41 (*TEX41*) gene which was significant only in men, and 6:152202621_GT_G in the estrogen receptor alpha (*ESR1*) where the effect on NEB in women was double that in men (**Table S7**). For all NEB-associated loci we provide a summary of individual study effect estimates (**Table S8**, **Figure S1**).

In the absence of well-powered studies of infertility, we performed a GWAS on lifetime childlessness (CL) in UK Biobank (N=450,082) and assessed the relevance of NEB-associated loci on susceptibility to CL. Effects on CL were modest, with the largest effect at the rs201815280-*CADM2* locus (sex combined OR=1.05, 95% CI [1.04-1.06], $P=6.8x10^{-18}$]). Using LD score regression, the genetic correlation between NEB and CL was very high but not perfect (r_g= -0.90 [-0.88 to -0.92). Accordingly, of the 16 independent loci identified for CL, eight were distinct from the NEB signals (**Table S7)**. Sex-stratified analyses revealed one additional female-specific CL signal (rs7580304, *PPP3R1*, **Table S7**). Several loci exhibited significantly smaller or larger effects on CL than expected given their effect on NEB (**Figure S2**, **Table S7**).

In summary, we identified 43 independent signals comprised of 28 from the sex-combined NEB meta-analysis, six sex-specific NEB signals, eight additional sex-combined CL signals, and one

sex-specific CL locus. We did not identify any genome-wide significant signals on the X-chromosome, likely due to a combination of chance and a slightly smaller discovery sample size (671,349 rather than 785,604 for autosomes). We note however that the heritability of all individual chromosomes – including the X chromosome – was broadly proportional to their size (**Figure S3**), suggesting that future expanded discovery efforts are likely to also identify signals there. To validate these findings, we examined associations of these signals in 34,367 women from the FinnGen study (**Methods**). Since NEB was not recorded for men in FinnGen, we only considered the 41 signals identified in sex-combined or female-specific analyses. Despite the small replication sample, 35 of 41 loci had the same direction of effect in FinnGen as in the discovery sample (binomial sign test $P=5x10^{-6}$; **Table S7 and Figure S4**).

Previous demographic research demonstrates NEB is strongly influenced by socio-environmental factors such as education, employment uncertainty, religiosity, housing and larger trends such as economic and unemployment trends, policy measures (e.g., childcare, taxes) and contraceptive technologies[6]. None of our identified signals exhibited genome-wide significant associations with educational attainment, church attendance or social deprivation indices (all of which reported genetic associations[18]) (**Table S9**). To investigate the possibility of subtle stratification, we explored the effect of increasing the number of principal components in the GWAS model from 10 to 40 in UK Biobank and found little change in the sizes of effect estimates at the 43 significant loci (average change: -0.6%; range -8.7%% to +12.7% **Table S10**). However, since this type of stratification can never be fully controlled, we turned to alternative sources of information to identify plausible mechanisms and candidate genes in order to prioritize the association signals more likely to be driven by biology than stratification.

Implicated genes and biological mechanisms

We used a combination of *in silico* fine-mapping and summary-based Mendelian randomization (SMR) using expression quantitative trait loci (eQTL) data integration to identify putatively causal genes (**Online Methods, Table S7**). Firstly, 4 of the 43 signals were highly correlated (pairwise $r^2=0.9$) with a non-synonymous variant, implicating *ARHGAP27* (rs12949256, p.Ala117Thr), *PIK3IP1* (rs2040533, p.Thr251Ser), *ZFP82* (rs17206365, p.Leu59Met) and *LRP4* (rs6485702, p.Ile1086Val). Of note, *PIK3IP1* is a negative regulator of the PI3K/Akt/mTOR pathway, which is an intracellular signalling pathway with well-established roles in cell cycle regulation. Oocyte-specific deletion of *Pten* in mice removes the inhibiting effect of the PI3K pathway on primordial follicle activation, leading to premature recruitment and exhaustion of the entire primordial follicle pool[19].

We extended the approach of implicating genes using predicted deleterious variants by performing MAGMA[20] multi-marker gene-burden analyses restricted to the same predicted deleterious variants (**Table S11-13**). This approach identified significant genes within 3 of our identified regions (**Table S7**), notably the gene encoding Melanocortin 1 receptor (*MC1R,* $P=1.6x10^{-8}$). This was driven by 13 independent non-synonymous variants, none of which were individually genome-wide significant (**Table S14**). *MC1R* is expressed on the surface of skin and hair melanocytes and produces the pigment melanin by binding α-melanocyte stimulating hormone (α-MSH). Genetic variation at *MC1R* accounts for ~73% of the heritability of red hair colour[21], including our lead non-coding variant in this region (rs8051733, **Table S7**) and the rarer coding alleles included in the MAGMA test. The NEB effect at this locus appeared significantly stronger in women than men (**Table S7**). Three sensitivity analyses suggested that hair colour association was not responsible for the observed NEB effect (either due to population structure or mating preference). Firstly, within women in the UK Biobank, the NEB effect remained significant when red-headed women were excluded from analyses and showed consistent direction of effect within women of the same hair colour (**Table S16**). Furthermore, inclusion of hair colour in the association model reduced the effect size by only 20-25% (Table S15), suggesting that mating preference based on hair colour is unlikely to fully explain the observed effect. Secondly, there was no concordance between individual SNP effects on hair colour and NEB (**Table S14**). For example, the red hair increasing allele at Arg151Cys decreased NEB (beta=-0.02, $P=1.3x10^{-4}$), whilst the red hair increasing allele at Val92Met increased NEB (beta=0.016, $P=3.9x10^{-3}$). Finally, we assessed the impact of *MC1R* loss of function using exome sequence data in ~450,000 UK Biobank participants. The 1511 carriers of *MC1R* loss of function alleles showed no difference in NEB (P=0.65), but these loss of function alleles had a very robust effect on presence of red hair ($P=1.8x10^{-792}$), suggesting an alternative mechanism. Collectively, these data suggest a novel role of the melanocortin 1 receptor in reproduction, consistent with the recent observations that other pigmentation genes are associated with puberty timing in males[22].

MAGMA also highlighted 3 genes outside regions identified by the 43 loci; *ATHL1 (*6 variants*)*, *GLDN* (11 variants) and *RPS11 (*2 variants). The association at *RPS11* was primarily driven by a single rare variant (rs739346, p.Thr77Ser), which had a relatively large effect on NEB ('T' allele frequency=0.15%, beta=0.21, $P=6.6x10^{-8}$). This gene encodes a key component of the complex which forms the ribosome and is one of the most differentially expressed genes in the sperm of men with asthenozoospermia[23].

Next, we systematically integrated publicly available gene expression QTL data with our GWAS meta-analysis results (**Online Methods**). To guide these analyses, we first assessed the relative genome-wide enrichment of NEB-associated variants near genes expressed in 53 GTEx cell types. In sex-combined analyses, a number of neuronal cell types reached significance (**Table**

S17). This pattern of enrichment is consistent with other reproductive traits such as puberty timing[13], likely reflecting the established role of genes in the hypothalamic-pituitary-gonadal (HPG) axis regulating fertility and reproductive ageing. Sex-stratified analyses demonstrated a similar enrichment, whilst also highlighting genes expressed in the testis for men (**Table S17**). Focusing on genes expressed in the brain, gonads and blood, SMR analyses suggest that 7 of our 43 genetic variants influence NEB through expression levels of one or more nearby genes (**Table S7**). This includes *IKZF3* which is a hematopoietic-specific transcription factor involved in B-cell differentiation and proliferation, where correlated variants were recently reportedly associated with mosaic Y chromosome loss[24].

