

1 The differential effects of common and rare genetic variants on 2 cognitive performance across development

3
4 Daniel S. Malawsky^{1,*}, Mahmoud Koko¹, Petr Danacek¹, Wei Huang¹, Olivia Wootton¹, Qin Qin
5 Huang¹, Emma E. Wade¹, Sarah J. Lindsay¹, Rosalind Arden², Matthew E. Hurles¹, Hilary C.
6 Martin^{1,*}

7
8 1. Human Genetics Programme, Wellcome Sanger Institute, Hinxton, UK
9 2. Centre for Philosophy of Natural and Social Science, London School of Economics and
10 Political Science, London, UK

11
12 * Correspondence to dm22@sanger.ac.uk and hcm@sanger.ac.uk

13 Abstract

14 Common and rare genetic variants that impact adult cognitive performance also predispose to
15 rare neurodevelopmental conditions involving cognitive deficits in children. However, their
16 influence on cognition across early life remains poorly understood. Here, we investigate the
17 contribution of common genome-wide and rare exonic variation to cognitive performance across
18 childhood and adolescence primarily using the Avon Longitudinal Study of Parents and Children
19 (n=6,495 unrelated children). We show that the effect of common variants associated with
20 educational attainment and adult cognitive performance increases as children age. Conversely,
21 the negative effect of deleterious rare variants attenuates with age. Using trio analyses, we show
22 that these age-related trends are driven by direct genetic effects on the individual who carries
23 these variants rather than indirect genetic effects mediated via the family environment. We further
24 find that the increasing effects of common variants are stronger in individuals at the upper end of
25 the phenotype distribution, whereas the attenuating effects of rare variants are stronger in those
26 at the lower end. Concordant results were observed in the Millenium Cohort Study (5,920 children)
27 and UK Biobank (101,232 adults). The effects of common and rare genetic variation on childhood
28 cognitive performance are broadly comparable in magnitude to those of other factors such as
29 parental educational attainment, maternal illness and preterm birth. The effects of maternal illness
30 and preterm birth on childhood cognitive performance also attenuate with age, whereas the effect
31 of parental educational attainment does not. We show that the relative contribution of these
32 various factors differs depending on whether one considers their contribution to phenotypic
33 variance across the entire population or to the risk of poor outcomes. Our findings may help
34 explain the apparent incomplete penetrance of rare damaging variants associated with
35 neurodevelopmental conditions. More generally, they also show the importance of studying
36 dynamic genetic influences across the life course and their differential effects across the
37 phenotype distribution.

38 Main text

39 Cognitive ability is an important predictor of life outcomes such as health, education, occupation,

40 and mortality¹⁻⁶. It is influenced both by genetic and environmental factors⁷, including extrinsic
41 factors that have both genetic and environmental components such as parental educational
42 attainment (EA), gestational age at birth, and maternal illness during pregnancy⁸⁻¹⁰. Results from
43 twin studies show that the heritability of cognitive ability increases throughout development, from
44 ~0.2 in infancy to ~0.6 in adulthood¹¹. In large meta-analysis of adults, the heritability of cognitive
45 performance attributable to common single-nucleotide polymorphisms (the ‘SNP heritability’) was
46 estimated to be 0.20¹². The SNP heritability is likely to be explained both by direct genetic
47 effects^{12,13} (the effects of genetic variants in an individual on that individual’s own phenotype) and
48 by a variety of other factors. These other factors include the indirect effects of “genetic nurture”
49 (a phenomenon whereby an individual’s rearing environment is influenced by their relatives’
50 genetic make-up), parental assortment (i.e. phenotypic correlation between partners, commonly
51 known as ‘assortative mating’), and uncontrolled population stratification in genome-wide
52 association studies (GWASs)¹⁴⁻¹⁷. The relative contribution of genetic nurture, parental
53 assortment and population stratification to genetic associations is still a matter of debate, and
54 likely differs between cognitive performance and related traits such as EA and academic
55 achievement^{14,18-21}. However, the role of direct versus indirect genetic effects on cognitive and
56 academic performance across different timepoints in development is under-studied, and we
57 explore this in this work.

58
59 It has become increasingly clear that the genetics of rare neurodevelopmental conditions (NDCs)
60 involving cognitive impairment overlap with genetic factors impacting cognitive ability in the
61 general population. While Mendelian-acting rare coding variants play a large role in NDCs,
62 explaining ~50% of probands in the Deciphering Developmental Disorders (DDD) study²²⁻²⁴,
63 common variants explain ~10% of variation in risk of these conditions on the liability scale^{25,26}.
64 This common variant risk is negatively genetically correlated with EA and adult cognitive
65 performance in the general population^{25,26}. Additionally, rare damaging coding variants in NDC-
66 associated genes are associated with lower fluid intelligence and lower EA in UK Biobank²⁷⁻²⁹.
67 Estimates suggest rare coding variants exome-wide explain 1% of the variance in adult cognitive
68 performance³⁰.

69
70 Multiple lines of evidence from both clinical³¹⁻³³ and population cohorts^{29,34} suggest that some rare
71 variants conferring risk of NDCs are often shared between affected children and their seemingly
72 unaffected parents. This so-called ‘incomplete penetrance’ is seen both for large copy-number
73 variants³² and for predicted loss of function variants (pLoFs) and damaging missense variants in
74 genes in which such variants have been under negative selection throughout recent human
75 evolution (‘constrained genes’³⁵)³³. It is frequently assumed that multiple genetic and
76 environmental factors contribute to an underlying liability for NDCs, and children are diagnosed if
77 they cross some phenotypic threshold³⁶. Under this model, incomplete penetrance of rare variants
78 is expected if they increase the mean of the liability distribution in carriers to a point at which some
79 but not all cross the threshold for diagnosis. It could also occur if the variants increase the *variance*
80 in liability even without a change in mean liability of carriers, by pushing more individuals to the
81 distribution extremes. Polygenic background affects liability^{25,26} and has been hypothesized to
82 modify the penetrance of these rare variants; there is evidence of this from the general
83 population²⁹, but recent within-family analyses in NDC cohorts failed to find support for it²⁶.

84 Another cause of apparent incomplete penetrance could be time-varying effects i.e., rare variants
85 (or at least a subset of them) influences cognitive ability in childhood more than they do in
86 adulthood. We investigate this here.

87
88 In this work, we sought to dissect the contribution of genome-wide common variants and rare
89 exonic variants to cognitive performance across childhood and adolescence, using genotype data
90 and new exome-sequence data (N=16,103 children and 10,140 parents) from two British birth
91 cohorts, the Avon Longitudinal Study of Parents and Children (ALSPAC)³⁷ and the Millenium
92 Cohort Study (MCS)³⁸. Our first aim was to examine the effects of genetic measures on average
93 cognitive performance and academic achievement as children age, then test whether these
94 change across time and the extent to which they are due to direct genetic effects. Our second
95 aim was to explore whether these genetic measures have differential effects at the tails of the
96 phenotype distribution (in other words, whether they impact the variance in cognitive ability).
97 Finally, our third aim was to compare these genetic effects to those of parental educational
98 attainment and perinatal exposures, and to explore whether they are robust to controlling for these
99 other measures. Our results illuminate the dynamic genetic architecture of cognitive performance
100 across development and shed light on the factors that best predict cognitive impairment in
101 childhood as opposed to average cognitive performance in adulthood.

102 Results

103 In this work, we used measurements of cognitive performance and/or school achievement from
104 three cohorts: ALSPAC, MCS and UK Biobank. Our primary analyses were based on measures
105 of IQ from ALSPAC that were collected at ages 4 (n=1,012), 8 (n=7,347), and 16 (n=5,270). These
106 IQ tests have good psychometric properties³⁹ and are longitudinally invariant, meaning that they
107 measure the same latent construct across ages⁴⁰, making them suitable for longitudinal analyses.
108 Since phenotypic missingness can induce ascertainment biases and reduce power, we imputed
109 missing IQ values across ages by leveraging the lower missingness of other cognitive and
110 behavioral tests and demographic variables (Supplementary Note 1; Extended Data Figure 1).
111 We carried out analyses on both the observed IQ values (“pre-imputation”) and on the observed
112 and imputed IQ values combined (“post-imputation”). Results were largely concordant between
113 these unless specifically noted otherwise.

114 Influence of common variants on cognitive performance across development

115 To assess the overall contribution of common variants to cognitive development, we first
116 considered the heritabilities and genetic correlations of IQ measured at different ages in ALSPAC,
117 using 6,495 unrelated children with SNP genotype data, genetically-inferred European ancestry
118 and observed or imputed IQ values. Using GREML-LDMS⁴¹, we estimated that the heritability of
119 IQ increased between ages 4 to 16 from 0.46 to 0.56 (post-imputation), though the increase was
120 not significant (p=0.053) (Table S2). The pairwise genetic correlations between IQ measures from
121 different ages were all indistinguishable from 1 (Table S2), suggesting the common variant
122 genetic architecture is largely stable across development. These results are consistent with
123 previous studies^{42,43}.

124
125 We then considered the associations between IQ and polygenic indices (PGIs), which capture
126 the effect of common SNPs ascertained for their association with a given trait. We first considered
127 a PGI for EA (PGI_{EA}). EA is known to have a cognitive component as well as a non-cognitive
128 component which may capture factors such as personality traits and socioeconomic status that
129 affect people's ability to progress through education⁴⁴. Genetic effects on these two components
130 have been previously derived from GWASs in adults using Genomic Structural Equation
131 Modeling⁴⁴, so we also included PGIs representing each of them (PGI_{Cog} and PGI_{NonCog}). For each
132 PGI, we modeled IQ across all ages as the outcome of individual-specific random effects and
133 several fixed effects, namely age (relative to age 4), the PGI, and their interaction, with sex and
134 genetic principal components as covariates. With this mixed-effects linear model, we were able
135 to estimate two parameters of interest: the effect of the PGI at baseline i.e. at age 4 (the "main
136 effect", estimated leveraging the data at all ages) and whether the effect changed with age (the
137 "interaction effect"; see Box 1). We detected significant ($p < .05/3$) positive main effects for all PGIs,
138 as expected ("population model" in Figure 1A, Table S4). Importantly, we found that common
139 variants ascertained for their association with cognitive performance and final educational
140 attainment in adults explain more variance in IQ as children age: there were significant positive
141 PGI-by-age interaction effects for all PGIs post-imputation (Figure 1A) as well as for PGI_{EA} pre-
142 imputation (Table S4).

