nature human behaviour

Article

https://doi.org/10.1038/s41562-023-01796-2

More than nature and nurture, indirect genetic effects on children's academic achievement are consequences of dynastic social processes

Received: 12 May 2022

Accepted: 29 November 2023

Published online: 15 January 2024

Check for updates

Michel G. Nivard[1,2], Daniel W. Belsky[3,4], K. Paige Harden[5,6], Tina Baier[7], Ole A. Andreassen[8,9], Eivind Ystrøm[10,11], Elsje van Bergen[1,2,12] & Torkild H. Lyngstad[7]

Families transmit genes and environments across generations. When parents' genetics affect their children's environments, these two modes of inheritance can produce an 'indirect genetic effect'. Such indirect genetic effects may account for up to half of the estimated genetic variance in educational attainment. Here we tested if indirect genetic effects reflect within-nuclear-family transmission ('genetic nurture') or instead a multi-generational process of social stratification ('dynastic effects'). We analysed indirect genetic effects on children's academic achievement in their fifth to ninth years of schooling in $N = 37{,}117$ parent–offspring trios in the Norwegian Mother, Father, and Child Cohort Study (MoBa). We used pairs of genetically related families (parents were siblings, children were cousins; $N = 10{,}913$) to distinguish within-nuclear-family genetic-nurture effects from dynastic effects shared by cousins in different nuclear families. We found that indirect genetic effects on children's academic achievement cannot be explained by processes that operate exclusively within the nuclear family.

Genetically informed research designs offer strong evidence that education is transmitted across generations via the inheritance of environmental advantage. Adoption[1], twin[2], molecular genetic[3] and genome-wide association studies (GWAS)[4–6] all offer evidence that the intergenerational transmission of educational attainment (EA) occurs via both genetic and environmental (that is, social) mechanisms (Table 1). Studies of adoptees show that children resemble their adoptive (social) parents in education, despite not being genetically related to them[7]. Studies of twins reveal educational similarities within monozygotic and dizygotic pairs that are in line with a role for both the genome and the environment[8]. Molecular genetics studies have also established evidence for environmental mechanisms of intergenerational transmission of EA[9,10]. Molecular genetic studies of parent–offspring trios, of adopted parent–child dyads and of

[1]Department of Biological Psychology, Vrije Universiteit, Amsterdam, the Netherlands. [2]Amsterdam Public Health Research Institute, Amsterdam, the Netherlands. [3]Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA. [4]Robert N. Butler Columbia Aging Center, Columbia University, New York, NY, USA. [5]Department of Psychology, University of Texas at Austin, Austin, TX, USA. [6]Population Research Center, University of Texas at Austin, Austin, TX, USA. [7]Department of Sociology and Human Geography, University of Oslo, Oslo, Norway. [8]NORMENT Centre, Institute of Clinical Medicine, University of Oslo, Oslo, Norway. [9]Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway. [10]PROMENTA Research Center, Department of Psychology, University of Oslo, Oslo, Norway. [11]Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway. [12]Research Institute LEARN!, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands. e-mail: t.h.lyngstad@sosgeo.uio.no

https://doi.org/10.1038/s41562-023-01796-2

Table 1 | Evidence from genetically informed studies for environmental transmission of educational outcomes

Design	Key comparison
Twins	Are dizygotic twins more similar in their EA than can be accounted for by their genetic relatedness?
Twins + offspring	Are children of (for example, female) monozygotic twins more similar in their EA to their mother than to their aunt?
Adoptees	Do adopted offspring resemble their adoptive parents more than their biological parents in their EA?
Adoptees + siblings	Do offspring adopted into more environmentally advantaged homes have higher EA than their siblings who were not adopted away?
Adoptees + PGIs	Do adoptive parents' PGIs predict adopted children's EA? Is the association between one's own PGI and one's own EA stronger if raised by biological parents than by adoptive parents?
Parent–offspring trios + PGIs	Is the portion of the parental genotype that is not inherited by the offspring (untransmitted PGI) associated with offspring's EA?
Siblings + PGIs	Is the PGI–EA association attenuated after controlling for a family-specific effect or family's SES?

biological siblings all show that people's EA is associated with the genetic variants they did not inherit—an association that can only operate via the environment/social context[11–14].

A new wave of molecular genetic studies measures genetic correlations with EA using polygenic indices (PGIs). The PGI method uses results from GWAS to summarize information about hundreds of thousands of genetic variants associated with a target trait or behaviour into a single number for each research participant. Taken at face value, PGIs based on GWAS of EA can predict as much as 12–16% of variation in EA in independent, population-based samples[15], a level of explanatory power similar to parental education. In these studies, PGIs are measured from DNA collected from the same individuals whose education is being measured. However, PGI associations with EA reflect more than direct genetic influences on the development of characteristics that promote success in school. Studies of gene–environment correlations reveal that children's EA PGIs are correlated with environments they inherit from their parents, including the social and economic circumstances of their families and neighborhoods[14,16,17]. EA PGIs therefore measure not just a child's genetic background, but their environment as well.

EA PGIs are associated with not just educational outcomes, but a range of social and economic behaviours, including where and how far people move from home, who they have children with and how they parent[16–19]. A parent generation's EA PGIs therefore become their children's environment. In family-based PGI studies, researchers can separate out the effects of genetics that are passed on from parents to children (transmitted genotypes) from those that parents' possess, but that their children do not inherit (non-transmitted genotypes). These studies find that children's EA is influenced by PGIs based on both the transmitted and non-transmitted genotypes[11,12]. The effects of the non-transmitted genotypes reflect a process of inheritance that is mediated by the environment. These 'indirect' genetic effects will be detected in GWAS and subsequently affect downstream PGI analyses of EA[20]. Indirect genetic effects are also viewed as a means to study how parental traits affect childhood outcomes, while accounting for the direct genetic effects on offspring outcomes[21].

Evidence for indirect genetic effects on EA come from PGI studies of siblings and adoptees. In sibling studies, GWAS of educational differences between siblings estimate lower heritability as compared with GWAS of unrelated individuals[4]. PGI studies based on these GWAS find that effect sizes for PGIs based on sibling-difference GWAS are smaller than effect sizes for PGIs based on between-family GWAS[6,22]. In adoption studies, two findings stand out. First, adoptive parents' PGIs are associated with their adopted children's outcomes[23], an association that could not be mediated by direct genetic transmission. Second, PGI effect sizes are larger for children living with their biological parents than for adoptees who live with social parents to whom they are not genetically related[24]. A key difference between these two groups of children is that the adoptee's PGIs are uncoupled from environments correlated with their parents' genotypes.

While there is consistent evidence for the presence of indirect genetic effects on EA, the mechanisms that produce these effects remain unclear. Among the most evocative descriptions of indirect genetic effects on EA are 'genetic nurture' and 'dynastic effects'[11,25]. 'Genetic nurture' invokes the transmission of skills and values from parents to children within nuclear families via 'nurturing' parenting behaviours. Evidence for genetic nurture-type mechanisms comes from studies that find parenting behaviours mediate indirect-genetic-effect associations of parental genotypes with their children's educational outcomes[26]. The term 'dynastic effects' invokes the transmission of wealth and social status within family lineages across generations. Evidence for dynastic-type mechanisms comes from studies that show multi-generational socioeconomic stratification correlated to the parental genotype[22], assortative mating[27–29] and subtle population stratification[30,31].