Whilst we were unable to putatively link all our significantly associated genetic variants to gene function, we note that many are in or near genes with established links to aspects of reproductive biology (**Table S7, S18**). This includes genes such as estrogen receptor 1 (*ESR1*), *ENO4* which is required for sperm motility and function as well as for male fertility in mice[25], and *WNT4* which regulates müllerian-duct formation and control of ovarian steroidogenesis. This signal in *WNT4* is the same as that previously reported for both uterine fibroids and endometriosis, with the disease risk increasing allele associated with lower NEB in women but not men (**Table S7**, $P_{all}=3.6x10^{-8}$, $P_{women}=4.5x10^{-7}$, $P_{men}=0.19$). We note that many of the gene-mapping approaches described above may identify multiple genes at individual loci, highlighting the challenges of moving from variant to gene function in complex trait genetics. Ultimately further experimental work will be required to fully elucidate which genes our identified signals implicate.

To further ascertain which loci might act directly on reproductive pathways, we integrated GWAS results from other reproductive traits (**Figure 2**). We confirm previously reported systematic correlations between NEB and Age at First Birth, and between NEB and Age at First Sex[10,11]. However, other associations are less consistent. For example, a missense allele (rs9730, $r^2=1$ with rs12949256 / p.Ala117Thr) in *ARHGAP27*, which encodes a Rho GTPase, a small family of molecules involved in axon guidance, was associated with increased NEB but shorter reproductive lifespan – later age at menarche ($P=1x10^{-11}$) and earlier menopause ($P=2x10^{-5}$) – and with earlier age at first birth ($P=5.5x10^{-8}$), lower circulating testosterone concentrations in women (both bioavailable [$P=2.1x10^{-4}$] and total [$P=2.1x10^{-3}$)], but higher testosterone concentrations in men (both bioavailable [$P=3.5x10^{-4}$] and total [$P=1.9x10^{-11}$]). Another NEB signal, rs4730673 near *MDFIC*, is correlated with the most significantly associated GWAS signal reported for same-sex sexual behaviour[26] (rs10261857; $r^2 = 0.74$). At this locus, the NEB-increasing allele was associated with lower likelihood of same-sex sexual behaviour.

Overlap between NEB and historical selection signals identifies the FADS locus

Another approach to prioritize the most likely associations is to search for variants that show evidence of natural selection – i.e., they also affected fitness in ancient populations. Effect estimates for the 34 genome-wide significant NEB loci ranged from 0.012-0.025 children per allele. The population mean NEB is ~1.8 in UK Biobank, thus an effect size of 0.02 per allele implies that a group of 25 people homozygous for an NEB-increasing allele would have, on average, 46 children, compared with 45 children for a group of 25 people without that allele. Assuming no effect on pre-reproductive mortality, these effect sizes can be directly translated to selection coefficients of 0.67-1.4% per allele, which is within the range detectable by genome-wide historical selection scans[27–29]. Accordingly, we compared our NEB/CL GWAS results with the results of scans testing selection over different timescales from ~2,000 to ~30,000 years before present[27,28,30] (**Online Methods**) and evaluated overlap using Bayesian co-localization analysis[31] (**Table S19-S20**).

The strongest overlap was observed at chr11:61.5Mb, which exhibited a posterior probability of 96% that the lead variants for ancient selection and NEB represent the same underlying signal (**Figure 3A).** This locus contains the genes *FADS1* and *FADS2,* which have been targeted by selection multiple times in human history[32–35]. In particular, the derived haplotype at this locus, which increases expression of *FADS1*, has increased from a frequency of <10% 10,000 years ago to 60-75% in present-day European populations (**Figure 3B**). While some of this increase is due to admixture, there is strong evidence of positive selection over the past few thousand years, even accounting for changes in ancestry[27,32–36]. *FADS1* and *FADS2* encode enzymes that catalyse the ω-3 and ω-6 lipid biosynthesis pathways that synthesize long chain polyunsaturated fatty acids (LC-PUFA) from short chain precursors. It has been hypothesised that selection in Europe was driven by dietary transitions, in particular the Bronze Age transition to a diet based intensively on agricultural products with relatively low LC_PUFA levels[35,36]. However, the mechanism through which this gene-environment interaction might affect fitness is unclear. Indeed, the *FADS* locus is highly pleiotropic. It is one of the strongest GWAS signals for circulating lipids[37] and blood metabolites[38], and is strongly associated with blood cell phenotypes, including erythrocyte and platelet sizes and counts (**Table S21**).

In our data, each 'C' allele of the lead NEB SNP rs108499, which tags the selected *FADS* haplotype, increased NEB by 0.0134 children, corresponding to a selection coefficient of 0.74% (0.0134 divided by mean NEB of 1.8). Consistent with this, in the "White British" subset of UK Biobank, the derived allele increased in frequency by 0.009% per-year between the 1938 and 1969 birth cohorts (**Figure 3C**), corresponding to a selection coefficient of 1.2% (95% CI -0.9-

3.2%). One caveat is that GWAS results can always be affected by stratification. To provide additional evidence, we turned to exome sequence data as an independent validation dataset. In UK Biobank exome sequence data, 242 *FADS1* loss of function carriers had lower NEB compared to non-carriers (**Table S22**, beta=-0.21, $P=5.3x10^{-3}$). The NEB-increasing allele increases *FADS1* expression, so this effect is directionally consistent.

Further support for the association comes from three lines of evidence directly connecting *FADS1* to reproductive biology. First, the NEB-increasing allele is associated with higher circulating sex-hormone binding globulin (SHBG, $P=2.3x10^{-20}$), total testosterone ($P=1.9x10^{-5}$) and estradiol concentrations ($P=4x10^{-4}$) in men, and lower bioavailable testosterone concentrations in women ($P=1.5x10^{-3}$). Second, *FADS1* is expressed in human oocytes and granulosa cells at various stages of development (**Table S23**). Finally, in mice, knockout of *Fads2*, which acts in the same pathway, leads to infertility in both sexes which can be rescued by dietary supplementation of LC-PUFA[40]. In contrast, when we assessed the dose-response relationship of all previously reported[39] HDL, LDL, total cholesterol or triglyceride associated variants on NEB using a Mendelian Randomization (MR) framework, we found no association ($P>0.05$ in all models) with or without inclusion of the *FADS* locus. This suggests that the FADS locus does not affect NEB indirectly via these phenotypes. Ultimately, while further experimental work will be required to elucidate the mechanisms linking NEB-associated variants at this locus to reproductive success; our results support the association between *FADS1* variation and NEB.

The two most significant NEB-associated genes exhibit signatures of balancing selection

The most significant NEB-associated variants in the genome, in *CADM2*, show no evidence of historical positive selection. However, *CADM2* is reported to exhibit one of the strongest genomic signals of long-term balancing selection[42]. Variants in *CADM2* are associated with a range of behavioural and reproductive traits, plausibly explained by a primary effect on risk taking propensity [9,18,43]. Variants that increase risk taking also increase NEB, with risk taking and behavioural disinhibition also linked to earlier reproductive onset[11]. *ESR1*–the gene identified with our second most significant association–also contains signals of balancing selection[44], while other NEB-associated loci with nominal evidence of balancing selection contain the genes *PTPRD* and *LINC00871* (**Table S24**). Balancing selection related to pleiotropy, time-varying or environmentally-varying selection might explain why variants with relatively large effects on NEB are able to remain segregating in the population.