143
144 These associations between children's PGIs and IQ, which we term 'population effects', likely
145 reflect a combination of direct genetic effects and other correlates of the child's genotype including
146 the influence of parental genotypes on the child's IQ ("genetic nurture") and potential confounding
147 due to population stratification and parental assortment. We next sought to specifically estimate
148 the direct genetic effects of PGIs on IQ across development. We repeated the above analyses
149 adding parental PGIs as covariates in a subset of 4,968 unrelated parent-offspring trios for which
150 we measured parental genotypes or could infer them via Mendelian imputation²⁰ ("trio model";
151 see Methods). In this model, the coefficient on the child's PGI provides an estimate of the direct
152 genetic effects¹⁴, while the coefficient on the parents' PGIs represents the association between
153 the child's phenotype and alleles in the parents that are not transmitted to the child. Our results
154 for PGI_{NonCog} suggested that its population effect reflects genetic nurture or potential confounding,
155 since its direct effect was not significantly different from 0 (Figure 1A). In contrast, PGI_{EA} and
156 PGI_{Cog} showed significant direct main effects on IQ, though weaker than the population effects
157 ($p < 0.005$, z test for difference between direct and population main effects) (Figure 1A). The
158 increasing effects of PGI_{EA} and PGI_{Cog} with age that we detected in the population model appear
159 to reflect direct genetic effects, since the PGI-by-age interaction effects were significant and
160 positive in the trio model ($p < .05/3$) and statistically indistinguishable from the population effects
161 ($p > 0.05$, z test). The cognitive component of EA fully captured the direct genetic effect of PGI_{EA}
162 on IQ: we found no significant differences between the direct effects for PGI_{EA} and PGI_{Cog} ($p > 0.05$
163 for difference in the main and interaction effects, z test).

164
165 In the trio model, all parental PGIs had significant main effects, indicating that genetic nurture
166 and/or confounding effects captured by the parental PGIs are associated with childhood IQ
167 (Figure 1B, Table S4). Only the parental alleles in PGI_{NonCog} showed evidence for a change with

168 age. The effects captured by parental PGI_{NonCog} were stronger in mothers than fathers (difference
169 in main effects post-imputation = 0.067, $p=0.014$, z test), but both effects increased with age to a
170 similar degree (Figure 1B bottom right panel). These associations between children's IQ and
171 parental non-transmitted alleles associated with the non-cognitive component of EA could be
172 because parenting behaviors correlated with these alleles increase children's cognitive ability, or
173 due to cross-trait parental assortment based on the cognitive and non-cognitive components of
174 EA. The significant difference between the maternal and paternal effects may suggest that there
175 is at least some 'genetic nurture' component which differs between parents.
176

Box 1: Genetic effects estimated in this study

We estimate a variety of genetic effects from mixed-effects models (see Methods section on “Associations between genetic measures and traits”), including either main or interaction effects combined with either population effects, direct effects or effects of non-transmitted parental alleles.

Population main effect: The effect of a child’s genetic score on the phenotype at the first measurement time point (e.g., age 4 in the ALSPAC IQ analysis), but leveraging information from all time points to estimate this (see Methods). This effect is estimated without adjusting for parental genetic scores in the regression, and reflects a combination of the effects of the child’s genetic score on their own phenotype, genetic nurture, parental assortment, and uncontrolled population stratification. It does **not** have a causal interpretation.

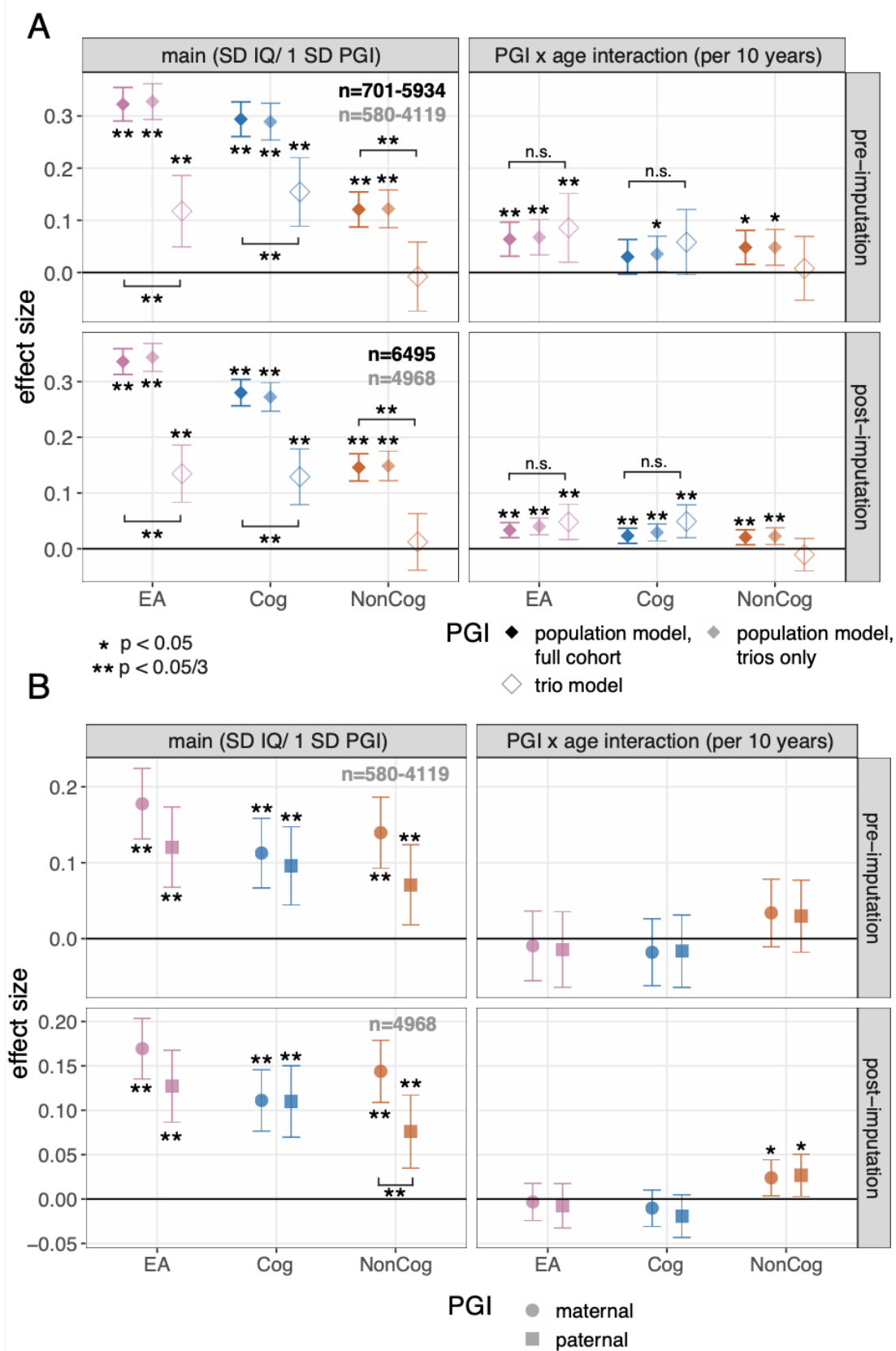
Population interaction effect: The *change* in the population effect of a genetic score on the phenotype with age. e.g., If the main effect is positive and the interaction effect is positive, this indicates that the effect of the genetic score increases with age, whereas if the main effect is negative and the interaction effect is positive, this indicates that the effect attenuates with age.

Direct main effect: The effect of a child’s genetic score on the phenotype at the first measurement time point, adjusting for the parental genetic scores (i.e., a trio model). This effect isolates the influence of the child’s own genetic score on their own phenotype and has a causal interpretation (see section IV in ⁴⁵ for further discussion).

Direct interaction effect: The *change* in the direct effect of a genetic score on the phenotype with age. The interpretation otherwise mirrors that of the population interaction effect.

Effects of non-transmitted alleles: These are the effects of the parental genetic scores on the phenotype, estimated in the trio model; in this model, the direct effect of transmitted alleles is captured by the coefficient on the child’s score, whereas the coefficients on the parents’ scores are mathematically equivalent to the effect of the non-transmitted alleles in the parents. These effects reflect a combination of genetic nurture, parental assortment, and uncontrolled population stratification. They do **not** have a causal interpretation.

Note that all of these effects are estimated both on the mean of the phenotype with linear regression (Figures 1, 2, 4A), and also on quantiles of the phenotype (e.g., the 5th percentile) with quantile regression (Figure 3, 4B).



178 **Figure 1**

179 Association between PGIs and IQ across ages. Standardized effects and 95% confidence intervals
180 estimated for the main effects and PGI-by-age-interaction effects for each PGI, either pre- (top) and post-
181 imputation (bottom) A) Results for the children's PGIs not controlling (i.e. 'population model') and controlling
182 (right) for parental PGIs (i.e. trio model). Population effect sizes are shown estimated in the full sample
183 (opaque) and in the subset of children with parental PGIs (translucent), with the corresponding sample
184 sizes shown in black and gray text respectively. B) Results for the parental PGIs from the trio model. SD:
185 standard deviation; PGI: polygenic index; EA: educational attainment; Cog: cognitive component of EA from
186 ¹⁹; NonCog: non-cognitive component of EA from ¹⁹. The square brackets indicate significant comparisons
187 highlighted in the text (z tests).