Indirect genetic effects are defined as the association of one individual's genotype with another individual's phenotype, above and beyond that individual's own genotype. Within the regression framework, indirect genetic effects on EA are estimated by including the child's genotype in a model testing the association between their parents' genotype and the child's EA. The control for the child's genotype isolates the environmentally mediated portion of the effect of the parent's genotype. We define nurture effects as the consequence of mechanisms, operating within the nuclear family, flowing from parental actions or status that introduce a correlation between parental genotypes and child outcomes above and beyond the effect of the genotypes transmitted to the child. We negatively define 'dynastic effects' as any social or historical process that introduces a correlation between parental genotype and offspring outcomes that is not nurture; as such, it includes processes like assortative mating and population stratification. Assortative mating may contribute to indirect effects by capturing the genetic component of the phenotype with which non-transmitted alleles of the parents are correlated[32]. Population stratification occurs when differences in genotype frequency spuriously correlate with environmental differences, and this induces confounding between genotype and outcome. We choose these specific definitions of nurture and dynastic effects as in our design we can sharply distinguish nurture from other causes of indirect genetic effects, but we cannot directly differentiate between other mechanisms behind the indirect effects.

An extended-pedigree design that includes multiple families in which some of the parents are siblings makes possible a further decomposition of the effect of the parent's genotype. In the extended-pedigree design, the indirect genetic effect isolated by the control for the child's genotype can be further divided into two components: (1) the between-family indirect genetic effect, identified as the effect of the average genotype among the siblings in the parental generation, and (2) the within-family indirect genetic effect, identified as the effect of the deviation of the parent's genotype from their sibship average. Within the regression framework, this is accomplished by including the parental-sibship-average genotype as a covariate in the model. This covariate effectively captures the effects of the grandparental genotype as well as effects of any environments that are shared within the extended-family pedigree (the parent and their siblings) and correlated to genotype. This design further includes a control for subtle population stratification, as the parent and their sibling have identical ancestry. An alternate specification is to regress the child's EA on their own genotype, their parents genotype and their uncle or aunts' genotype. If the parental genotype correlates with the child's

outcome because of nurture in the nuclear family, the genotype of the child's aunt/uncle should be independent from the child's outcome conditional on the parents genotype.

Within the extended-pedigree genetic-nurture model, the within-family indirect genetic effect represents 'nurture', that is, environmentally mediated effects operating within the nuclear family environment. This could include effects mediated by parenting behaviours or direct investments by parents in their children. In parallel, the estimate of the between-family indirect genetic effect captures both indirect effects that operate via 'dynastic transmission' and those that act via nurture. This could include effects mediated by multi-generational stratification in environments. Between-family indirect genetic effects could also be a function of a bias introduced in the estimated relations between genotype and phenotype introduced by people systematically selecting mates that are similar in terms of education or related traits ('assortative mating') across multiple generations. This bias does not persist within families. Therefore assortment among spouses would introduce a between-family indirect genetic effect but not a within-family indirect genetic effect.

We conducted extended-pedigree analysis of indirect genetic effects on academic achievement in the Norwegian Mother, Father, and Child Cohort Study (MoBa), in which both children and their parents are genotyped, and that includes over 10,000 sibling pairs in the parental generation. MoBa recorded children's grade 5, 8 and 9 standardized test scores on three subjects (reading comprehension in Norwegian (for almost all children their first language), maths and English). We computed four PGIs from the most recent GWASs of EA and related phenotypes; a PGI for EA (GWAS $N > 3$ million)[15], PGIs for cognitive- (GWAS $N = 257{,}700$) and non-cognitive-skill (GWAS $N = 510{,}795$) contributions to EA[33], and a PGI based on a GWAS of EA performed within sibling pairs (GWAS $N = 128{,}777$)[4].

We structure our analysis around four models (outlined in detail in Methods). We first establish that the children's own PGI's are related to their academic achievement in this sample (model 1). Next, we establish the presence of indirect genetic effects following the standard approach of regressing children's achievement on parental PGIs while conditioning on children's own PGIs (model 2). Then, to test whether the indirect effects reflect genetic nurture or dynastic effects, we specify a model that includes the mean of one parent and sibling PGIs, the parent–sibling's deviation from that mean, and the PGI of the parent that does not have a sibling in the data (model 3). The presence of a within-family indirect effect is consistent with nurture-like processes, while its absence in the presence of a between-family indirect effect is consistent with dynastic-like processes. Finally, we consider an alternative specification where the child's achievement is regressed on their own PGI, their parents' PGI's and the PGI of a sibling of one of the parents (that is, the child's aunt's or uncle's PGI). In this final model, the logic is that the PGI of the aunt or uncle would not relate to the child's achievement through nurture within the nuclear family (model 4).

Results

Figure 1 shows results from all four models estimated on the parent–sibling sample, with numerical results available in Supplementary Tables 2–4.

Associations of children's PGIs with academic achievement

In the MoBa child cohort included in our analysis, the effect size (standardized β) for the association of the EA4 PGI with academic achievement was 0.24 ($t(10{,}287) = 30.3$, standard error (s.e.) 0.008, $P < 0.001$). For the PGIs of cognitive (Cog) and non-cognitive (Non-Cog) contributions to education, which were analysed as concurrent predictors, effect sizes were 0.26 ($t(10{,}318) = 32.9$, s.e. 0.008, $P < 0.001$) for Cog and 0.14 ($t(10{,}281) = 17.1$, s.e. 0.008, $P < 0.001$) for Non-Cog. For the PGI from the within-family GWAS of EA (WFEA), the effect size was 0.17 ($t(10{,}315) = 22.1$, s.e. 0.008, $P < 0.001$).

Indirect genetic effect estimates from parent–offspring data

In all models that include an indirect genetic effect, the direct genetic effects remained significant, but were attenuated, with standardized betas that were reduced by 15–35% compared with models that did not include indirect genetic effects (for numerical results, see Fig. 1 and Supplementary Tables 2–4). The parent–offspring model (model 2) includes PGIs for parents and their child as concurrent predictors of the child's academic achievement. In these models, the effect estimate for the parental PGIs can be interpreted as an indirect genetic effect (because directly inherited genetic influences are captured by the child's PGI). Effect sizes for indirect genetic effects were modest, but in the expected direction and statistically different from zero at the $\alpha = 0.05$ level. For the EA4 PGI, effect sizes for fathers and mothers were 0.05 (mothers: $t(10{,}271) = 5.22$, s.e. 0.01, $P < 0.001$; fathers: $t(10{,}293) = 4.72$, s.e. 0.01, $P < 0.001$); for the within-family GWAS PGI, the effect size for fathers was 0.04 ($t(10{,}283) = 4.2$, s.e. 0.02, $P < 0.001$) and for mothers was 0.05 ($t(10{,}341) = 4.7$, s.e. 0.01, $P < 0.001$). For the Cog and Non-Cog PGIs, which were tested in the same model, Cog effect sizes were 0.03 ($t(10{,}277) = 3.3$, s.e. 0.01, $P < 0.001$) for fathers and 0.04 ($t(10{,}295) = 4.0$, s.e. 0.01, $P < 0.001$) for mothers and Non-Cog effect sizes were 0.05 ($t(10{,}285) = 4.8$, s.e. 0.01, $P < 0.001$) for fathers and 0.04 ($t(10{,}267) = 4.1$, s.e. 0.01, $P < 0.001$) for mothers. The model confirms an indirect genetic effect.