Lack of contemporary selection at historical selection signals

We next tested whether there was any evidence of ongoing selection (as measured by association with NEB/CL) at regions identified by the three genome-wide historical selection scans (**Table S25**). Other than the *FADS* locus, none of the other 53 regions tested exhibited an association with NEB, suggesting that few of the strong historical selective sweeps in humans are ongoing. For example, the sweep associated with lactase persistence – one of the strongest signals of selection in any human population – is not ongoing in the European ancestry populations in this study. Other sweeps, such as those associated with skin pigmentation-decreasing loci, are likely not detected in the NEB GWAS because the selected variants are now virtually fixed among European ancestry populations.

Different methods for detecting selection in humans are sensitive to selection across very different timescales – ranging from thousands to millions of years. Our GWAS can be interpreted as a genome-wide selection scan over the shortest timescale – living generations. The limited overlap between this and historical selection scans is consistent with the limited overlap among historical scans, reflecting a highly dynamic landscape of selection. Positively selected loci either fix, or stop being selected due to changing environmental pressures and remain at intermediate frequency. Balancing selection related to changing environment, or pleiotropic effects on other components of fitness also help to maintain NEB-associated variants at intermediate frequency. The *FADS* locus is unique in the sense that the selective sweep – starting at least several thousand years ago is still ongoing.

In summary, our study identifies 38 signals that have not been previously reported for NEB and represent potential targets of ongoing natural selection. Further work should aim to parse these effects into mechanisms that directly influence reproductive biology, in contrast to those which affect behaviour or reduce fitness through premature morbidity or mortality. Finally, we note that our analysis includes only European ancestry individuals and is heavily weighted by the UK Biobank, which is not representative of the UK population[41]. It remains to be seen which of these effects are consistent across cohorts and populations.

Acknowledgements

This research was conducted using the UK Biobank Resource under application 9905. This work was supported by the Medical Research Council [Unit Programme number MC_UU_12015/2 and MC_UU_00006/2], ERC grants 615603, 835079 and 865356, ESRC ES/N011856/1, The

Leverhulme Trust, Leverhulme Centre for Demographic Science and LabEx Ecodec ANR Grant ANR-11-LABX-0047. Full study-specific and individual acknowledgements can be found in the supplementary information. The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funders. This study received ethical approval from the Department of Sociology, University of Oxford, and relevant ethical approval was obtained at the local level for the contributing datasets.

Conflicts of interest

John Perry is an employee of Adrestia Therapeutics. MMcC (Mark McCarthy) has served on advisory panels for Pfizer, NovoNordisk and Zoe Global, has received honoraria from Merck, Pfizer, Novo Nordisk and Eli Lilly, and research funding from Abbvie, Astra Zeneca, Boehringer Ingelheim, Eli Lilly, Janssen, Merck, NovoNordisk, Pfizer, Roche, Sanofi Aventis, Servier, and Takeda. As of June 2019, MMcC is an employee of Genentech, and a holder of Roche stock.

Data Availability

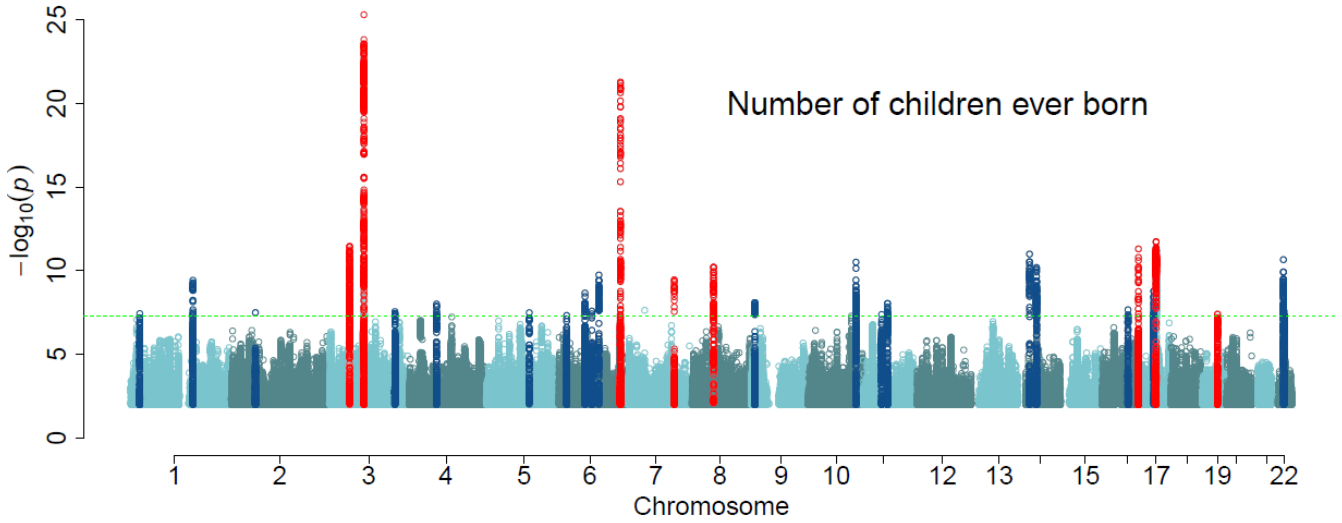
Upon publication, GWAS summary statistics will be made available at the GWAS Catalog (www.ebi.ac.uk/gwas/downloads/summary-statistics). Access to individual level data from the multiple sources used in this GWAS can be obtained by bona fide scientists through application to each specific data providers, with each data source described in the Supplementary Note.

Code availability

All analyses and modelling using standard software as described in the Methods section and in detail in the Supplementary Information.

Tables and Figures

Figure 1 |Manhattan plots for genome-wide association analyses of NEB and CL. The green line indicates the genome-wide significance threshold ($P=5x10^{-8}$). Variants that fall within 300 kb of an identified signal are highlighted: those in blue are specific to either NEB or childlessness, those in red were significant for both. Only SNPs with a p-value less than 0.01 are represented.