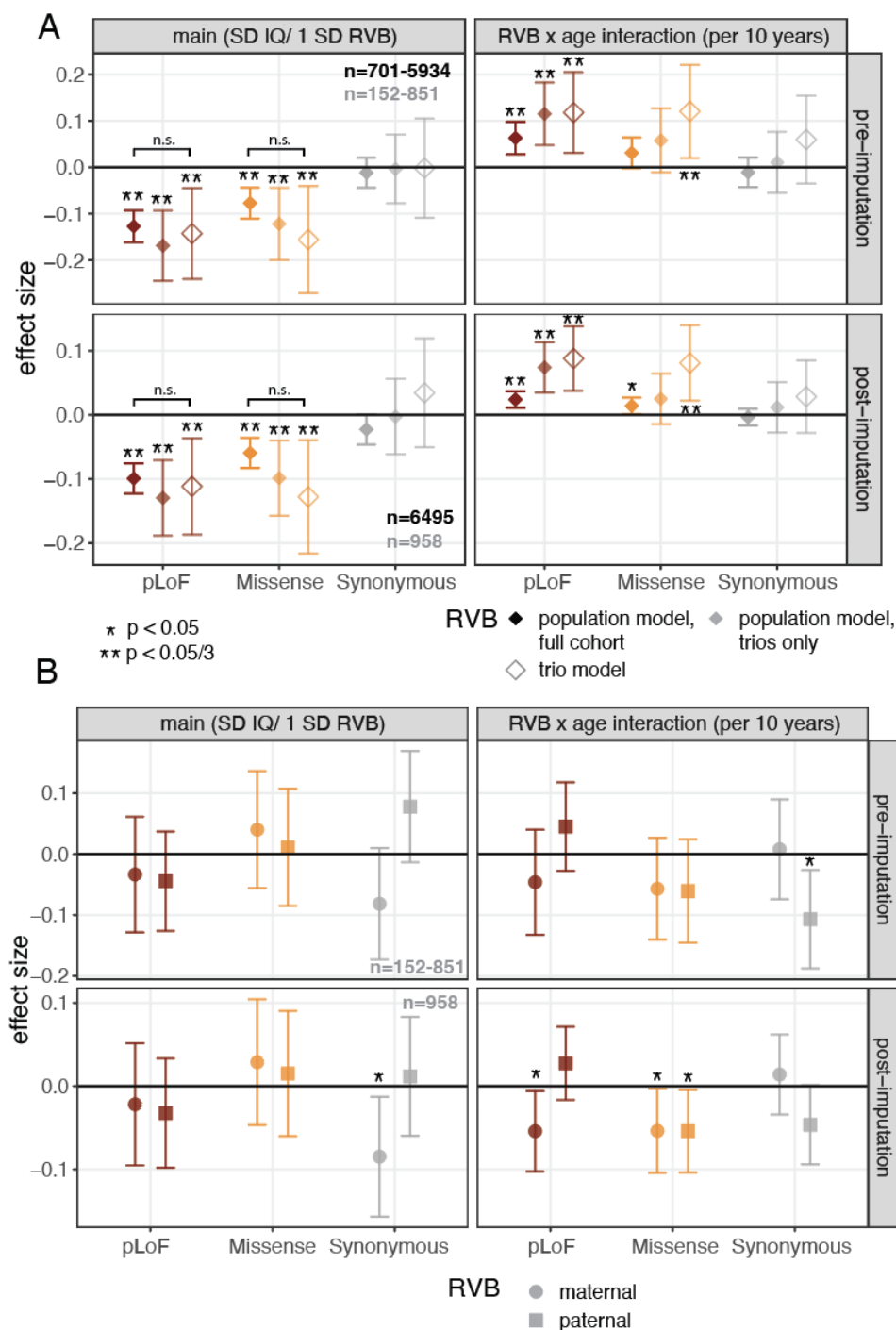
188
189 Repeating these PGI analyses using a measure of academic performance in ALSPAC gave
190 similar results (Supplementary Note 2; Extended Data Figure 2, Table S5).

191 Influence of rare variant burden on cognitive performance across development

192 In order to query the effects of rare protein-altering variants, we generated new exome
193 sequencing data on ALSPAC (8,436 children and 3,215 parents) and MCS (7,667 children and
194 6,925 parents) to supplement the SNP genotype and longitudinal phenotype data already
195 available. We first used these to assess the associations between IQ and rare variant burden
196 (RVB) in ALSPAC. We quantified RVB as the sum of gene-specific selection coefficients⁴⁶
197 (reflecting evolutionary constraint, or negative selection) for genes in which an individual carries
198 a rare predicted loss-of-function or deleterious missense variant (RVB_{pLoF} , $RVB_{Missense}$; within-
199 sample minor allele frequency, MAF <0.1%; see Methods), and, as a negative control, calculated
200 an equivalent score for synonymous variants ($RVB_{Synonymous}$).

201
202 Deleterious rare variant burden measured by RVB_{pLoF} and $RVB_{Missense}$ was significantly negatively
203 associated with IQ (main effects $p < .05/3$), whereas the burden of rare synonymous variants was
204 not (Figure 2A left, Table S6). As expected, the unstandardized effect of deleterious LoF variants
205 was stronger than that of deleterious missense variants, with the main effect of RVB_{pLoF} being
206 3.2-times larger than that of $RVB_{Missense}$ ($p = 4 \times 10^{-7}$, z test) (Extended Data Figure 3). In contrast
207 to the results seen above for common variants, the effect of deleterious rare variants attenuated
208 with age ($p < .05/3$ for RVB_{pLoF} , and $p = 0.033$ for $RVB_{Missense}$ post-imputation) (Figure 2A right).
209 Repeating these analyses using an academic performance measure in ALSPAC yielded similar
210 results (Supplementary Note 2).

211
212 Next we considered the effects of RVB_{pLoF} calculated for specific gene sets. We found that genes
213 prioritized via common variant-based GWAS of educational attainment⁴⁷ showed a stronger
214 association with IQ in childhood than expected for similarly evolutionarily constrained genes,
215 highlighting the convergence of common and rare variant associations for these related
216 phenotypes on the same genes (Supplementary Note 3, Extended Data Figure 5). We also found
217 suggestive evidence that the attenuation of the effect of RVB_{pLoF} on IQ with age was strongest in
218 genes that showed preferential expression in the prenatal rather than postnatal brain ⁴⁸
219 (Supplementary Note 3, Extended Data Figure 5A), although this was less clear when examining
220 academic performance (Figure S5).



221
 222 **Figure 2**
 223 Association between rare variant burden (RVB) and IQ across ages. Standardized effects and 95%
 224 confidence intervals estimated for the main effects and RVB-by-age-interaction effects for RVB calculated
 225 with three different consequence classes (pLoF, missense or synonymous), either pre- (top) and post-
 226 imputation (bottom). A) Results for the children's RVBs not controlling (i.e. 'population model') and
 227 controlling (right) for parental RVBs (i.e. trio model). Population effect sizes are shown estimated in the full
 228 sample (opaque) and in the subset of children with parental RVBs (translucent), with the corresponding

229 sample sizes shown in black and gray text respectively. B) Results for the parental RVBs from the trio
230 model. The square brackets indicate comparisons highlighted in the text (z tests).

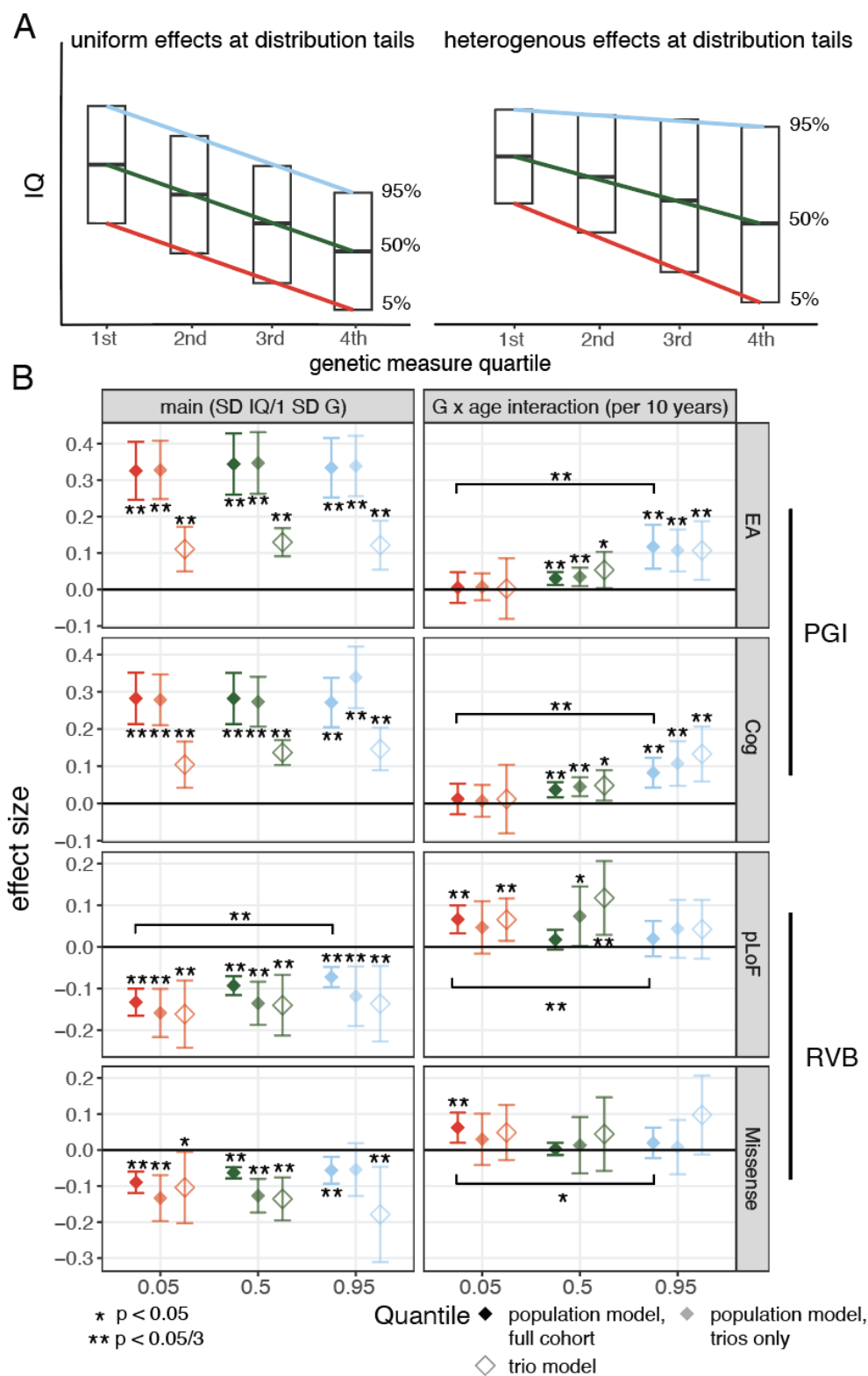
231
232 We then sought to estimate the direct genetic effects of deleterious rare variants on IQ across
233 development in a subset of 958 unrelated exome-sequenced trios by adding parental RVBs as
234 covariates. Our findings suggested that the population effects are dominated by the direct effects
235 of rare variants, in contrast to what we observed for PGIs: the estimated population and direct
236 main effects were not significantly different from each other for either RVB_{pLoF} nor $RVB_{Missense}$
237 ($p > 0.05$, z test) (Figure 2A, Table S6) and no parental RVB measure was significantly associated
238 with IQ after accounting for multiple testing (Figure 2B, Table S6), though this may be in part due
239 to the smaller sample size. We also found that the direct effects of these rare variants significantly
240 attenuated with age ($p < .05/3$ for RVB_{pLoF} -by-age and $RVB_{Missense}$ -by-age interactions in the trio
241 model). In summary, our results show that higher exome-wide burden of rare damaging variants
242 is associated with lower IQ in childhood, though with attenuated effects at later age, and that this
243 pattern is recapitulated when examining the direct genetic effects. When partitioning the effect of
244 the burden of rare pLoFs in constrained genes into inherited variants versus *de novo* mutations,
245 we estimated that the former explained 3.5-4 times more variance in IQ than the latter in ALSPAC
246 (Supplementary Note 4; Extended Data Figure 6), with similar results when considering a
247 composite measure of childhood cognitive performance in MCS (see Methods; Figure S1).