Indirect genetic effects estimates from extended pedigrees

Models 3 and 4 are extended-family models. Model 3 includes PGIs for parents and their child as well as the mean PGI for parental siblings as concurrent predictors of the child's academic achievement. In model 3, the estimate for the parental PGIs can be interpreted as a within-family indirect genetic effect because directly inherited genetic influences are captured by the child's PGI and between-family indirect genetic effects are captured by the parental-sibship-mean PGIs.

Considering the PGI based on the EA4 GWAS we find a large effect of the child's PGI ($\beta = 0.184$, $t(10{,}273) = 16.2$, s.e. 0.011, $P < 0.001$) on their test score, a modest but significant ($P < 0.05$) effect for the PGI of the parent for whom no sibling is in the data ($\beta = 0.051$, $t(10{,}294) = 5.3$, s.e. 0.011), a similar effect of the mean sibling PGI for the parent and their sibling (who is an aunt/uncle to the child) ($\beta = 0.053$, $t(10{,}261) = 5.0$, s.e. 0.011, $P < 0.001$), while the deviance of their parent relative to their sibling is not significant ($\beta = 0.014$, $t(10{,}324) = 0.82$, s.e. 0.017, $P = 0.412$).

Using the PGI based on the WFEA GWAS we also find a large effect of the child's PGI ($\beta = 0.126$, $t(10{,}305) = 10.8$, s.e. 0.011, $P < 0.001$) on their test score, a modest but significant effect for the PGI of the parent for whom no sibling is in the data ($\beta = 0.036$, $t(10{,}329) = 3.6$, s.e. 0.010, $P < 0.001$), a similar effect of the mean sibling PGI for the parent and their sibling ($\beta = 0.058$, $t(10{,}294) = 5.7$, s.e. 0.010, $P < 0.001$), while the deviance of their parent relative to their sibling is not significant ($\beta = 0.020$, $t(10{,}292) = 1.14$, s.e. 0.017, $P = 0.252$).

We consider the Cog and Non-Cog PGI jointly, and observe a substantial effect of the child's Cog PGI ($\beta = 0.222$, $t(10{,}276) = 19.3$, s.e. 0.011, $P < 0.001$) and a modest effect of the child's Non-Cog PGI ($\beta = 0.089$, $t(10{,}271) = 7.71$, s.e. 0.011, $P < 0.001$). The between parent-sibling pair PGIs were significant for both the Cog PGI ($\beta = 0.042$, $t(10{,}280) = 3.87$, s.e. 0.011, $P < 0.001$) and the Non-Cog PGI ($\beta = 0.053$, $t(10{,}277) = 4.94$, s.e. 0.011, $P < 0.001$). Crucially, neither the effect of the within parent-sibling pair Cog ($\beta = -0.002$, $t(10{,}270) = -0.14$, s.e. 0.017, $P = 0.882$) nor Non-Cog ($\beta = 0.027$, $t(10{,}339) = 1.59$, s.e. 0.016, $P = 0.112$) PGI was significantly different from zero.

We performed one-tailed tests of the hypothesis that the difference between the between parent-sibling pair coefficient(s) and the within parent-sibling pair coefficient(s) is equal to or smaller than zero (cf. Supplementary Table 5). Tests reject the hypothesis for results obtained with all three PGIs: EA4 ($t = 2.16$, $P = 0.0155$, $\beta_{\text{difference}} = 0.0394$), Cog/Non-Cog ($t = 2.38$, $P = 0.0087$, $\beta_{\text{difference}} = 0.0696$) and WFEA ($t = 2.09$, $P = 0.0183$, $\beta_{\text{difference}} = 0.038$).