Number of children ever born
$-\log_{10}(p)$
0
5
10
15
20
25
1
2
3
4
5
6
7
8
9
10
12
13
15
17
19
22
Chromosome

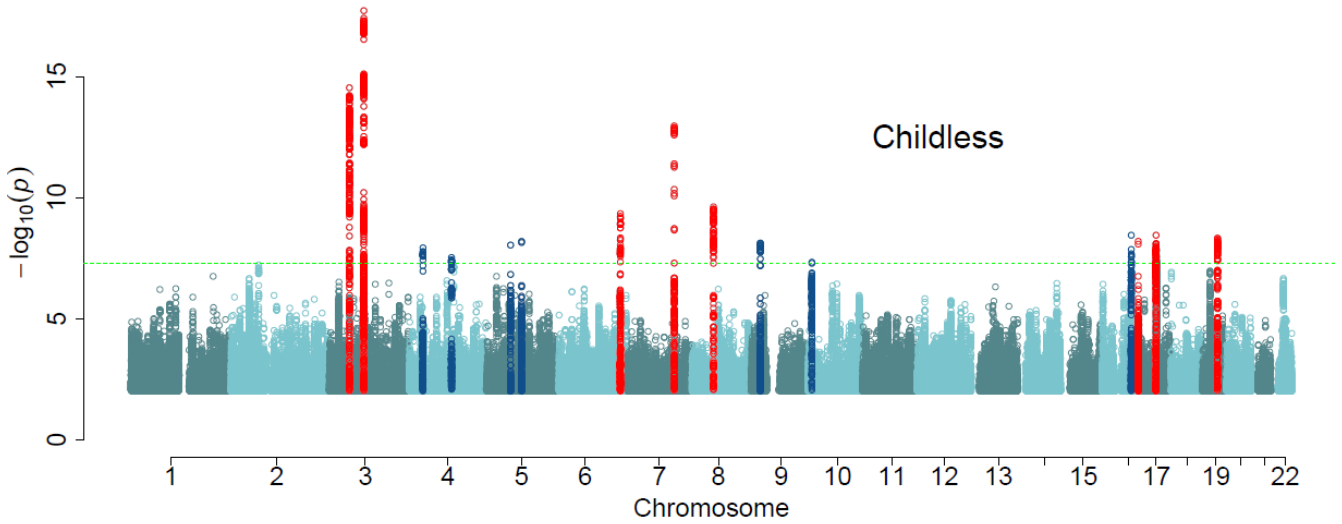

Childless
$-\log_{10}(p)$
0
5
10
15
1
2
3
4
5
6
7
8
9
10
11
12
13
15
17
19
22
Chromosome

Figure 2 |Heat map of the effects of the 43 independent signals identified for NEB or CL on other reproductive traits. All associations based on trait-specific Z-scores aligned to NEB-increasing allele, with 0 (white) denoting no association. SHBG = sex hormone binding globulin, Bioavailable T = bioavailable testosterone.

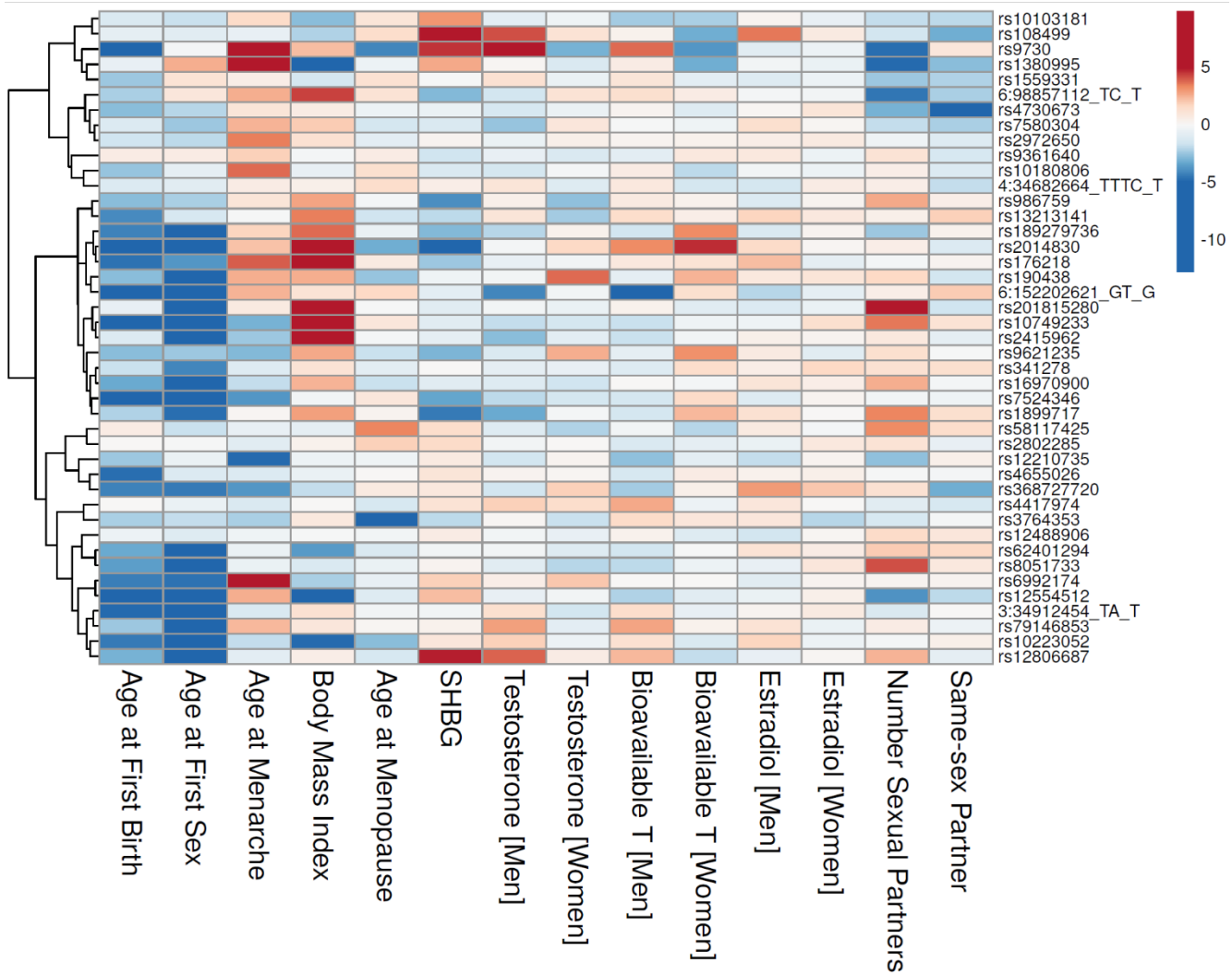
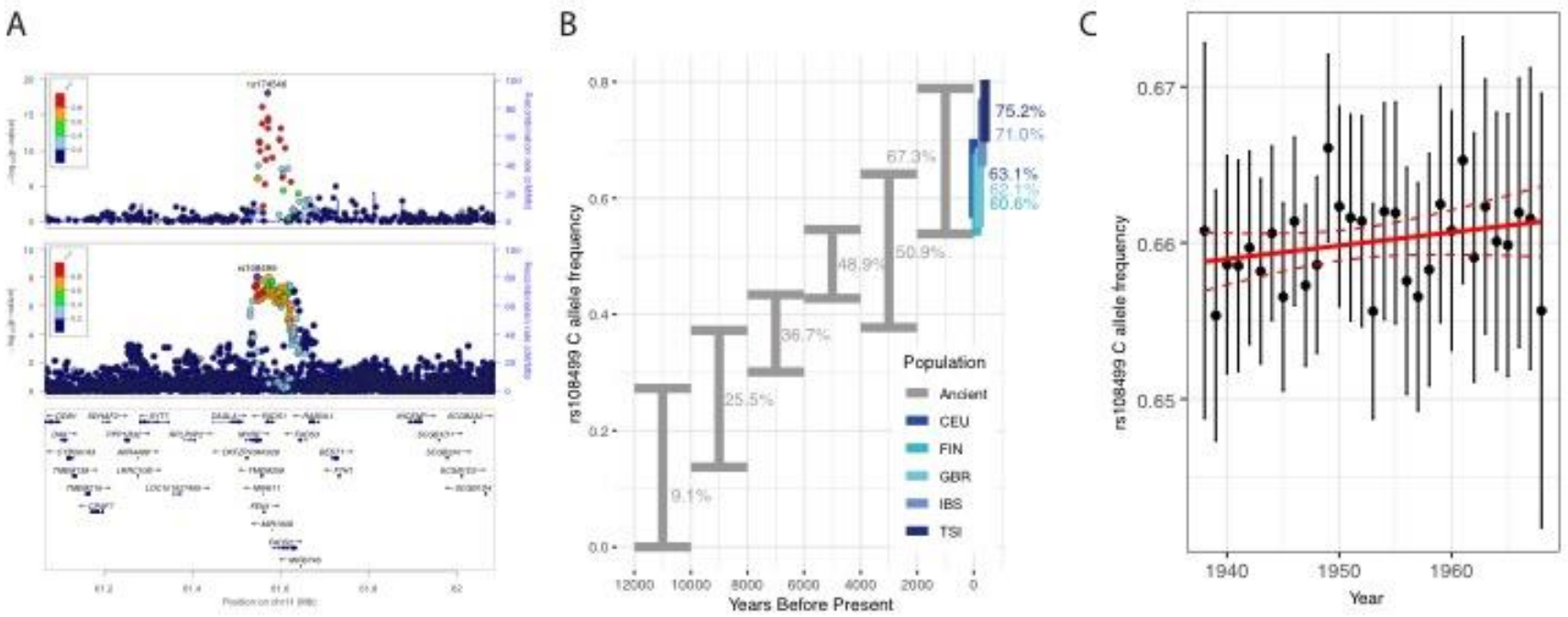