248 Influence of genetic factors on the tails of the IQ distribution

249 Our results thus far consider genetic effects on mean IQ. However, we have yet to consider
250 whether they differentially impact IQ at different quantiles in the IQ distribution. Importantly,
251 deleterious rare variants in constrained genes are known to affect risk of neurodevelopmental
252 conditions²², hinting that they may possibly have larger effects at the lower tail end of the IQ
253 distribution. This would reflect a statistical phenomenon known as “heteroscedasticity”, in which,
254 in a linear regression of phenotype on genotype, the variance of the residuals varies by genotype.
255 To test whether there are heterogenous genetic effects at the tails of the IQ distribution, we used
256 quantile regression to estimate the influence of PGIs and RVB on both the median IQ as well as
257 the bottom and top 5th percentiles (Figure 3A), which roughly correspond to IQs of 75 and 125,
258 respectively³⁹.

259
260 First, we considered the influence of PGL_{EA} and PGL_{Cog} , which had significant direct genetic effects
261 on IQ in Figure 1. Broadly, our results suggested that the observed increase of these PGI effects
262 on IQ with age was driven by increasing genetic effects in the upper half of the IQ distribution.
263 Although the effects of the PGIs on these different quantiles in the IQ distribution were largely
264 concordant at any given time point (Figure S2, Table S8), we found significant positive age
265 interaction effects for both PGIs at the median and 95th quantile for population effects ($p < .05/3$),
266 which were at least nominally significant when considering the direct effects (Figure 3B). Notably,
267 there was no evidence for significant age interactions at the 5th percentile for either PGI.
268 Repeating this analysis using academic performance in ALSPAC replicated the results, showing
269 that the effect of PGL_{EA} only significantly increased at the median and nominally at the 95th
270 percentile (Extended Data Figure 7A). We observed concordant results when using cognitive
271 ability measures from MCS and UK Biobank (Supplementary Note 5; Extended Data Figure 7B).

272 In ALSPAC, parental PGI associations with IQ were largely uniform across the phenotype
 273 distribution and consistent across development (Extended Data Figure 8).
 274



275

276 **Figure 3**

277 Influence of common and rare variants on the tails of the IQ distribution, using IQ measures post-imputation
278 in ALSPAC. A) Schematic showing two scenarios where a genetic measure is associated with IQ and has
279 uniform effects across the IQ distribution (left) versus a scenario in which the genetic measure has
280 heterogeneous effects on IQ across the IQ distribution (right). B) Standardized effects and 95% confidence
281 intervals for quantile regression of the 5th, 50th, and 95th percentile estimated from mixed-effects modeling
282 with post-imputation IQ at ages 4, 8, and 16 for EA and EA-cog PGIs before (left) and after (right) controlling
283 for parental genetic measures. C) Same as (B) using RVB_{pLoF} and $RVB_{Missense}$ as genetic measures. See
284 Figure S3 for results from this mixed-effects model pre-imputation, and Figure S2 and Figure S4 for cross-
285 sectional estimates at each age. The square brackets indicate significant comparisons highlighted in the
286 text (z tests).

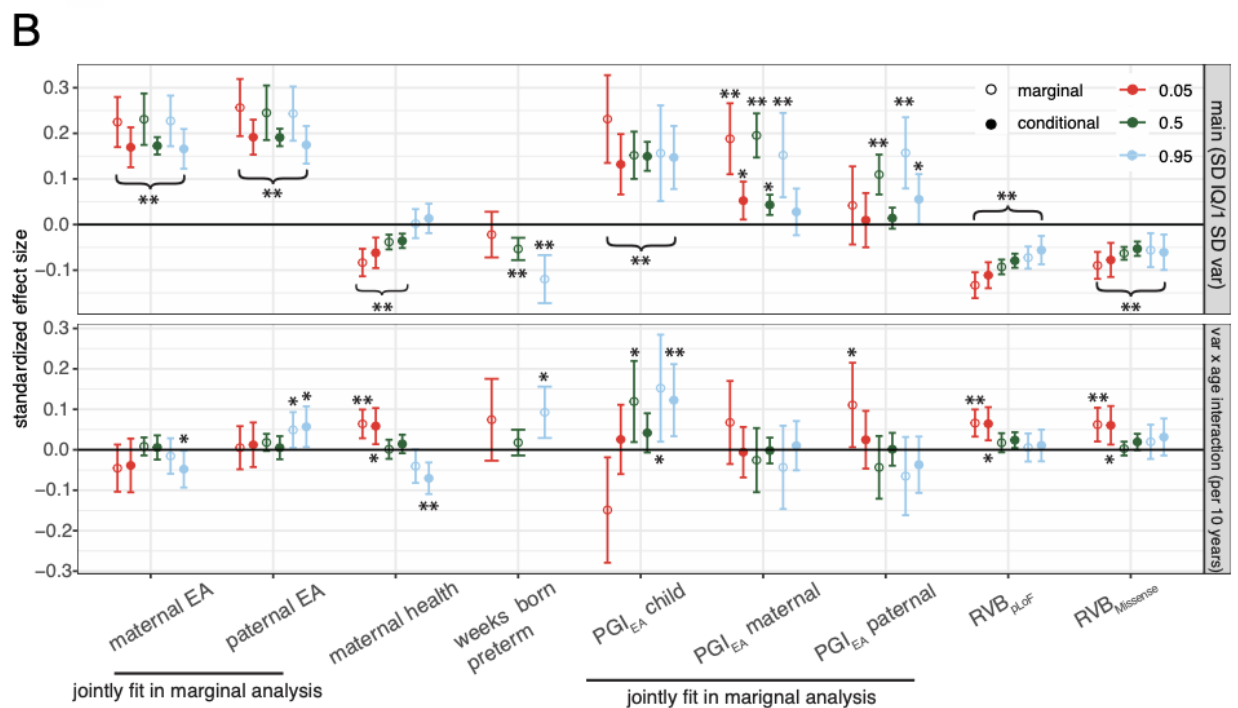
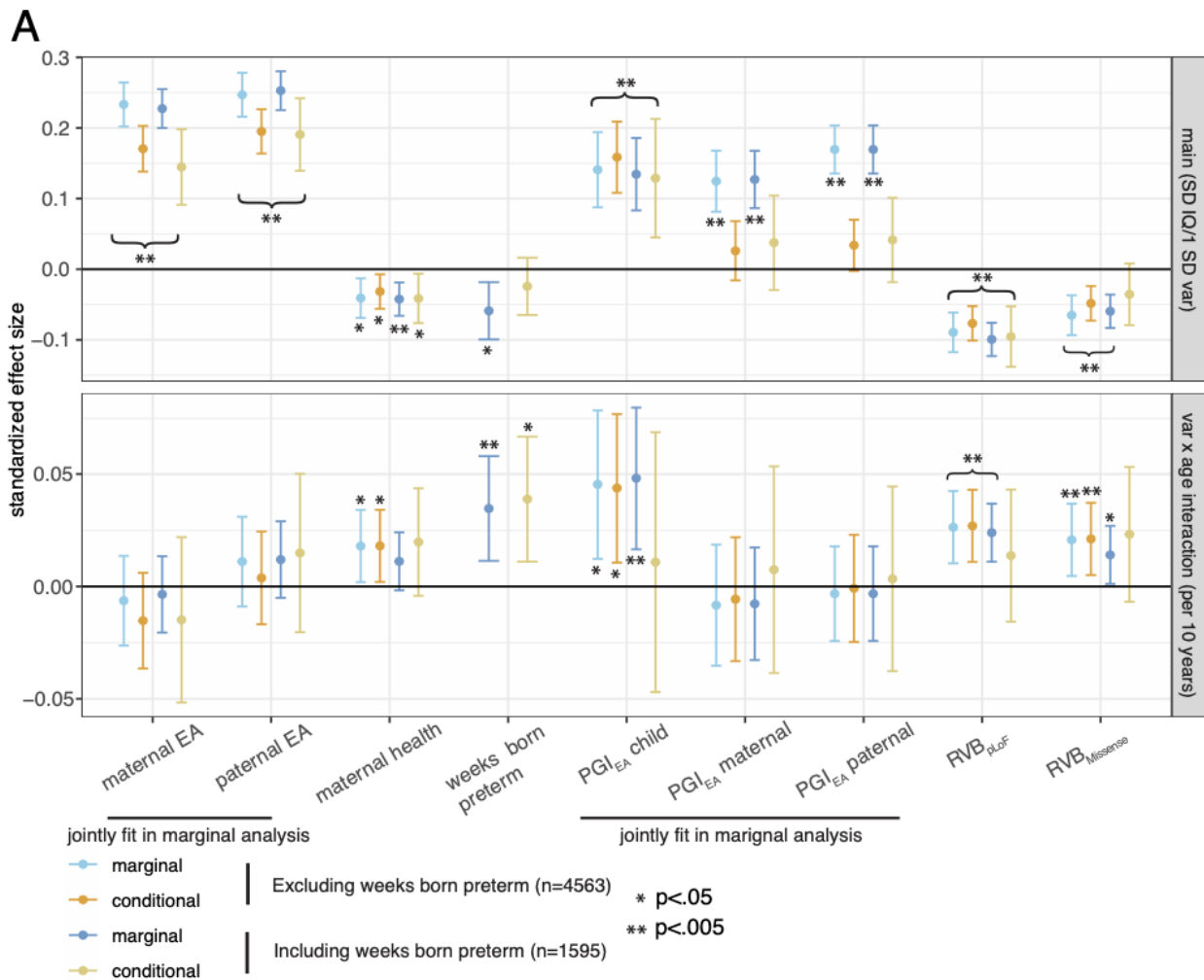
287
288 Secondly, we considered the influence of rare variant burden on these different quantiles of the
289 IQ distribution at different ages in ALSPAC. Broadly, we found that the effects of deleterious rare
290 variants were strongest on those with low IQ, and that these effects attenuate more markedly with
291 age than those on the upper tail of the distribution, reducing the heterogeneity in effects across
292 the phenotype distribution with time (Figure S4; Table S9). Notably, at age 4, the population effect
293 of RVB_{pLoF} at the 5th percentile was 3.9-times greater than that at the 95th percentile ($p=9 \times 10^{-5}$,
294 z test) and 2.2-times greater than at the median ($p=7 \times 10^{-4}$), whereas by age 16, the estimates at
295 the different quantiles were not significantly different from one another. Put another way, this
296 implies that, at age 4, the variance in IQ in people with high RVB_{pLoF} is greater than in those with
297 lower RVB_{pLoF} , but this is driven by there being more people with lower rather than higher IQ, as
298 illustrated in the right-hand plot in Figure 3A; however, by age 16, RVB_{pLoF} shows uniform effects
299 at the distribution tails. As a consequence of this, the significant attenuation of population effects
300 with age was observed only at the 5th percentile and median, but not at the 95th percentile, for
301 both RVB_{pLoF} and $RVB_{Missense}$ (Figure 3B). This finding was recapitulated when examining direct
302 genetic effects of RVB_{pLoF} although not $RVB_{Missense}$ (Figure 3B). Although we had detected no
303 significant effects of parental RVB on mean IQ in the trio model (Figure 2B), we detected a
304 nominally significant negative effect of maternal RVB_{pLoF} on the bottom 5th percentile, suggesting
305 that the effect of genetically-influenced parental behaviors on IQ may vary across the IQ
306 distribution (Extended Data Figure 8). This is in line with our recent finding of a significant effect
307 of non-transmitted rare damaging maternal alleles on risk of neurodevelopmental conditions
308 (Extended Data Figure 10B in ²⁶). Repeating the RVB_{pLoF} analysis on school grades in ALSPAC,
309 we similarly find stronger associations between RVB_{pLoF} and academic performance at the 5th
310 percentile that attenuate as children age (Extended Data Figure 7A). An analysis using cognitive
311 ability measures from MCS and UK Biobank yielded concordant observations (Supplementary
312 Note 5; Extended Data Figure 7B).