https://doi.org/10.1038/s41562-023-01796-2

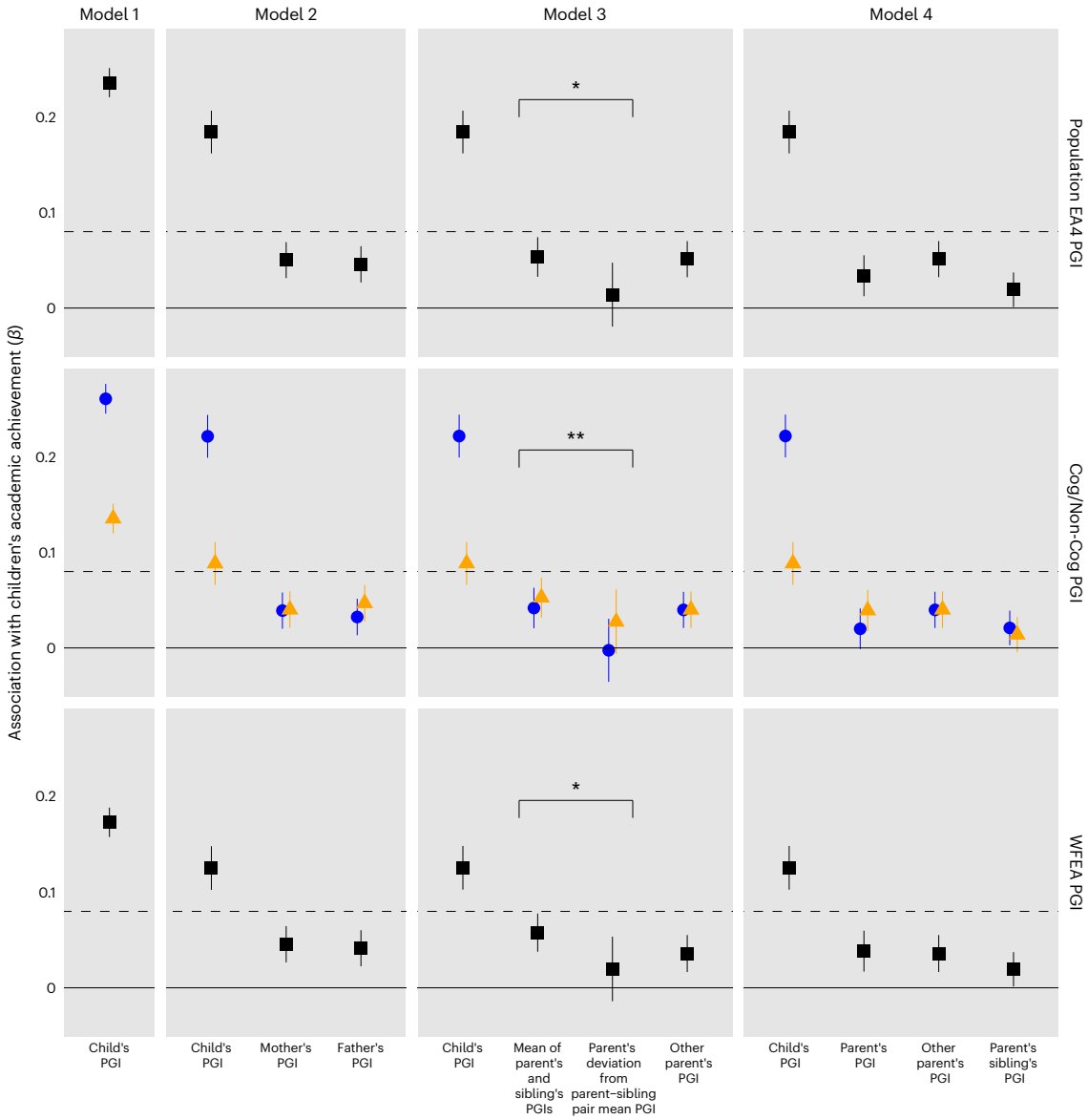

Fig. 1 | Results from four models of academic achievement using three definitions of polygenic scores. From left to right results for the models 1 and 2 (N = 37,117 families) and 3 and 4 (N = 10,913 families) where achievement is regressed on a set of PGIs and covariates (age, year, test and genomic principal components), with a child-specific random effect. Figure only shows the effect of the PGIs of interest. Top: coefficients for the EA (EA4) PGI. Middle: coefficients for cognitive skills (blue) and non-cognitive skills (orange) PGIs. Bottom: coefficients for within-family PGI for EA. Symbols represent point estimates, and vertical error bars represent 95% confidence intervals. Dashed line is a reference value for the indirect genetic effect as established in previous meta-analysis (that did not include MoBa) of educational outcomes. Brackets and stars in Model 3 panel indicate results from tests of parameter estimates for Mean of parent and sibling PGIs and Parent's deviation from parent-sibling mean PGI (* p < 0.05, ** p < 0.01, *** p < 0.0001).

The alternate specification, model 4, includes the PGIs of the child, the parent for whom no sibling is in the data, the parent who has a sibling in the data, and that parent's sibling. The results are as expected (for expectations, see Table 2). Considering the PGI based on the EA4 GWAS, we find a large direct effect of the child's PGI (β = 0.184, t(10,273) = 16.2, s.e. 0.011, P < 0.001), a modest but significant effect for the PGI of the parent for whom no sibling is in the data (β = 0.051, t(10,294) = 5.3, s.e. 0.011, P < 0.001), a smaller and significant effect for the parent for whom a sibling is available (β = 0.034, t(10,298) = 3.09, s.e. 0.011, P < 0.001) and a significant effect of the PGI of the aunt/uncle on the child (β = 0.019, t(10,315) = 2.08, s.e. 0.009, P = 0.038).

We find a large effect of the child's WFEA PGI (β = 0.126, t(10,305) = 10.7, s.e. 0.012, P < 0.001), a modest but significant effect for the PGI of the parent for whom no sibling is in the data (β = 0.036, t(10,330) = 3.63, s.e. 0.010, P < 0.001), a smaller, and significant effect for the parent for whom a sibling is available (β = 0.038, t(10,296) = 3.5, s.e. 0.011, P < 0.001), and a significant effect of the PGI of the aunt/uncle on the child (β = 0.019, t(10,286) = 2.11, s.e. 0.009, P = 0.035).

Finally, the effect of the child's Cog PGI is quite pronounced (β = 0.222, t(10,276) = 19.4, s.e. 0.011, P < 0.001), the effect of the Non-Cog PGI is modest (β = 0.088, t(10,271) = 7.71, s.e. 0.011, P < 0.001), the effect of the PGI of the parent for whom no sibling is in the data is significant and modest (Cog: β = 0.040, t(10,298) = 4.12, s.e. 0.01, P < 0.001; Non-Cog: β = 0.040, t(10,268) = 4.12, s.e. 0.01, P < 0.001), the effects of the PGIs for an aunt/uncle are significant for Cog but not Non-Cog (Cog: β = 0.021, t(10,282) = 2.27, s.e. 0.009, P = 0.023; Non-Cog: β = 0.014, t(10,333) = 1.48, s.e. 0.009, P = 0.139), while the Cog and Non-Cog effects of the parent for whom a sibling is in the data is modest (and insignificant for Cog: β = 0.020, t(10,265) = 1.84, s.e. 0.011, P = 0.066; Non-Cog: β = 0.039, t(10,307) = 3.57, s.e. 0.011, P < 0.001).

Table 2 | Expectations for indirect genetic effect parameter estimates for two model specifications under three conditions

Condition	Model specification	
	Model 3: mean of parent and sibling PGIs and parent's deviation PGI	Model 4: uncle and aunt PGI
Only genetic nurture	$\beta_{within} = \beta_{between} = \beta_{other} = \beta_{moth}$ and β_{fath}	$\beta_{focal} = \beta_{other}$ and $\beta_{u/a} = 0$
Only assortative mating	$\beta_{within} = 0$	$\beta_{focal} = \beta_{u/a} = 0.5\beta_{other}$
Only dynastic effects	$\beta_{within} = 0$	$\beta_{focal} = \beta_{u/a} = 0.5\beta_{other}$

Results from models 1 and 2 estimated on the largest possible sample, and stratified by school grade and test subject, are shown in Supplementary Fig. 1. Results from models 1–4 estimated on the sample of all parent–sibling families, and stratified by school grade and test subject, are reported in Supplementary Figs. 2–9.

Discussion

The discovery of specific genetic variants associated with EA has given researchers a new tool for investigating the intergenerational transmission of education. In particular, the observation of indirect genetic effects, whereby the parental genotype is associated with offspring outcomes beyond the child's genotype, illustrates the role of the environment in the intergenerational transmission. Our indirect genetic effect estimates are consistent with previous analysis of MoBa[34], and of Dutch[35] and UK[22] cohorts, but somewhat smaller than the average effect from a recent meta-analysis[13]. The meta-analytic average may be stronger because it included studies of educational outcomes obtained from later stages of the life course, such as adult EA.

We used an unique extended-pedigree dataset with genotyped relatives from multiple generations to study the environmental processes driving the indirect genetic effect. In contrast to the processes implied by the phrase 'genetic nurture', we did not find evidence that a large portion of the indirect genetic effects repeatedly established in previous work, predominantly runs through environmental mechanisms within the nuclear family, such as parental behaviours or investments. Instead, our findings suggest that the majority of the indirect genetic effect in academic achievement, though not necessarily all of it, does not arise within the nuclear family, but instead reflects processes shared across families with common grandparents. Specifically, after accounting for genetics shared at the extended-family level (between a child's parent and their aunt or uncle), and the child's own genetics, the 'genetic nurture' association of a parent's PGI with their child's educational achievement was not statistically different from zero. Another important source of indirect genetic effects that is consistent with our findings is a major role for assortative mating, where repeated spousal selection on characteristics that are related to educational success would introduce indirect genetic effects that in our extended-family design were fully controlled with parent-sibling pairs, as those are matched on their history of genetic assortment.