rs10103181
rs108499
rs9730
rs1380995
rs1559331
6:98857112_TC_T
rs4730673
rs7580304
rs2972650
rs9361640
rs10180806
4:34682664_TTTC_T
rs986759
rs13213141
rs189279736
rs2014830
rs176218
rs190438
6:152202621_GT_G
rs201815280
rs10749233
rs2415962
rs9621235
rs341278
rs16970900
rs7524346
rs1899717
rs58117425
rs2802285
rs12210735
rs4655026
rs368727720
rs4417974
rs3764353
rs12488906
rs62401294
rs8051733
rs6992174
rs12554512
3:34912454_TA_T
rs79146853
rs10223052
rs12806687
5
0
-5
-10
Age at First Birth
Age at First Sex
Age at Menarche
Body Mass Index
Age at Menopause
SHBG
Testosterone [Men]
Testosterone [Women]
Bioavailable T [Men]
Bioavailable T [Women]
Estradiol [Men]
Estradiol [Women]
Number Sexual Partners
Same-sex Partner

Figure 3 | Evidence for historical and ongoing selection at the FADS locus. **A**: Colocalization of the ancient DNA selection signal[27] (upper panel) and the NEB GWAS signal (lower panel). **B**: Estimated frequency (95% confidence intervals) for the derived *FADS* allele in Europe, based on direct evidence from ancient DNA. Present-day frequencies in 1000 Genomes European populations shown in blue. **C**: Frequency (95% confidence intervals) of the derived *FADS* allele in UK biobank as a function of birth year from 1938 to 1968.



References

1. Day, F. R. *et al.* Causal mechanisms and balancing selection inferred from genetic associations with polycystic ovary syndrome. *Nat. Commun.* **6**, 8464 (2015).
2. Day, F. *et al.* Large-scale genome-wide meta-analysis of polycystic ovary syndrome suggests shared genetic architecture for different diagnosis criteria. *PLoS Genet.* **14**, e1007813 (2018).
3. Censin, J. C., Bovijn, J., Holmes, M. V & Lindgren, C. M. Commentary: Mendelian randomization and women's health. *Int. J. Epidemiol.* **48**, 830–833 (2019).
4. O'Connor, L. J. *et al.* Extreme Polygenicity of Complex Traits Is Explained by Negative Selection. *Am. J. Hum. Genet.* **105**, 456–476 (2019).
5. Weissbrod, O. *et al.* Functionally informed fine-mapping and polygenic localization of complex trait heritability. *Nat. Genet.* **52**, 1355–1363 (2020).
6. Balbo, N., Billari, F. C. & Mills, M. Fertility in Advanced Societies: A Review of Research: La fécondité dans les sociétés avancées: un examen des recherches. *Eur. J. Popul.* **29**, 1–38 (2013).
7. Mills, M., 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. Tropf, F. C. *et al.* Hidden heritability due to heterogeneity across seven populations. *Nat. Hum. Behav.* **1**, 757–765 (2017).
9. Day, F. R. *et al.* Physical and neurobehavioral determinants of reproductive onset and success. *Nat. Genet.* **48**, 617–623 (2016).
10. Barban, N. *et al.* Genome-wide analysis identifies 12 loci influencing human reproductive behavior. *Nat. Genet.* **48**, 1462–1472 (2016).
11. Mills M. *et al*. Identification of 371 genetic variants for age at first sex and birth linked to externalising behaviour. *Nat. Hum. Behav.* **5**, 1717-1730 (2021)
12. Tropf, F. C. *et al.* Human fertility, molecular genetics, and natural selection in modern societies. *PLoS One* **10**, e0126821 (2015).
13. Day, F. R. *et al.* Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. *Nat. Genet.* **10**, 1–19 (2017).
14. 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**, 1294–303 (2015).
15. Kong, A. *et al.* Selection against variants in the genome associated with educational attainment. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E727–E732 (2017).
16. Beauchamp, J. P. Genetic evidence for natural selection in humans in the contemporary United States. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 7774–9 (2016).
17. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–5 (2015).
18. Day, F. R., Ong, K. K. & Perry, J. R. B. Elucidating the genetic basis of social

interaction and isolation. *Nat. Commun.* **9**, 2457 (2018).

19. Reddy, P. *et al.* Oocyte-specific deletion of Pten causes premature activation of the primordial follicle pool. *Science* **319**, 611–3 (2008).

20. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput. Biol.* **11**, (2015).

21. Morgan, M. D. *et al.* Genome-wide study of hair colour in UK Biobank explains most of the SNP heritability. *Nat. Commun.* **9**, 5271 (2018).

22. Hollis, B. *et al.* Genomic analysis of male puberty timing highlights shared genetic basis with hair colour and lifespan. *Nat. Commun.* **11**, 1536 (2020).

23. Bansal, S. K., Gupta, N., Sankhwar, S. N. & Rajender, S. Differential Genes Expression between Fertile and Infertile Spermatozoa Revealed by Transcriptome Analysis. *PLoS One* **10**, e0127007 (2015).

24. Thompson, D. J. *et al.* Genetic predisposition to mosaic Y chromosome loss in blood. *Nature* **575**, 652–657 (2019).

25. Nakamura, N. *et al.* Disruption of a spermatogenic cell-specific mouse enolase 4 (eno4) gene causes sperm structural defects and male infertility. *Biol. Reprod.* **88**, 90 (2013).

26. Ganna, A. *et al.* Large-scale GWAS reveals insights into the genetic architecture of same-sex sexual behavior. *Science* **365**, (2019).

27. Mathieson, I. *et al.* Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).

28. Field, Y. *et al.* Detection of human adaptation during the past 2000 years. *Science* **354**, 760–764 (2016).

29. Grossman, S. R. *et al.* A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science* **327**, 883–6 (2010).

30. Grossman, S. R. *et al.* Identifying recent adaptations in large-scale genomic data. *Cell* **152**, 703–13 (2013).

31. Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).