313
314 In summary, we show that the negative effects of RVB are strongest on the lower tail of the
315 distribution of cognitive ability, and that these effects attenuate with age. These results stand in
316 contrast to those for the PGIs for which the significant age interactions were seen only in the top
317 half of the IQ distribution.

318 Relative contribution of genetic and other exposures to IQ

319 Finally, we compared the longitudinal effects of common and rare genetic variants on IQ (Figures
320 1 and 2) to the effects of other factors that have previously been associated with children's
321 academic and cognitive outcomes (see Methods). We tested the effects of two perinatal factors,
322 namely maternal illness during pregnancy and premature birth (as measured by the number of
323 weeks born preterm), and the effect of realized parental educational attainment on IQ in ALSPAC.
324 When fitting the variables separately (i.e., marginal associations shown in blue points in Figure
325 4A), we found strong positive effects of paternal and maternal EA (Figure 4A top) that were
326 constant with age (Figure 4A bottom). In contrast, the perinatal factors were negatively associated
327 with IQ (Figure 4A top) but their effects attenuated with age (Figure 4A bottom), consistent with
328 previous findings using school grades in ALSPAC and MCS^{49,50}. The standardized effects for
329 these perinatal exposures were weaker than those for rare variants, with maternal illness and
330 weeks born preterm explaining 0.18% and 0.35% of the variance in IQ at age 4 respectively,
331 versus 1.40% for RVB_{pLoF} and $RVB_{Missense}$ collectively ($p=5.2 \times 10^{-4}$ and 0.027, respectively) (Figure
332 4A top). In a joint model including parental EA, perinatal exposures and genetic scores (offspring
333 and parental PGI_{EA} and the child's RVB_{pLoF} and $RVB_{Missense}$), most effect sizes changed minimally,
334 with the exception of the parental PGI_{EA} effects which became non-significant. Importantly, the
335 effects of the child's genetic scores did not significantly change (Figure 4A top, Supplementary
336 Note 6), and the direct effect of the child's PGI_{EA} was similar to the effects of parental EA.
337 Conclusions were similar when considering a composite measure of childhood cognitive
338 performance in MCS (Extended Data Figure 10A). Collectively, these results show that the
339 magnitude of effects of these genetic measures on childhood IQ is comparable to those of other
340 well-established influences and that the changing effects with age persist after controlling for other
341 environmental factors (Figure 4A bottom).

342
343 Given our observation above of the differential effects of genetic measures on the tails of the IQ
344 distribution, we further considered the influence of parental EA and these perinatal exposures on
345 the different quantiles of the distribution, fitting models for each additional variable independently
346 as well as joint models as mentioned above (Figure 4B; Extended Data Figure 10B;
347 Supplementary Note 6). Our conclusions from the marginal analysis in Figure 3 remain
348 unchanged when controlling for these other factors; specifically, deleterious rare variants show
349 larger effects on the tail end of the IQ distribution in early childhood which attenuate with age,
350 whereas the attenuation of the direct effect of PGI_{EA} is strongest in the top half of the IQ distribution
351 These results indicate that the relative influences of factors that best predict which children will
352 have cognitive difficulties or will excel cognitively across childhood differ from those that best
353 predict average IQ.



355 **Figure 4**

356 Associations between genetic and other factors and IQ across ages in ALSPAC. A) Standardized effects
357 and 95% confidence intervals for the main effects (top) and measure-by-age-interaction effects (bottom)
358 obtained from mixed-effects models fitted for IQ post-imputation. These were obtained from four models,
359 as indicated in the key: marginal associations from models in which only the indicated variable(s) were
360 included in the model (in addition to standard covariates), and conditional associations from models in
361 which all the variables were included in the same model. Note that we estimated the marginal and
362 conditional effects either using only individuals who had gestational age information (n=1,595) or all
363 individuals (n=4,563), dropping the “weeks born preterm” variable in the latter case. See Methods for
364 details. B) Standardized effects and 95% confidence intervals for the main effects (top) and measure-by-
365 age-interaction effects (bottom) for quantile regression of the 5th, 50th, and 95th percentile estimated from
366 mixed-effects quantile regression modeling for IQ post-imputation using both marginal and conditional
367 associations as in (A). In (B), marginal effect sizes for the child’s PGI_{EA}, RVB_{pLoF}, and RVB_{Missense}, which
368 were previously shown in Figures 1 and 2 are replotted to facilitate comparisons of effect size with the
369 conditional analysis results. Asterisks indicate whether the estimate is significantly different from 0; note
370 that curly brackets with asterisks indicate that the estimates spanned by the bracket are all significant.