In contrast, the extended-family-level PGI did show a statistically significant association with the child's educational achievement that was comparable in size to the original indirect-genetic-effect estimate. This result does not rule out the presence of within-nuclear-family indirect genetic effects on EA. But it does suggest that genetic nurture processes unique to the nuclear family are likely to be a minor contributor to the indirect genetic effects observed in studies of trios or parent–child dyads.

We acknowledge limitations of the study. There are known biases to models that use PGIs to separate direct from indirect genetic effects. The GWAS from which we derived the per-single nucleotide polymorphism (SNP) effect estimates for the Cog, Non-Cog and EA PGIs are influenced by unmodelled indirect genetic effects. Thus, for each SNP, we rely on an effect estimate that is a mix of direct and indirect effects. This mixture can result in bias to within-family analysis of PGIs derived from these GWAS[36,37]. However, our findings persist in analysis using a PGI derived from within-family GWAS, in which the bias in SNP effects that go into the PGI, induced by gene–environment correlation, are sharply attenuated, though not entirely absent[38].

Our analysis may be under-powered to detect very small within-nuclear-family indirect genetic effects. Our analysis utilizes the largest sample for extended-pedigree indirect-genetic-effect analysis currently available. Our results are consistent across analysis of three PGIs and two different specifications. Nevertheless, there could be a non-zero within-family indirect genetic effect undetected in our analysis. Power calculations reveal that our analysis was powered to detect effects of half the size (0.04) of the indirect genetic effect on educational outcomes estimated in a recent meta-analysis[13]. Finally selection bias could affect results. MoBa has relatively high participation rates. Our use of national register data to determine educational achievement limits loss to follow-up as a source of bias. However, the MoBa sample is healthier and wealthier than the Norwegian population. Replications with other samples with the appropriate data structure (for example, the HUNT study[39]) should be a primary concern. Replication should be closely followed by generalization to other educational outcomes, such as high school completion or college enrolment. These outcomes are believed to be more strongly influenced by the nuclear family environment than are children's scores on standardized tests[35]. A recent meta-analysis of indirect genetic effects on a variety of educational outcomes including adult attainment reported an effect size (marked by dashed line in Fig. 1) stronger than those obtained in our analysis. The specific educational tests we use have limited consequences for educational careers and therefore could be considered low-stakes tests, while parents may be more invested (and seek more influence) for educational outcomes that are closer linked to the child's future social position. It would further be desirable to triangulate our result across alternate designs, for example, adoption in the parental generation or directly observed grandparental genotypes. Finally, there is a need to generalize beyond contemporary Norway, which has relatively low income inequality, a high-quality tuition-free public education system and a generous welfare state.

Our results are consistent with the interpretation of indirect genetic effects on academic achievement as in part or largely due to 'dynastic effects'. Such effects could reflect subtle socioeconomic and genetic-ancestry stratification co-occurring within homogeneous populations[4,30,31].According to this interpretation, the extended-family-level PGI is correlated with a set of inherited social circumstances that affect children's academic achievement. An alternative interpretation is that dynastic effects reflect extended-family-level behaviours and investments that contribute to children's academic achievement. Our results are further consistent with a bias in the population GWAS and PGI estimates introduced by assortative mating. Our analysis cannot isolate the precise mechanisms of indirect genetic effects on EA. However, we can conclude that, for childhood academic achievement in the context of contemporary Norway, the mechanisms that give rise to indirect genetic effects, as indexed by current PGIs, operate mostly beyond the boundaries of nuclear families.

Children from families of higher socioeconomic status (SES) perform better on standardized tests of academic achievement. As with PGI associations, correlations between children's test scores and SES might reflect 'nurture' processes occurring within the nuclear family (for example, parents actively using their resources to support their children's educational careers with higher incomes pay for private tutoring) as well as dynastic processes that persist across generations (for example, accumulation of wealth and access to social capital).

Our results suggest that, to the extent that currently available PGIs come to be correlated with child academic achievement because they are systematically associated with SES-related environmental variation, they are capturing multi-generational effects of socioeconomic privilege, rather than the more local advantages conferred by individual parents on their own offspring.

Our results do not imply that parenting behaviours or a nurturing family environment do not affect school performance. Instead, they shed light on the mechanisms behind the widely observed indirect genetic effect of parental education-related PGI on offspring education outcomes[11,13,22–24]. Any effects of parenting that are not correlated to the parental educational PGIs are not detected in studies of the indirect genetic effect. By focusing on parental PGI for education-related traits, we omit potentially important parental influences. The education PGI used here would for example not index all parental life events or circumstances that may relate to worse educational outcomes for children. While our analysis can speak to the widely studied effect of parental educational PGI on childhood academic achievement, and is well designed to avoid genetic confounding, it does not represent a comprehensive evaluation of parental influences on their children's educational outcomes.

There are strategies to leverage genetic data to study parenting without relying on education-related PGI. One follow-up would be to repeat the current analysis as a GWAS, regressing child outcomes on each SNP in the child, the same SNP in the parent and include the mean SNP of the parent and their sibling as a third covariate. A GWAS of parental effects on childhood outcomes, while using the parental sibling structure to control for confounding (that is, a within-sibling GWAS), would yield SNP level summary statistics that would allow analytical techniques like linkage disequilibrium score regression to test genetic correlations between the indirect effects and hundreds of heritable parental traits such as personality, psychopathology, wellbeing and physical health. The primary constraint on this type of analysis is sample size. However, with the continued development of national genetic databases, such extended-family GWAS of genetic nurture may soon be possible. Ultimately, a better understanding of the environmental/social intergenerational transmission of education will benefit from a tighter integration between social scientific data and genetic data.

Methods

Participants

The Norwegian Mother, Father and Child Cohort Study (MoBa) is a population-based pregnancy cohort study conducted by the Norwegian Institute of Public Health[40]. Participants were recruited from all over Norway from 1999 to 2008. Women consented to participation in 41% of the pregnancies. The cohort now includes 114,500 children, 95,200 mothers and 75,200 fathers. Not all participants have yet been genotyped, and legal restrictions related to consent reduce our effective sample size relative to some other versions of the data. The current study is based on version 12 of the quality-assured survey data files released for research in January 2019 and MoBaPsychGen v.1. The establishment of MoBa and initial data collection was based on a licence from the Norwegian Data Protection Agency and approval from The Regional Committees for Medical and Health Research Ethics. The MoBa cohort is now based on regulations related to the Norwegian Health Registry Act. The current study was considered by The Regional Committees for Medical and Health Research Ethics.