32. Fumagalli, M. *et al.* Greenlandic Inuit show genetic signatures of diet and climate adaptation. *Science* **349**, 1343–7 (2015).

33. Ameur, A. *et al.* Genetic adaptation of fatty-acid metabolism: a human-specific haplotype increasing the biosynthesis of long-chain omega-3 and omega-6 fatty acids. *Am. J. Hum. Genet.* **90**, 809–20 (2012).

34. Ye, K., Gao, F., Wang, D., Bar-Yosef, O. & Keinan, A. Dietary adaptation of FADS genes in Europe varied across time and geography. *Nat. Ecol. Evol.* **1**, 167 (2017).

35. Buckley, M. T. *et al.* Selection in Europeans on Fatty Acid Desaturases Associated with Dietary Changes. *Mol. Biol. Evol.* **34**, 1307–1318 (2017).

36. Mathieson, S. & Mathieson, I. FADS1 and the Timing of Human Adaptation to Agriculture. *Mol. Biol. Evol.* **35**, 2957–2970 (2018).

37. Teslovich, T. M. *et al.* Biological, clinical and population relevance of 95 loci for

blood lipids. *Nature* **466**, 707–13 (2010).

38. Draisma, H. H. M. *et al.* Genome-wide association study identifies novel genetic variants contributing to variation in blood metabolite levels. *Nat. Commun.* **6**, (2015).

39. Klarin, D. *et al.* Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nat. Genet.* **50**, 1514–1523 (2018).

40. Stoffel, W. *et al.* Dietary ω3-and ω6-Polyunsaturated fatty acids reconstitute fertility of Juvenile and adult Fads2-Deficient mice. *Mol. Metab.* **36**, 100974 (2020).

41. Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am. J. Epidemiol.* **186**, 1026–1034 (2017).

42. Siewert, K. M. & Voight, B. F. Detecting Long-Term Balancing Selection Using Allele Frequency Correlation. *Mol. Biol. Evol.* **34**, 2996–3005 (2017).

43. Boutwell, B. *et al.* Replication and characterization of CADM2 and MSRA genes on human behavior. *Heliyon* **3**, e00349 (2017).

44. Bitarello, B. D. *et al.* Signatures of Long-Term Balancing Selection in Human Genomes. *Genome Biol. Evol.* **10**, 939–955 (2018).

45. Wood, A. R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46**, 1173–86 (2014).

46. 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).

47. Winkler, T. W. *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nat. Protoc.* **9**, 1192–212 (2014).

48. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–9 (2006).

49. Loh, P.-R. *et al.* Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **47**, 284–290 (2015).

50. Rietveld, C. A. *et al.* GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* **340**, 1467–71 (2013).

51. Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539–42 (2016).

52. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–1 (2010).

53. Evangelou, E. & Ioannidis, J. P. A. Meta-analysis methods for genome-wide association studies and beyond. *Nat. Rev. Genet.* **14**, 379–89 (2013).

54. 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).

55. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature Genetics* vol. 44 369–375 (2012).

56. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–73 (2010).

57. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).

58. Finucane, H. K. *et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* **50**, 621–629 (2018).

59. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–60 (2015).

60. Ruth, K. S. *et al.* Using human genetics to understand the disease impacts of testosterone in men and women. *Nat. Med.* **26**, 252–258 (2020).

61. Li, L. *et al.* Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. *Cell Stem Cell* **20**, 858-873.e4 (2017).

62. Zhang, Y. *et al.* Transcriptome Landscape of Human Folliculogenesis Reveals Oocyte and Granulosa Cell Interactions. *Mol. Cell* **72**, 1021-1034.e4 (2018).

63. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

64. Chen, S. *et al.* AfterQC: automatic filtering, trimming, error removing and quality control for fastq data. *BMC Bioinformatics* **18**, 80 (2017).

65. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

66. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **11**, 1650–67 (2016).

67. Vilhjálmsson, B. J. *et al.* Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *Am. J. Hum. Genet.* **97**, 576–92 (2015).

68. 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

69. Bulik-Sullivan B. *et al*. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015)

70. Võsa U. *et al*. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* **53**, 1300-1310 (2021)

71. Backman J. D. *et al*. Exome sequencing and analysis of 454,787 UK Biobank participants. *Nature* **599**, 628–634 (2021)

Online Methods

Phenotype definitions

Number of children ever born (NEB) is treated as a continuous measure that was either asked directly or could be created from several survey questions (e.g., birth histories). A standard question in most surveys asks: *How many children have you given birth to?* Or another variant is*: How many children do you have?* In most cases it was possible to distinguish between biological, adopted or step-children and when this was possible, we refer to live born biological children. Individuals were eligible for inclusion in the analysis if they were assessed for NEB and were at least age 45 for women and age 55 for men.

Childlessness (CL) is a binary measure, derived from NEB, with 1 referring to childless and 0 if an individual had children with the same inclusion rules of biological live born children and age restrictions that applied to NEB. Detailed measures for both phenotypes per cohort are described in **Table S2**.

Participating cohorts and analysis plan

A total of 45 cohorts participated in our study (**Table S1**). **Table S2** provides an overview of cohort-specific details, including an adjusted pooled analysis of women and men in the case of family data (see below). Cohorts who agreed to participate followed an Analysis Plan posted on the Open Science Framework preregistration site https://osf.io/b4r4b/ on February 08, 2017.

Cohort-level data were quality-controlled and meta-analyzed by two separate independent centres at the University of Oxford and University of Cambridge. We follow the QC protocol of the GIANT consortium's study of human height[45] and employed the software packages QCGWAS[46] and EasyQC[47], 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. In the case where apparent errors could not be amended by stringent QC and correspondence with the local analyst of the respective cohort, cohorts were excluded from the meta-analysis. (See section below for details on cohort inclusion and errors).

For NEB the total number of individuals in the pooled meta-analysis was 785,604. Not all cohorts provided data about the X chromosome (**Table S3)** meaning that the X chromosome analysis included only included 671,349 individuals. CL analysis was restricted to UK Biobank with 450,082 for both the autosomal and X chromosomes.

Sample exclusion criteria

Individuals were eligible for inclusion if they met the following conditions:

a. They were assessed for NEB at least at age 45 for women, age 55 for men.
b. All relevant covariates (e. g. year of birth) were available for the individual.
c. They were successfully genotyped genome-wide (recommended individual genotyping rate > 95%).
d. They passed the cohort-specific standard quality controls, e.g. excluding individuals who are genetic outliers in the cohort.
e. They were of European ancestry.

Genotyping and Imputation

Table S4 provides an overview of the cohort-specific details of the genotyping platform, pre-imputation quality control filters applied to the genotype data, imputation software used, the reference panel used for imputation and the presence of X chromosome data. We asked cohorts to include all autosomal SNPs imputed from the 1000G panel (at a minimum) to allow analyses across different genotyping platforms. Cohorts with denser reference panels we asked to communicate this to our team. Cohorts were asked to provide unfiltered results since filters on imputed markers were applied at the meta-analysis stage.