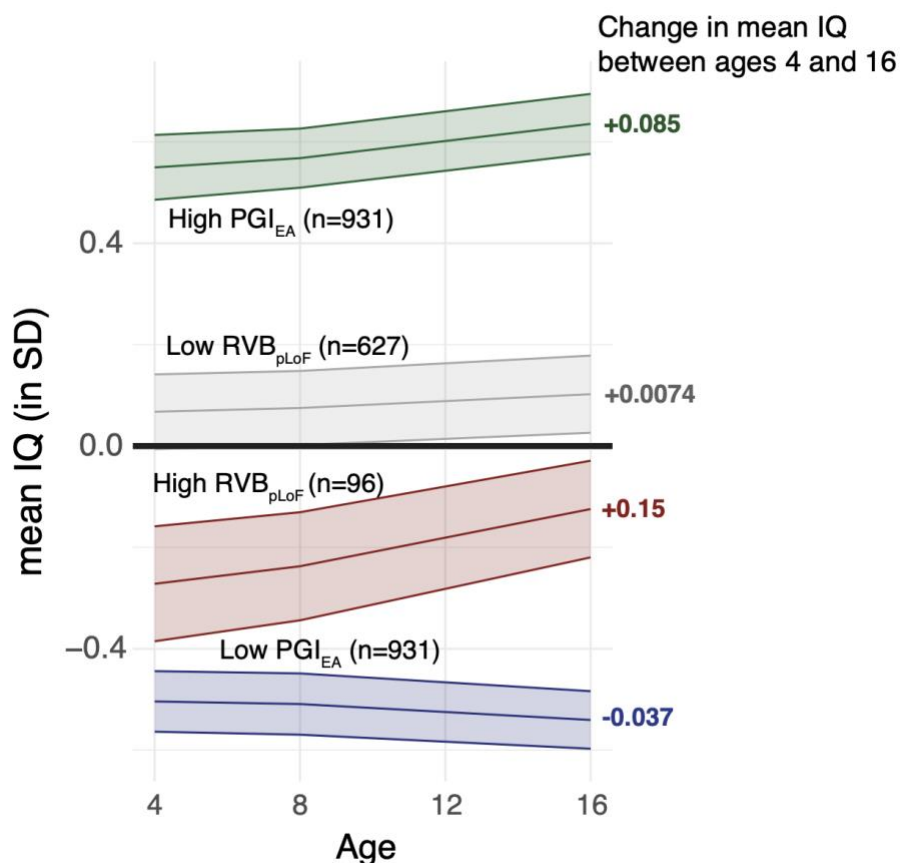
371 **Discussion**

372 Here we explored the contribution of common and rare variants to cognitive performance
373 longitudinally across childhood and adolescence. We showed that common and rare variants
374 differ in their patterns of effects on IQ and school achievement across time. While the effect of
375 rare damaging variants attenuates between childhood and adolescence (Figure 2), the effect of
376 common variants associated with educational attainment and cognitive performance in adulthood
377 increases (Figure 1). In both cases, these changes over time are due to direct genetic effects,
378 with no evidence for an age interaction effect of non-transmitted common or rare alleles in parents
379 (Figure 1B), nor indeed for any effects of non-transmitted rare alleles on average IQ at all (Figure
380 2B). Our common variant findings accord with those of Malanchini *et al.* who showed that the
381 effect of a polygenic score for non-cognitive skills on academic achievement increased over
382 development in a different cohort⁵¹.

383
384 In theory, the increasing effect of polygenic indices for educational attainment and adult cognitive
385 performance with age could simply be because the underlying SNP effects have been estimated
386 on adult phenotypes which are better correlated with IQ in later childhood than early childhood^{52,53}.
387 Having said that, we observed suggestive evidence for an increase in the total SNP heritability of
388 IQ with age in ALSPAC (p=0.053) (Table S2), consistent with the trend reported in another
389 cohort⁴². Additionally, the common variant genetic correlations between IQ measured at the
390 different ages and educational attainment measured in adulthood were not significantly different
391 from 1 (Table S3), suggesting the increase in PGI effects are most likely due to genome-wide
392 amplification of genetic effects as opposed to different common variants impacting cognitive ability
393 differentially at different ages¹¹. By contrast, our finding that the effects of rare variants attenuate
394 as children age cannot be explained by the way the variant weights were obtained; this is because
395 the weights are based on estimates of evolutionary negative selection against deleterious variants
396 in each gene in large population-based cohorts of adults⁴⁶, rather than through genotype-
397 phenotype association studies.

398
399 Genetic studies typically test for genetic effects on the mean of the phenotype, although some
400 have also considered effects on variance^{54–58} or differential effects on different quantiles of
401 anthropometric traits and biochemical measures⁵⁹. Here, we examined the longitudinal effects of
402 common and rare variants on different quantiles of the IQ distribution, which essentially tests for
403 an effect on the phenotypic variance and skew. We found that the main effects of polygenic scores
404 and of RVB_{Missense} did not differ by quantile, whereas the main effect of RVB_{pLoF} was significantly
405 stronger at the 5th than 95th percentile (Figure 3). However, the increasing effects of PGI_{EA} and
406 PGI_{Cog} on mean IQ with age were strongest for the top half of the phenotype distribution, whereas
407 the attenuating effects of RVB_{pLoF} were strongest in the lower tail (Figure 3). The latter observation
408 appears to be driven by RVB_{pLoF} having a larger effect on the 5th than the 95th percentile of IQ at
409 age 4 and 8, but similar effects at the distribution tails by age 16 (Figure S4). Results based on
410 measures of cognitive performance in MCS and UK Biobank were concordant with these
411 observations from ALSPAC (Extended Data Figure 7B). These results imply that the genetic
412 profile of children in the highest and lowest ranges of cognitive ability differ across development
413 (Extended Data Figure 9). Figure 5 summarizes our key findings for the IQ trajectories of
414 individuals with different common and rare variant genetic backgrounds; children in the high
415 versus low PGI_{EA} groups become more stratified across development, with more of the
416 stratification occurring due to increasing IQ in the high PGI_{EA} group, while the stratification in IQ
417 due to RVB_{pLoF} is strongest at the earlier ages and attenuates particularly in the high RVB_{pLoF}
418 group.

419
420 Our results imply that there is higher variance in IQ in early childhood amongst individuals with
421 higher RVB_{pLoF} , which could be potentially due to gene-by-environment or gene-by-gene effects,
422 or simply to stochasticity. Due to this higher variance, there are more people who cross the
423 diagnostic threshold for intellectual disability than there are amongst those with low RVB_{pLoF} , but
424 there are also relatively more people within the typical IQ range, which could explain why these
425 variants have variable and reduced penetrance. Multiple neurodevelopmental conditions exhibit
426 a high degree of phenotypic variability, even amongst carriers of the same causal mutation, such
427 as in neurofibromatosis type 1⁶⁰. Although quantitative measures of intellectual impairment are
428 relatively rare in studies of these conditions, there is evidence, for example, that patients with
429 tuberous sclerosis complex (TSC) have not only lower *mean* IQ but also higher *variance* in IQ
430 than unaffected sibling controls (F-test for a difference in variance between cases and controls;
431 by our analysis, $p=0.004$ for TSC1, $p=0.033$ for TSC2)⁶¹. Taken together, our results support two
432 hypotheses for the incomplete/variable penetrance of rare NDC-associated variants. Firstly, these
433 variants impact variance in cognitive ability early in life, consistent with heterogeneous
434 phenotypes of children with NDCs. Secondly, they have larger effects on average cognitive ability
435 in early childhood than later in life, consistent with apparently clinically unaffected parents passing
436 on these pathogenic variants to affected children.



437

438 Figure 5

439 Summary of phenotypic trajectories across development, as inferred in this work. The plot shows average
440 trajectories of IQ (inferred post-imputation) for individuals grouped by genetic measures. The high/low
441 PGI_{EA} groups are individuals with a value that is one or more standard deviations above/below 0. The low
442 RVB_{pLoF} group are individuals with RVB_{pLoF} below 0.01 (approximately 9.6% of individuals) and the high
443 RVB_{pLoF} group are individuals with RVB_{pLoF} greater than 0.4 (approximately 1.5% of individuals); these
444 cutoffs were chosen arbitrarily for illustrative purposes. The change in mean genetic measure/IQ between
445 ages 4 and 16 are shown on the right hand side of the plots. Bands indicate 95% CI. The decrease in mean
446 IQ in the low PGI_{EA} group is explained by the fact that, due to PGI_{EA} effects becoming stronger towards the
447 top of the IQ distribution, those that had a high IQ earlier in development will tend towards their genetically
448 predicted lower IQ at later ages, thereby lowering the average IQ in this group.

449

450 The increasing heritability of IQ with age is well established but its causes remain unclear¹¹. In
451 contrast to our results for rare variants, we found evidence across cohorts for increasing *direct*
452 genetic effects of PGIs with age, particularly at the top of the IQ distribution; to our knowledge,
453 this has not been previously demonstrated. Recent work by Morris et al. using longitudinal twin
454 modeling⁶² found evidence that the increase in heritability of cognitive performance across
455 development is due to pervasive genetic innovation (i.e., new genetic effects coming into play
456 across time, resulting in genetic correlations between ages that are less than one). In contrast,
457 our molecular genetic results based on common variants suggest that amplification (i.e., identical

458 genetic effects increasing proportionally with age) is the primary process. However, it is likely that
459 these heterogeneous effects across the IQ spectrum that we have observed violate the
460 assumptions of the twin model of Morris et al., and may lead to the appearance of genetic
461 innovation where none exists. The increasing direct effect of common variants on IQ may be
462 relevant to the so-called ‘Matthew effect’, the phenomenon whereby individual differences in
463 ability compound over time and increase gaps in cognitive and academic outcomes as children
464 age^{63–66}. If this effect is particularly driven by those with high cognitive performance, this could
465 explain the increase in common variant effects over time that we observe for the upper tail of the
466 phenotype distribution. Several potential mechanisms could drive this, including evocative gene-
467 environment correlations, such as high-performing children being selected into more from
468 cognitively-stimulating environments, or active gene-environment correlations, such as high-
469 performing children being more effective learners and hence improving their cognitive
470 performance further relative to their peers⁶⁷.

471
472 When comparing the relative magnitude of the main effects of genetic versus other factors on IQ
473 in ALSPAC, we found that the RVB contribution was slightly larger than but broadly comparable
474 to those of the perinatal exposures, and the direct effect of common variants associated with EA
475 was comparable to that of parental EA (Figure 4A top). This highlights the important contribution
476 of genetic variation across the allele frequency spectrum to variance in cognitive performance.
477 We showed that inherited rare damaging coding variants explain more variance in childhood IQ
478 in the general population than do *de novo* mutations (Supplementary Note 4), in contrast to what
479 is observed in cohorts of children with neurodevelopmental disorders^{22,33}. We further showed that
480 the effects of both rare damaging variants and perinatal factors on cognitive ability attenuate as
481 children age (Figure 4A bottom). This may be due to acute, time-limited effects during early
482 neurodevelopment that can be attenuated by later plasticity, potentially in response to
483 environmental influences. If this is true, one might predict that the attenuation of these effects with
484 age are strongest in children from less deprived households. Testing this hypothesis will need
485 larger datasets than studied here. In contrast to these other factors, the effect of parental EA on
486 average IQ does not change across development, whereas the direct effect of EA-associated
487 common variants increases (Figure 4A bottom), suggesting a cumulative influence. These
488 observations have important implications for how we identify people at risk of poor outcomes in
489 early versus in later life and highlight the value of considering genetic data in studies of social and
490 environmental factors influencing cognitive outcomes.