In our version of the data, 39,230 nuclear families have genotype information on complete trios (mother, father and child) where both parents were born in Norway. The Norwegian system of personal ID numbers facilitates linking of data from MoBa to register-based information for educational outcomes, basic demography and links between parents and their siblings. The data structure is illustrated in Supplementary Fig. 10. For 37,117 complete trio families, children have one or more educational outcomes available. In our analytic sample, there are 10,913 nuclear families where the child and both parents are genotyped, the child has taken at least one standardized test, and one parent has a sibling that is a genotyped parent in another MoBa family in the dataset.

Measures

Academic achievement is measured by children's results on national standardized tests ('Nasjonale prøver') in reading (that is, reading comprehension in Norwegian, for almost all children their first language), maths and English. Reading and maths were administered in 5th, 8th and 9th grades and English in 5th and 8th grades. Nearly all MoBa children have data on 5th grade tests, while the youngest cohorts do not yet have data on 8th and 9th grade tests. The test scores were obtained from Norwegian administrative registries. The scores were standardized within the test and year to control for test version and changes over time.

PGIs were computed for all individuals using the LDpred2 software. GWAS summary statistics were obtained from a GWAS-by-subtraction, for cognitive (Cog) and non-cognitive (Non-Cog) SNP effects on EA[33], and from the WFEA[15,4]. For the EA PGIs, we relied on the top 10,000 publicly reported SNPs[15].

Statistical analysis

Regression models. We fit four models. All eight test scores are included, with test fixed effects and a child-specific random intercept included in all models. The first two models establish the presence of an indirect genetic effect in the sample of all genotyped parent–offspring trios. Model 1 estimates the total genetic effect measured by the PGI.

$$\mathrm{Edu}_{ij} = \beta_{\mathrm{PGI}}\mathrm{PGI}_{\mathrm{child}\,i} + \ldots + u_i + e_{ij} \qquad \text{(model 1)}$$

Model 2 adds additional parameters for parents' PGIs and decomposes the total genetic effect into a direct component, measured by β_{dir} (for the child's PGI) and indirect components, measured by β_{fath} and β_{moth} (for the parents' PGIs).

$$\begin{aligned}\mathrm{Edu}_{ij} = {} & \beta_{\mathrm{dir}}\mathrm{PGI}_{\mathrm{child}i} + \beta_{\mathrm{fath}}\mathrm{PGI}_{\mathrm{father}i} \\ & + \beta_{\mathrm{moth}}\mathrm{PGI}_{\mathrm{mother}i} + \ldots u_i + e_{ij}\end{aligned} \qquad \text{(model 2)}$$

We then decompose the indirect genetic effect into within- and between-family components. First, we select parents with one or more siblings in the MoBa sample. We next compute the following predictors:

$$\mathrm{PGI}_{\mu} = \frac{\sum_1^m \mathrm{PGI}_m}{m},$$

where PGI_{μ} is the mean PGI for a sibship of size m;

$$\mathrm{PGI}_{\Delta} = \mathrm{PGI}_m - \mathrm{PGI}_{\mu},$$

where PGI_{Δ} is the deviation of each parent's PGI from their sibship mean PGI (PGI_{μ}), and $\mathrm{PGI}_{\mathrm{other}}$, which is the PGI of parents who do not have a sibling in the data. Finally, we combine these parameters in the equation for model 3:

$$\begin{aligned}\mathrm{Edu}_{ij} = {} & \beta_{\mathrm{dir}}\mathrm{PGI}_{\mathrm{child}i} + \beta_{\mathrm{within}}\mathrm{PGI}_{\Delta} + \beta_{\mathrm{between}}\mathrm{PGI}_{\mu} \\ & + \beta_{\mathrm{other}}\mathrm{PGI}_{\mathrm{otherparent}} \ldots + u_i + e_{ij}\end{aligned} \qquad \text{(model 3)}$$

Children for whom neither parent has a sibling in the data are omitted (N = 28,317). The model specification, which follows previous work[22], results in identical between-family and within-family effects in the absence of population stratification and/or the absence of a multi-generational effect on childhood academic achievement[41]. In the

presence of either, we expect the between parent-sibling pair effect to be larger than the within parent-sibling pair effect.

As an alternative specification to model 3, we fit a parallel model that parametrizes within- and between-family indirect genetic effects using a different approach. In this alternative specification, model 4, we include the PGIs of the parents and their siblings (that is, the uncle or aunt of the child) in the regression.

$$\mathrm{Edu}_{ij} = \beta_{\mathrm{dir}}\mathrm{PGI}_{\mathrm{child}i} + \beta_{\mathrm{focal}}\mathrm{PGI}_{\mathrm{focalparent}} + \beta_{\mathrm{uncle/aunt}}\mathrm{PGI}_{\mathrm{u/a}} + \beta_{\mathrm{other}}\mathrm{PGI}_{\mathrm{otherparent}} \ldots + u_i + e_{ij} \quad \text{(model 4)}$$

Here, an indirect genetic effect consisting of purely within-family ('genetic nurture') mechanisms would result in a parameter estimate of zero for $\beta_{\mathrm{uncle/aunt}}$. In contrast, an indirect effect consisting of only between-family ('dynastic effect') mechanisms would result in a parameter estimate of zero for β_{focal}.

All regression models include a set of child covariates indicated by the ellipsis: sex, birth year, test subject and grade fixed effects (to account for systematic differences in achievement between tests), the first ten genetic principal components (to account for population stratification) and genotyping-batch fixed effects (to account for batch-to-batch variation in genotype processing and measurement). We performed a power analysis and established we had >80% power to detect a β_{within} sibling-pair effect half that of previous meta-analytic indirect genetic effects[13]. All tests are two-tailed unless otherwise specified.

Relation between parameter and conceptual processes. In the regression models above we define three parameters that relate in the following way to underlying mechanisms that generate associations between parental PGIs and child's outcome conditional on the child's own PGI. In model 2 we define: β_{moth} and β_{fath} which are the sum of influences of genetic nurture, dynastic effects and assortative mating. In models 3 and 4 we define β_{within} which is a consequence of genetic nurture but not dynastic effects or assortment. β_{between} and β_{other} are again the sum of the effects of influences of genetic nurture, dynastic effects and assortative mating.

If β_{within} is not different from zero we find no evidence for 'genetic nurture', while if β_{within} is not different from to β_{between} and not different from the average of β_{moth} and β_{fath} (which are estimated in a larger sample, and hence with more power) this would be consistent with the absence of the influence of a 'dynastic effect' or 'assortative mating'. For convenience summarize the relations between the mechanisms that can generate PGI-phenotype associations and the regression parameters we estimate in Table 2. The relationships are confirmed through simulations available on the GitHub repository that accompanies this paper.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data analysed in the study are administrative data maintained by Statistics Norway and genotype data from MoBa Genetics. The data are not publicly available, but available to researchers upon application to the respective data owners. Such applications require approval by the appropriate ethics/research data access authorities. Access to administrative data from Statistics Norway can be applied for at Statistics Norway (http://www.ssb.no/mikrodata/) and access to MoBa Genetics can be applied for at the Norwegian Public Health Institute (http://www.fhi.no/studier/moba/). In Norway, the appropriate ethics and research data boards are the Regional Committee on Medical Research Ethics (REK) or SIKT. The consent given by the MoBa participants does not open for storage of data on an individual level in repositories or journals.