Association testing models

Analysts ran linear regression models for NEB and logistic regression models for CL. Analysts were asked to include birth year of the respondent (represented by birth year minus 1900), its square and cube to control for non-linear birth cohort effects. For cohorts with family-based data, we suggested controlling for family structure or excluding relatives. Combined analyses that included both men and women included interactions of birth year and its polynomials with sex. We asked cohorts to include top principal components to control for population stratification[48] and cohort specific covariates if appropriate. Some cohorts only used birth year and not the polynomials because of multi-collinearity issues/convergence of the GWA analysis. Per-chromosome heritability estimates were calculated using restricted maximum likelihood (REML) implemented in BOLT-LMM[49]. This analysis was performed for NEB in UK Biobank, using directly genotyped variants in unrelated individuals of European ancestry.

X chromosome analysis

Analysis of X chromosome variants was performed using one of two approaches, XWAS or SNPtest, the results of which could be combined by meta-analysis. In XWAS software

(http://keinanlab.cb.bscb.cornell.edu/content/xwas) we used the *--var-het-weight* command. In SNPtest, we used the -method *newml*. Since this assumes complete X-inactivation (i.e., a hemizygous male is considered the same as a homozygous female) the effect estimates and standard errors approximate ½ of those produced by XWAS.

Quality Control: filters & diagnostic checks

We followed the quality control (QC) protocol described by the GIANT consortium's GWAS of height[45]. We used an adapted version of the software package QCGWAS[46], which allows the inclusion of structural variants, in order to standardize files across cohorts and we used EasyQC[47] to filter variants by QC criteria and to produce diagnostic graphs and statistics as described below. Where errors could not be amended by combining stringent QC with file-inspections, queries to cohorts and corrections, cohorts were excluded from the meta-analysis. See also Supplementary **Tables S5-S6** for QC results on autosomal and X chromosomes for NEB and CL. Specific individual filters were:

a) Missing data. We filtered variants where information on both reference and other allele were missing, where the estimated effect, *p*-value, standard error, expected allele frequency or number of observations were missing.

b) Implausible values. We filtered variants where *p*-values > 1 or < 0, standard errors = 0 or infinite, expected allele frequency > 1 or < 0, N < 0, call rate > 1 or < 0, an SE of the effect estimate which was approximately 40% greater than the expected SE based on MAF and standard deviation and for those with an R^2 >10% (see Winkler et al[47] for the approximation for quantitative and Rietveld et al[50] for quantitative and binary traits).

c) Quality thresholds. We filtered variants where expected allele frequency = 1 or = 0 (monomorphic variants), N < 100 to guard against spurious associations due to overfitting of the model, minor allele count <6 to guard against spurious associations with low frequency-SNPs and genotyped SNPs which were not in Hardy-Weinberg Equilibrium (HWE), with significant thresholds of 10^{-3} in case N < 1,000, 10^{-4} in case 1,000 ≤ N < 2,000, 10^{-5} in case 2,000≤N<10,000 and no filter in case N>10,000, imputed markers with imputation quality < 40% and SNPs with a call rate < 95%, if discrepancies between reported and expected *p*-value based on effect estimates and standard errors are detected (see also next section on diagnostic graphs).

d) Data harmonization. We matched the cohort-based summary statistics with a 1000 Genome reference panel phase 1 version 3 reference panel provided by Winkler et al[47]. EasyQC drops

mismatched variants which cannot be resolved such as duplicates, allele mismatches or missing or invalid alleles. Based on graphical inspections (see below), we applied cohort specific filters to drop variants with obvious deviations between expected allele frequency based on the reference panel and observed allele frequency.

Quality Control: Filtering results

a) Autosomal chromosomes

Overall, the quality of studies was good (for full results of the QC-filters described above see **Tables S5** and **S6** for autosomal SNPs). One cohort was excluded (INGI-Carlantino) due to the filter on sample size. For autosomal chromosomes and NEB, the remaining 45 cohorts provided 81 files, 39 for women only, 28 for men only and 13 pooled (from cohorts with family data). Two studies did not provide imputation quality (KORA F3, N =1.066; and KORA F4, N =1,111) and contributed only 584,866 and 496,556 SNPs respectively. For the two HPFS cohorts, results from our last discovery GWAS[10] based on HapMap 2 reference panels were recycled with number of SNPs between 2,394,353 and 2,412,487. For all other cohorts, the number of variants in the analysis range between 6,691,978 for men in LBC 1921 and 20,783,286 for women in EPIC with an average of 10,574,721. For CL, between 25,555,939 and 25,554,098 variants from the UK Biobank entered the GWAS and between 13,539,540 and 13,661,642 passed QC.

b) X chromosome

For NEB, 12 cohorts provided information on the X chromosome. Overall, we received 27 files, 12 for women, 10 for men and 5 for the pooled analysis if there were relatives in the data. On average 325,872 variants passed QC with a minimum of 191,880 in men from LBC 1921 and a maximum 991,081 for the pooled UK Biobank sample. For CL, the UK Biobank provided results for between 980,779 and 991,081 variants on the X chromosome after QC.

GWAS meta-analysis, signal selection and replication

Cohort association results (after applying the QC filters) were combined using sample-size weighted meta-analysis, implemented in METAL[52]. Sample-size weighting is based on Z-scores and can account for different phenotypic measurements among cohorts[53]. The two QC centres agreed in using sample-size weighting to allow cohorts to introduce study-specific covariates in their cohort-level analysis. For each study, 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^{-8} (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. In order to identify independent signals, distance-based clumping (using a 1Mb window) was used to identify the most significant SNPs in associated regions (termed "lead SNPs"). This was then supplemented by approximate conditional analysis implemented in GCTA[54,55], where we required additional independent signals to be genome-wide significant in both pre and post conditional models.

We meta-analysed GWAS results for NEB and CL both in sex-combined and sex-specific models. 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. **Table S8** provides the forest plots of all genome-wide significant SNPs to provide a visualization of the effect size estimates for each cohort and the summary meta-analysis in addition to the 95% confidence intervals. The genetic correlation between these two results was assessed using linkage disequilibrium score regression[69].

Replication

Replication was performed using the FinnGen study - a public-private partnership project combining genotype data from Finnish biobanks and digital health record data from Finnish health registries (https://www.finngen.fi/en). Six regional and three country-wide Finnish biobanks participate in FinnGen which also incorporates data from previously established populations and disease-based cohorts. Release 4 of Finngen includes 176,899 participants.

In this analysis we included women that participated to Finngen release 4 and were at least 45 years of age by 31st December 2017. This was the last date we had information from national registries. We also excluded women younger than 16 in 1969 (the start of the registries). Using these inclusion criteria, we included women born between 1953 and 1973 and children born

between 1969 and 2017. We also excluded women that emigrated from Finland in the study period.

To determine if a woman delivered a child we used the following codes obtained from the national inpatient registry (HILMO):

- ICD-10 codes: O80-O84
- ICD-9 code: 6440B, 6450B, 650[0-9]B-659[0-9]B
- ICD-8 codes: 650-662

When multiple codes were used within a 10-month period we counted as a single delivery. There were 37,741 women, the average (SD) number of children was 1.72 (1.32) and 20.4% of the women were childless.

For principal components analysis, FinnGen data was combined with 1000 genomes data. Related individuals (<3rd degree) were removed using King software[56]. We considered common (MAF >= 0.05) high quality variants: not in chromosome X, imputation INFO>0.95, genotype imputed posterior probability>0.95 and missingess<0.01. LD-pruned (r^2<0.1) common variants were used for computing PCA with Plink 1.92. SAIGE mixed model logistic regression (https://github.com/weizhouUMICH/SAIGE/releases/tag/0.35.8.8) was used for association analysis. Age and 10 PCs and genotyping batch were used as covariates. Each genotyping batch was included as a covariate to avoid convergence issues.