491
492 There are several limitations of this study. Firstly, our primary findings about changing genetic
493 effects with time are based on measurements of IQ at only three time points in a single cohort
494 (ALSPAC). There is a dearth of cohorts with longitudinal measures of IQ in childhood and
495 adolescence with sufficient sample size and appropriate genetic data for us to attempt direct
496 replication. Thus, we have relied on replication with school grades measured at two timepoints in
497 ALSPAC (not totally independent of our IQ results, given the overlapping sample) (Extended Data
498 Figures 2 and 4; Figure S5), or, for some analyses, a single measure of cognitive performance
499 from MCS and UK Biobank (Extended Data Figures 7 and 10, Figures S1 and S6), which we
500 found to show strong genetic correlations with IQ in ALSPAC (r_g 1.09, 95% CI [0.97-1.21]) and
501 with adult cognitive performance⁴⁴ (r_g 0.89, [0.69-1.09]) and EA⁴⁷ (r_g 0.89, [0.72-1.08]). Broadly,

502 the results from these replication analyses supported our primary findings in ALSPAC. Secondly,
503 IQ tests are less reliable in early life⁶⁸, and those used in ALSPAC may be less accurate at
504 measuring the tails of the distribution⁶⁹. If the accuracy at the tails differed between the tests used
505 at different ages, this may have induced spurious interaction effects between the genetic scores
506 and age. However, this does not seem a likely scenario given that we see age interaction effects
507 in opposite directions for common versus rare variants for the same trait, and replicate trends in
508 other cohorts and with other cognitive performance measures. Thirdly, the genetic scores we
509 have used to capture the effects of common and rare variants under-estimate the total liability
510 conferred by these variants. Thus, the questions we have addressed here should be revisited in
511 future once much larger common and rare variant association studies are available, including
512 using whole-genome sequence data to capture the effects of noncoding variants. . A final
513 limitation is that our results may be affected by non-random missingness within the cohorts and
514 biased ascertainment and attrition across time, known to be present in both birth cohorts^{37,70}. We
515 have attempted to mitigate non-random missingness by imputation of missing IQ measures in
516 ALSPAC (Figure S1) and by Mendelian imputation of unmeasured parental genotypes for our trio-
517 based PGI analyses. To mitigate ascertainment bias and attrition, we incorporated weights in
518 MCS to render the sample representative of the whole UK population. In general, our conclusions
519 in ALSPAC were unchanged when analyzing IQ before versus after imputation, although the latter
520 was obviously better powered, and there were some differences particularly in the findings from
521 the quantile regressions (Supplementary Note 7; Figure 3BC *versus* Figure S3; Figure S4A *versus*
522 S4B). One analysis that is likely to have been impacted by ascertainment bias is the estimate of
523 the fraction of the rare variant effect on IQ that is due to *de novo* mutations, which we have
524 probably over-estimated, as noted in Supplementary Note 4 (Extended Data Figure 6).

525
526 Our study prompts several strands of future work. Perhaps most importantly, future research
527 should extend upon our findings to investigate how genetic and other factors might be best
528 employed to identify children at risk of poor cognitive outcomes, so that interventions may be
529 targeted at those who most need help. It should also seek to confirm the relevance of our
530 observations to the incomplete penetrance of rare damaging variants in NDCs. If we had NDC
531 cohorts with genetic and longitudinal phenotype data from both affected children and parents who
532 appear to be clinically unaffected, we could directly test whether in fact, these variants were
533 associated with reduced cognitive ability or educational achievement earlier in the parents' lives,
534 and whether their effects on these traits attenuated as they aged. Finally, future work should seek
535 to better understand why there are differential time-varying genetic effects on children depending
536 on their level of cognitive ability, and why the effect of our rare variant burden scores attenuates
537 with time. For example, is the latter observation purely driven by the expression patterns of
538 evolutionarily constrained genes that are more heavily weighted in this score (Extended Data
539 Figure 5A), with these prenatally-expressed genes becoming progressively less important in brain
540 function as children age? Or is it that children with particularly marked cognitive deficits in early
541 childhood, who appear to be particularly influenced by these rare variants (Figure 3C), are
542 targeted for interventions which help to mitigate the effects of these rare variants over time? More
543 broadly, our results suggest that heteroscedastic and time-varying effects of genetic variants on
544 human phenotypes deserve more exploration.

545 Abbreviations

546 GWAS: Genome-wide association study; NDCs: neurodevelopmental conditions; SNP: single
547 nucleotide polymorphism; ALSPAC: Avon Longitudinal Study of Parents and Children; MCS:
548 Millennium Cohort Study; IQ: Intelligence quotient; PGI: polygenic index; EA: educational
549 attainment; Cog: the cognitive component of educational attainment; NonCog: the non-cognitive
550 component of educational attainment; RBV: rare variant burden; pLoF: predicted loss of function;
551 TSC: tuberous sclerosis complex; QC: quality control; PCA: principal component analysis; MAF:
552 minor allele frequency; HWE: Hardy-Weinberg Equilibrium; LD: linkage disequilibrium.

553 Acknowledgements

554 We thank the Human Genetic Informatics group at the Wellcome Sanger Institute for assisting in
555 the generation and preparation of the whole exome sequencing data for the ALSPAC and MCS
556 cohorts and general technical support. We thank Alexandra Havdhal for helpful discussions.
557

558 ALSPAC: We are extremely grateful to all the families who took part in ALSPAC, the midwives for
559 their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer
560 and laboratory technicians, clerical workers, research scientists, volunteers, managers,
561 receptionists and nurses. The UK Medical Research Council and Wellcome (Grant ref:
562 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. This publication is
563 the work of the authors and Hilary Martin will serve as a guarantor for the contents of this paper.
564 Genome-wide genotyping data was generated by Sample Logistics and Genotyping Facilities at
565 the Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support
566 from 23andMe. This research was specifically funded by the UK Medical Research Council and
567 Wellcome (Grant ref: 076467/Z/05/Z) and the Department for Education and Skills (Grant ref:
568 EOR/SBU/2002/121). A comprehensive list of grants funding is available on the ALSPAC website:
569 <http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf>
570

571 MCS: We are grateful to the Centre for Longitudinal Studies (CLS), UCL Social Research Institute,
572 for the use of these data and to the UK Data Service for making them available. However, neither
573 CLS nor the UK Data Service bear any responsibility for the analysis or interpretation of these
574 data.
575

576 UKB: This research has been conducted using the UK Biobank Resource under application
577 number 44165.
578

579 This research was funded in part by Wellcome (grant no. 220540/Z/20/A, "Wellcome Sanger
580 Institute Quinquennial Review 2021–2026"). For the purpose of open access, the authors have
581 applied a CC-BY public copyright license to any author accepted manuscript version arising from
582 this submission. D.S.M. is supported by a Gates Cambridge Scholarship (OPP1144).
583

584 Author contributions

585 DSM conducted all of the analyses. DSM, MK, PD, QH, WH, and EW carried out data preparation
586 and quality control with supervision by SL, MEH, and HCM. MK led the whole exome sequencing
587 data preparation for ALSPAC and MCS with assistance from DSM and EW. PD conducted the *de*
588 *novo* variant calling in ALSPAC and MCS and conducted variant quality control for the UK Biobank
589 whole exome sequencing data. WH conducted sample quality control for the UK Biobank whole
590 exome sequencing data and assisted with preparation of UK Biobank exome data. OW assisted
591 with selection of other non-genetic factors and generated PGI SNP weights. DSM, OW, QH, RA,
592 MEH and HCM provided key intellectual input. HCM supervised the analyses and directed the
593 study. DSM and HCM wrote the first draft of the manuscript with input from MEH. All authors read
594 and commented on the final manuscript.
595

596 Declarations

597 MEH is a co-founder of, consultant to and holds shares in Congenica, a genetics diagnostic
598 company. The remaining authors declare no competing interests.
599

600 Data availability

601 Researchers can apply to access data from ALSPAC
602 (<https://www.bristol.ac.uk/alspac/researchers/access/>), MCS ([https://cls.ucl.ac.uk/dataaccess-](https://cls.ucl.ac.uk/dataaccess-training/data-access/)
603 [training/data-access/](https://cls.ucl.ac.uk/dataaccess-training/data-access/)), and UK Biobank ([https://www.ukbiobank.ac.uk/enable-your-](https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access)
604 [research/apply-for-access](https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access)).
605

606 Online Methods

607 Cohorts

608 We report results from the Avon Longitudinal Study of Parents and Children (ALSPAC)^{37,71},
609 Millennium Cohort Study³⁸ (MCS), and UK Biobank cohorts.
610

611 ALSPAC

612 In ALSPAC, pregnant women with expected deliveries between April 1st 1991 and December 31
613 December were recruited in the greater Bristol area (formerly Avon county), resulting in an initial
614 sample of 14,541 pregnancies enrolled into the study, of which 13,988 resulted in live births of
615 children surviving to age 1. Data collected after the age of 7 was available for an additional 906
616 pregnancies from other phases of enrollment, which resulted in an additional 913 children that
617 survived to age 1. At initial enrollment, 14,203 unique mothers were in the study, which increased

618 to 14,833 unique mothers enrolled after the additional phases of enrollment (G0 mothers). The
619 partners of G0 mothers (G0 partners) were also invited to participate in the study, of which 12,113
620 provided data at one point in the study and 3,807 are currently enrolled. From birth to early
621 adulthood, mother and children were followed up with questionnaires and clinical and
622 psychometric data collection across several time points. Biosamples used for genotyping and
623 exome sequencing were obtained from most children and some of the mothers and fathers.