Code availability

No custom computer code was used in the study. The software used in the data preparation and analysis were R 4.0, LDpred2 and plink 1.9. R scripts for data preparation and analysis are available at http://github.com/torkildl/nurturenature.

References

1. Lundborg, P., Nordin, M. & Rooth, D. O. The intergenerational transmission of human capital: the role of skills and health. *J. Popul. Econ.* **31**, 1035–1065 (2018).
2. Silventoinen, K. et al. Genetic and environmental variation in educational attainment: an individual-based analysis of 28 twin cohorts. *Sci. Rep.* **10**, 12681 (2020).
3. Young, A. I. et al. Relatedness disequilibrium regression estimates heritability without environmental bias. *Nat. Genet.* **50**, 1304–1310 (2018).
4. Howe, L. J. et al. Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nat. Genet.* **54**, 581–592 (2022).
5. Lee, J. J. et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018).
6. Okbay, A. et al. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539–542 (2016).
7. Sacerdote, B. Nature and nurture effects on children's outcomes: what have we learned from studies of twins and adoptees? *Handb. Soc. Econ.* **1A**, 1–30 (2011).
8. Branigan, A. R., McCallum, K. J. & Freese, J. Variation in the heritability of educational attainment: an international meta-analysis. *Soc. Forces* **92**, 109–140 (2013).
9. Demange, P. A. et al. Estimating effects of parents' cognitive and non-cognitive skills on offspring education using polygenic scores. *Nat. Commun.* **13**, 2022.
10. Liu, H. Social and genetic pathways in multigenerational transmission of educational attainment. *Am. Sociol. Rev.* **83**, 278–304 (2018).
11. Kong, A. et al. The nature of nurture: effects of parental genotypes. *Science* **359**, 424–428 (2018).
12. Bates, T. C. et al. The nature of nurture: using a virtual-parent design to test parenting effects on children's educational attainment in genotyped families. *Twin Res. Hum. Genet.* **21**, 73–83 (2018).
13. Wang, B. et al. Robust genetic nurture effects on education: a systematic review and meta-analysis based on 38,654 families across 8 cohorts. *Am. J. Hum. Genet.* **108**, 1780–1791 (2021).
14. Belsky, D. W. et al. Genetic analysis of social-class mobility in five longitudinal studies. *Proc. Natl Acad. Sci. USA* **115**, E7275–E7284 (2018).
15. Okbay, A. et al. Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nat. Genet.* **54**, 437–449 (2022).
16. Belsky, D. W. et al. Genetics and the geography of health, behaviour and attainment. *Nat. Hum. Behav.* **3**, 576–586 (2019).
17. Abdellaoui, A. et al. Genetic correlates of social stratification in Great Britain. *Nat. Hum. Behav.* **3**, 1332–1342 (2019).
18. Wertz, J. et al. Genetics of nurture: a test of the hypothesis that parents' genetics predict their observed caregiving. *Dev. Psychol.* **55**, 1461–1472 (2019).
19. Belsky, D. W. et al. The genetics of success: how single-nucleotide polymorphisms associated with educational attainment relate to life-course development. *Psychol. Sci.* **27**, 957–972 (2016).
20. Morris, T. T., Davies, N. M., Hemani, G. & Smith, G. D. Population phenomena inflate genetic associations of complex social traits. *Sci. Adv.* **6**, eaay0328 (2020).

21. Koellinger, P. D. & Harden, K. P. Using nature to understand nurture. *Science* **359**, 386–387 (2018).
22. Selzam, S. et al. Comparing within- and between-family polygenic score prediction. *Am. J. Hum. Genet.* **105**, 351–363 (2019).
23. Domingue, B. W. & Fletcher, J. Separating measured genetic and environmental effects: evidence linking parental genotype and adopted child outcomes. *Behav. Genet.* **50**, 301–309 (2020).
24. Cheesman, R. et al. Comparison of adopted and nonadopted individuals reveals gene–environment interplay for education in the UK Biobank. *Psychol. Sci.* **31**, 582–591 (2020).
25. Brumpton, B. et al. Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. *Nat. Commun.* **11**, 3519 (2020).
26. Wertz, J. et al. Using DNA from mothers and children to study parental investment in children's educational attainment. *Child Dev.* **00**, cdev.13329 (2019).
27. Yengo, L. et al. Imprint of assortative mating on the human genome. *Nat. Hum. Behav.* **2**, 948–954 (2018).
28. Robinson, M. R. et al. Genetic evidence of assortative mating in humans. *Nat. Hum. Behav.* **1**, 0016 (2017).
29. Border, R. et al. Assortative mating biases marker-based heritability estimators. *Nat. Commun.* **13**, 660 (2022).
30. Zaidi, A. A. & Mathieson, I. Demographic history mediates the effect of stratification on polygenic scores. *eLife* **9**, e61548 (2020).
31. Mostafavi, H. et al. Variable prediction accuracy of polygenic scores within an ancestry group. *eLife* **9**, e48376 (2020).
32. Young, A. S. Estimation of indirect genetic effects and heritability under assortative mating. Preprint at *bioRxiv* https://doi.org/10.1101/2023.07.10.548458 (2023).
33. Demange, P. A. et al. Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. *Nat. Genet.* **53**, 35–44 (2021).
34. Isungset, M. et al. Social and genetic effects on educational performance in early adolescence. *Proc. Natl. Acad. Sci.* **119** https://doi.org/10.3386/w28498 (2021).
35. Zeeuw, E. L. de et al. Intergenerational transmission of education and ADHD: effects of parental genotypes. *Behav. Genet.* **50**, 221–232 (2020).
36. Trejo, S. & Domingue, B. W. Genetic nature or genetic nurture? Introducing social genetic parameters to quantify bias in polygenic score analyses. *Biodemography Soc. Biol.* **64**, 187–215 (2018).
37. Fletcher, J., Wu, Y., Li, T. & Lu, Q. Interpreting polygenic score effects in sibling analysis. Preprint at *bioRxiv* https://doi.org/10.1101/2021.07.16.452740 (2021).
38. Veller, C. & Coop, G. Interpreting population and family-based genome-wide association studies in the presence of confounding. Preprint at *bioRxiv* https://doi.org/10.1101/2023.02.26.530052 (2023).
39. Krokstad, S. et al. Cohort profile: the HUNT study, Norway. *Int. J. Epidemiol.* **42**, 968–977 (2013).
40. Magnus, P. et al. Cohort profile update: The Norwegian Mother and Child Cohort Study (MoBa). *Int. J. Epidemiol.* **45**, 382–388 (2016).
41. Carlin, J. B., Gurrin, L. C., Sterne, J. A., Morley, R. & Dwyer, T. Regression models for twin studies: a critical review. *Int. J. Epidemiol.* **34**, 1089–1099 (2005).