Prioritizing and characterizing putatively functional genes in GWAS highlighted regions

We utilized three distinct approaches to identify putatively functional genes at each genome-wide significant locus. First, we assessed the coding impact of any variant in LD with our 43 lead index variants. We restricted assessment to variants with r^2>0.8 and predicted moderate or high impact effects based on Variant Effect Predictor (VEP) annotations. We calculated LD using PLINK v1.9 from best guess genotypes for 1000 Genomes Phase 3/HRC imputed variants in ~10,000 unrelated UK Biobank participants of white British ancestry. Second, we used MAGMA v1.08[20] to collapse multiple predicted deleterious variants (using the same VEP categories above) into a single gene score. Finally, we integrated our genome-wide summary statistics with eQTL data using Summary data-based Mendelian Randomization (SMR)[57]. Publicly available expression datasets for ovary and testis tissues in GTEx v7, in addition to a meta-analysis of eQTL brain tissues, were downloaded from the SMR website (https://cnsgenomics.com/software/smr/#eQTLsummarydata). Whole-blood data in an eQTL meta-analysis of 31,684 samples was available from the eQTLGen consortium[70]. A Bonferroni corrected p-value threshold was used in each expression dataset individually and only associations with HEIDI P > 0.01 were considered to avoid coincidental overlap due to extended patterns of LD. This resulted in a total of 11 (SMR $P<6.6\times10^{-6}$) significant transcriptions in the brain, 12 in whole blood ($P<3.2\times10^{-6}$) and 9 in the female-specific GTEx ovary analysis (SMR $P<3.2\times10^{-5}$). We additionally performed tissue enrichment analysis using linkage-disequilibrium

(LD) score regression to specifically expressed genes (LDSC-SEG)[58]. We used three datasets available on the LDSC-SEG resource page (https://github.com/bulik/ldsc/wiki/Cell-type-specific-analyses), relating to cell and tissue-specific annotations from GTEx[59].

We characterized the phenotypic consequences of *MC1R*, *FADS1* and *FADS2* loss of function using up to 454,756 exome sequences in the UK Biobank study[71]. The exome-based association with variants in MC1R and hair colour was assessed with an interim release of 184,135 individuals. All variants were annotated using Variant Effect Predictor and we only considered those predicted to be high impact loss of function defined by VEP. Individuals carrying one or more rare (MAF <=0.1%) loss of function alleles in a given gene were grouped together. We created dummy variables based on this definition for each gene and tested for association using BOLT-LMM[49].

All lookup data for additional traits was taken from previously described UK Biobank analyses. This includes sex hormones[60], number of sexual partners and same-sex behavior[26], age at menarche and BMI[13], townsend deprivation index[18], religious group attendance[18] and years of education[18].

Assessment of FADS1-3 expression in human oocytes and granulosa cells at various stages of development

We used processed RNA-seq data of Fetal Primordial Germ Cells from two studies:

- Li *et al*[61] (Accession code: GSE86146) report data from 17 human female embryos ranging from 5-26 weeks post-fertilisation.
- Zhang *et al*[62] (Accession code GSE107746), report data from follicles at 5 different stages of development from fresh ovarian tissue from 7 adult donors, separated into oocytes and granulosa cell fractions.

We also generated novel single-cell RNA-seq data from human MII Oocytes. We performed sample QC and filtering of reads to remove low quality reads, adaptor sequences and low quality bases with trimmomatic version 0.36[63] in two steps using ILLUMINACLIP:/ /Trimmomatic-0.36/adapters/NexteraPE-PE.fa:2:30:10 (SLIDINGWINDOW:4:20 CROP:72 HEADCROP:10 MINLEN:40 followed by and extra trim of headbases with HEADCROP:10). Subsequent to filtering, we used the remaining paired reads for alignment by hisat2[64] to the human genome GeneCode v.27 release with the paired GenCode v.27 gtf file containing gene annotations using: ($HISAT2 -p 22 --dta -x .gencode.v27 -1 R1.fastq -2 R2.fastq -S sample.sam) (Pertea *et al.* 2016). The resulting sam files were sorted, indexed and transformed to bam files using samtools[65]. QC measures of aligned reads was generated using picard metrics (https://slowkow.github.io/picardmetrics) and the CollectRnaSeqMetrics tool from picard tools (http://broadinstitute.github.io/picard). We filtered the bam files for mitochondrial reads and applied Stringtie to merge and assemble reference guided transcripts for gene level quantifications of raw counts, and transcripts per million (TPM)[66]. Gene expression levels in

TPM are presented in **Table S20** as this unit allows efficient comparison of gene expression levels between samples from different studies.

Identifying overlap between NEB hits and previously-identified selection signals

We assessed overlap of our NEB signals with the results of three genome-wide selection scans. First, the Composite of Multiple Signals test[30] which combines information from different statistics to detect selection on the order of the past 50,000 years. Second, an ancient DNA based scan[27] that uses direct inference of allele frequency from ancient populations to infer selection over the past 10,000 years. Finally, the Singleton Density Score[28], which uses patterns of singleton variants to infer very recent selection – on the order of a few thousand years.

For the Composite of Multiple Signals test[30], we used the rankings of CMS_{GW} statistics to obtain an empirical P-value for each SNP. For the Singleton Density Score [28], we converted normalized SDS scores to two-tailed P-values of the standard normal distribution. Finally, for the ancient DNA based selection scan[27], we used the genomic control corrected P-values from the original scan. For each NEB hit, for each scan, we identified the SNP within 10kb, with a P-value < 10^{-6} for NEB that had the most significant selection scan signal (SNP1 and PVAL1 in **Table S19-20**). We also identified the SNP within 10kb that had the most significant selection scan signal, regardless of its NEB P-value (SNP2 and PVAL2 in **Table S19-20**). Finally, we performed a Bayesian Colocalization analysis using the "*coloc*" package[31] using all SNPs within 10kb of the lead NEB SNP. This computes posterior probabilities for the hypotheses: H0 No causal SNPs, H1 Causal SNP for selection but not NEB, H2 Causal SNP for NEB but not selection, H3 One independent causal SNP for each trait, H4 One shared causal SNP for both traits. We report the hypothesis with the maximum posterior probability (COLOC in **Table S19-20**).

We also tested for overlap with a scan for balancing selection using the NCD2 statistic[44] for the GBR population of 1000 genomes. We used the target frequency of 0.5 for these tests. For all SNPs, we report the value of the window that overlaps that SNP or, if more than one window overlaps a SNP, we report the lowest P-value of any window within 10kb. Finally, we report the lowest P-value for genes within 10kb of each SNP.

Estimating *FADS1* allele frequencies from ancient DNA

We downloaded combined data from https://reich.hms.harvard.edu/allen-ancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dna-data (v37.2 accessed 14 May 2019) and restricted to 652 samples west of 40°E, north of 35°S, more recent than 12,000 years before present and with coverage at rs108499. We binned them into 2000-year bins, and

computed estimated allele frequencies and bootstrap confidence intervals. We also include the European sub-populations from phase 3 of the 1000 Genomes Project[68].