Acknowledgements

We thank the Norwegian Institute of Public Health (NIPH) for generating high-quality genomic data. This research is part of the HARVEST collaboration, supported by the Research Council of Norway (#229624). We also thank the NORMENT Centre for providing genotype data, funded by the Research Council of Norway (#223273), South East Norway Health Authority and KG Jebsen Stiftelsen. We further thank the Center for Diabetes Research, the University of Bergen for providing genotype data and performing quality control and imputation of the data funded by the ERC AdG project SELECTionPREDISPOSED, Stiftelsen Kristian Gerhard Jebsen, Trond Mohn Foundation, the Research Council of Norway, the Novo Nordisk Foundation, the University of Bergen and the Western Norway health authorities (Helse Vest) The Norwegian Mother, Father and Child Cohort Study is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research. We are grateful to all the participating families in Norway who take part in this on-going cohort study. M.G.N. is supported by ZonMW grants 849200011 and 531003014 from The Netherlands Organisation for Health Research and Development, a VENI grant awarded by NWO (VI. Veni.191G.030), and NIH grant R01MH120219. K.P.H. is supported by grant R01HD092548 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), and by NICHD grant P2CHD042849 awarded to the Population Research Center at The University of Texas at Austin. D.W.B. was supported by US National Institute on Aging grants R01AG066887, R01AG073402, Russell Sage Foundation BioSS Grant 1810-08987, and the Canadian Institute for Advanced Research. E.v.B. is supported by ZonMw grant 531003014 and NWO Gravitation grant 024.001.003. T.H.L. and T.B. are supported by Horizon2020 ERC Consolidator grant #818420 OPENFLUX. M.G.N., D.W.B., K.P.H. and E.v.B. are all past or present Jacobs Foundation Research Fellows. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

M.G.N., D.W.B., K.P.H. and T.H.L. designed the study. T.H.L. prepared data. M.G.N. and T.H.L. analyzed data. M.G.N., D.W.B., K.P.H., T.B., O.A.A., E.Y., E.v.B. and T.H.L. interpreted results. M.G.N., D.W.B., K.P.H. and T.H.L. wrote the paper. All authors provided critical comments and feedback on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-023-01796-2.

Correspondence and requests for materials should be addressed to Torkild H. Lyngstad.

Peer review information *Nature Human Behaviour* thanks Andrea Allegrini, Qiongshi Lu, Hilary C. Martin and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Torkild H. Lyngstad

Last updated by author(s): November 28th 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study.
For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☑	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☑	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☑	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☑	A description of all covariates tested
☐	☑	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☑	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☑	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☑	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☑	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☑	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No special software was used to collect data for the study
Data analysis	R 4.3, LDpred 2, plink 1.9

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The data analyzed in the study are administrative data maintained by Statistics Norway and genotype data from MoBa Genetics. The data are not publicly available, but available to researchers upon application to the respective data owners. Such applications require approval by the appropriate ethics / research data access authorities. Access to administrative data from Statistics Norway can be applied for at Statistics Norway (http://www.ssb.no/mikrodata/) and access to MoBa Genetics can be applied for at the Norwegian Public Health Institute (http://www.fhi.no/studier/moba/). In Norway, the appropriate ethics and research data boards are the Regional Committee on Medical Research Ethics (REK) or SIKT. The consent given by the MoBa participants does not open for storage of data on an individual level in repositories or journals.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Study does not report on sex and gender (identify/presentation).
Reporting on race, ethnicity, or other socially relevant groupings	Study does not report on race, ethnicity or other socially relevant groupings.
Population characteristics	Pregnant mothers, their partners and children giving birht 1999-2008
Recruitment	Pregnant mothers were recruited during routine hospital check-ups by midwives.
Ethics oversight	SIKT Personverntjenester, The Norwegian Regional Committee on Medical Research Ethics, and the University of Oslo.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☑ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	
Data exclusions	
Replication	
Randomization	
Blinding	

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	A quantitative study of genetic trios (children and their parents) and how parents' genotype relates to children's test scores
Research sample	The MoBa survey linked with population-wide administrative data on children born 2002-2008; Sample is fairly representative; Sample meets requirements for doing a analysis using polygenic scores (parents & children plus siblings of parents) and is very large.
Sampling strategy	The sample chosen was the largest sample where test scores and genotype information was available for full trios. No power analysis was done prior to first submission, but a subsequent power analysis confirmed the sample is large enough to detect relevant effect sizes.
Data collection	Genotype data were inferred from blood collected from the participants, and stored in the MoBa biobank, demographic data were collected by the Norwegian Central Population Register, and test scores collected by the Norwegian Directorate for Education.
Timing	Blood samples were collected at hospital check-ups 1999-2008. Administrative register data were continuously collected by administrative systems from 1967 to 2018, before being extracted in 2018.
Data exclusions	Missing data and individuals of non-European ancestry were excluded from the study.
Non-participation	Study only uses data already collected. No participants dropped out of the study.
Randomization	Covariates include child sex, child cohort and genomic principal components. These were included in statistical models where appropriate.

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	
Research sample	
Sampling strategy	
Data collection	
Timing and spatial scale	
Data exclusions	
Reproducibility	
Randomization	
Blinding	

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	
Location	
Access & import/export	
Disturbance	

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☑	☐ Antibodies
☑	☐ Eukaryotic cell lines
☑	☐ Palaeontology and archaeology
☑	☐ Animals and other organisms
☑	☐ Clinical data
☑	☐ Dual use research of concern
☑	☐ Plants

Methods

n/a	Involved in the study
☑	☐ ChIP-seq
☑	☐ Flow cytometry
☑	☐ MRI-based neuroimaging

Antibodies

Antibodies used	
Validation	

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	
Authentication	
Mycoplasma contamination	
Commonly misidentified lines (See ICLAC register)	

Palaeontology and Archaeology

Specimen provenance	
Specimen deposition	
Dating methods	

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight	

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	
Wild animals	
Reporting on sex	
Field-collected samples	
Ethics oversight	

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	
Study protocol	
Data collection	
Outcomes	

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☐	☐	Public health
☐	☐	National security
☐	☐	Crops and/or livestock
☐	☐	Ecosystems
☐	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☐	☐	Demonstrate how to render a vaccine ineffective
☐	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☐	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☐	☐	Increase transmissibility of a pathogen
☐	☐	Alter the host range of a pathogen
☐	☐	Enable evasion of diagnostic/detection modalities
☐	☐	Enable the weaponization of a biological agent or toxin
☐	☐	Any other potentially harmful combination of experiments and agents

Plants

Seed stocks

Novel plant genotypes

Authentication

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

Files in database submission

Genome browser session
(e.g. UCSC)

Methodology

Replicates

Sequencing depth

Antibodies

Peak calling parameters

Data quality

Software

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Instrument

Software

Cell population abundance

Gating strategy

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type

Design specifications

Behavioral performance measures

Imaging type(s)

Field strength

Sequence & imaging parameters

Area of acquisition

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software

Normalization

Normalization template

Noise and artifact removal

Volume censoring

Statistical modeling & inference

Model type and settings

Effect(s) tested

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference

(See Eklund et al. 2016)

Correction

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Graph analysis

Multivariate modeling and predictive analysis