624
625 In the current study, data from 6,495 unrelated children with European inferred genetic ancestry
626 (G1 children; described below) were used, along with genetic and/or survey data collected on
627 4,968 G0 mothers and 4,563 G0 partners. Please note that the study website contains details of
628 all the data that is available through a fully searchable data dictionary and variable search tool:
629 <http://www.bristol.ac.uk/alspac/researchers/our-data/>. Ethical approval for the study was obtained
630 from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees.
631 Consent for biological samples has been collected in accordance with the Human Tissue Act
632 (2004). Informed consent for the use of data collected via questionnaires and clinics was obtained
633 from participants following the recommendations of the ALSPAC Ethics and Law Committee at
634 the time. At age 18, study children were sent 'fair processing' materials describing ALSPAC's
635 intended use of their health and administrative records and were given clear means to consent or
636 object via a written form. Data were not extracted for participants who objected, or who were not
637 sent fair processing materials.

638

639 MCS

640 MCS recruited 18,552 pregnant mothers (2000-2002) using a sampling scheme to ensure a
641 nationally representative sample across the UK, as previously described³⁸. Mothers and children
642 were followed longitudinally and had genetic data collected at age 14 from children, mother, and
643 fathers where available. Ethical approval for the collection of saliva samples from these individuals
644 as part of the sixth sweep was obtained from London-Central Research Ethics Committee.
645 Genotype and newly generated exome data are available for ~8,000 and ~7,000 children,
646 respectively (~13,000 and ~7,000 parents).

647

648 UK Biobank

649 The UK Biobank is a prospective cohort of over 500,000 individuals sampled throughout the UK
650 between 2006 and 2010. Individuals have extensive phenotype data and genetic data including
651 genotype array⁷² and whole exome sequence data⁷³.

652 Cognitive performance measures across cohorts

653 In this study we used various cognitive performance and school performance measures across
654 the three cohorts.

655

656 In ALSPAC, children had IQ measured at ages 4 (Wechsler Preschool and Primary Scale of
657 Intelligence), 8 (Wechsler Intelligence Scale for Children), and 16 (Wechsler Abbreviated Scale

658 of Intelligence). Linked educational records are available, including national standardized exam
659 scores in English, Math, and Science administered at the end of Key Stage 2 (Year 6 at around
660 11 years old) and Key Stage 3 (Year 9 at around 14 years old). To create a composite measure
661 of academic performance, we standardized each score and, at a given Key Stage examination
662 point, computed a one-factor model score using *factanal* function in R using the Bartlett method
663 for scoring, explaining 74% and 77% of the variance at Key Stage 2 and 3 exams, respectively.
664 The factor loadings and factor variance explained for the three exams were nearly identical at
665 Key Stage 2 and 3 (Table S1), implying longitudinal invariance.

666
667 In MCS, children completed various cognitive performance tests at several ages. Those at later
668 ages had a greatly reduced sample size, so we focused on those at ages 3 to 7. These included
669 Bracken School Readiness at age 3, reading vocabulary at ages 3 and 5, pattern construction at
670 ages 5 and 7, and word reading and progress in math at age 7. These were summarized into a
671 single cognitive performance measure using a one-factor model score using the *factanal* function
672 in R and Bartlett method for scoring, explaining 39% of the variance.

673
674 In UK Biobank, individuals were asked thirteen questions testing verbal-numerical reasoning in a
675 limited time frame during the initial assessment visit. We used the sum of questions answered
676 correctly as a measure of fluid intelligence (data field 20016.0), as has been used previously in
677 genetic studies of cognitive ability ⁷⁴.

678 Imputation of IQ values in ALSPAC

679 Imputation of missing IQ values was carried out using *SoftImpute*⁷⁵ using `rank.max = 4.5` and
680 `lambda = 4.5`. Before imputation, all variables were standardized to have mean 0 and variance
681 1. To assess the imputation accuracy obtained using three potential sets of variables (described
682 below), we set 100 random individuals with measured IQ values at a given age to missing,
683 conducted phenotype imputation, and calculated correlations between the true measured values
684 and the imputed values, repeating this procedure 100 times.

685
686 We considered the following sets of variables:

- 687 1) Base set: sex (kz021), birth weight (kz030), maternal and paternal socio-economic groups
688 (b_seg_m and b_seg_p), and IQ values at ages 4 (cf813), 8 (f8ws112), and 16 (fh6280).
- 689 2) Expanded set: base set plus development score at age 2 (cf783), sociability score at age
690 3 (kg623a/c), additional verbal and performance IQ at age 4 (cf811, cf812), nonword
691 repetition and multisyllabic word repetition at age 5 (cf470, cf480), communication score
692 at age 6 (kq517), Skuse social cognition score at age 7 (kr554a/b), cognitive scores at
693 age 8 including Sky Search, Diagnostic Analysis of Nonverbal Accuracy, nonword
694 repetition, verbal and performance IQ, children communication checklist score (f8at062,
695 f8dv443, f8dv444, 8dv445, f8dv446, f8sl100, f8sl101, f8sl102, f8ws110, and f8ws111,
696 ku506a).
- 697 3) Auxiliary set: expanded set excluding base set variables.

698

699 When calculating the imputation accuracy using the expanded or auxiliary sets, we removed
700 verbal and performance IQ at each age, although in practice very few children had only one
701 measured at a given age.

702
703 For the imputation of the final IQ values, we used the expanded set and required that each
704 individual had nonmissing values for at least one IQ test and at least 20% of the variables used
705 for imputation. Final IQ values were standardized to mean 0 and standard deviation of 1.

706 Genotype data preparation and imputation

707 ALSPAC

708 ALSPAC genotype data were generated and processed as described in ⁷¹. We further removed
709 individuals with high genotype missingness >3%. We restricted analyses to autosomal SNPs with
710 MAF >0.5%, missingness rate <3%, and that had HWE test $p > 1 \times 10^{-5}$.

711
712 We used KING⁷⁶ for relatedness inference and removed 152 samples who had first degree
713 relationships but were reported as coming from different families and another 16 samples who
714 did match available exome sequencing data supposedly for the same individuals resulting in
715 8,831 children, 9,302 mothers, and 1,706 fathers. To identify a set of unrelated children, we
716 iteratively identified the child with the greatest number of genetically inferred relatives (3rd degree
717 or closer) and removed them, recalculated the number of inferred relatives per child, and repeated
718 the first step until no children were inferred to have any relatives. This resulted in a total of 6,495
719 unrelated children with genotype array and whole exome sequence data.

720
721 To identify individuals of genetically inferred European ancestry, we projected samples onto 1,000
722 Genomes phase 3 individuals⁷⁷ using the smartpca function from EIGENSOFT version 7.2.1⁷⁸.
723 We used linkage disequilibrium (LD)-pruned SNPs (pairwise $r^2 < 0.2$ in batches of 50 SNPs with
724 sliding windows of 5) with MAF > 5% and removed 24 regions with high or long-range LD,
725 including the HLA⁷⁹. All ALSPAC samples projected onto European ancestry samples. We then
726 performed PCA identically on the unrelated ALSPAC samples and projected all individuals into
727 the PCA space.

728
729 Prior to imputation, we removed palindromic SNPs, SNPs that were not in the imputation
730 reference panel, and SNPs with mismatched alleles. We imputed the samples to the TOPMed r2
731 reference panel using the TOPMed imputation server⁸⁰⁻⁸². We kept well imputed common variants
732 with Minimac4 $R^2 > 0.8$ and MAF >1%.

733 MCS

734 MCS genotyped data were generated and processed as described in ⁸³. A set of unrelated
735 individuals with genetically-inferred European ancestry were described in a similar manner to
736 ALSPAC, as described previously ²⁶. We kept well imputed common variants with Minimac4 R^2
737 >0.8 and MAF >1%.

738

739 UK Biobank

740 UK Biobank genotype data were generated, processed, and imputed by UKB as described in ⁷².
741 Sample quality control consisted of excluding individuals with >3% missingness, inconsistent sex,
742 sex aneuploidy, or withdrawn consent and relatedness was similarly calculated using KING as
743 previously described. As in the other cohorts, we kept autosomal SNPs with MAF >1%,
744 missingness rate <3%, and passed the HWE test ($p\text{-value} > 1 \times 10^{-5}$). To identify UKB individuals
745 with genetically inferred European ancestry, we similarly projected samples onto the 1,000
746 Genomes Phase 3 individuals and assigned individuals to a genetic ancestry based on their
747 Mahalanobis distance to the nearest continental ancestry group centroid using the top 6 PCs.
748 Those more than 6 standard deviations from the centroid along any axis were removed. Unrelated
749 individuals were identified using the iterative exclusion procedure as previously described,
750 resulting in 387,531 unrelated individuals. PCs were recalculated in this subset of unrelated
751 individuals using smartpca as previously described, and related individuals were projected into
752 this PCA space.

753 Calculating polygenic scores

754 SNP weights for polygenic scores were estimated using LDpred2-auto⁸⁴, which does not require
755 a tuning dataset and automatically estimates the required hyper-parameters from the discovery
756 sample⁸⁵. The LD reference panel was computed from unrelated individuals from the target
757 dataset for a set of 1,444,196 HapMap3+⁸⁶ variants⁷². GWAS summary statistics for EA¹⁶
758 excluding 23andMe or 23andMe only (used for PGI_{EA} UK Biobank samples), EA-Cog and EA-
759 NonCog⁴⁴ were matched with the list of overlapping SNPs.

760
761 Once the weights were generated, we scored individuals using the --score function in PLINK v1.9,
762 which calculates the weighted sum of genotypes across a set of SNPs for each individual.

763 Mendelian imputation of parental genotypes

764 When an individual had at least one parent with genotype data in the cohort, we imputed the
765 expected parental genotype of the other parent. To do so, we used *snipar*²⁰ across the three
766 cohorts. We supplied relatedness inference using KING output and used the default options
767 throughout. We used the previous weights derived from each GWAS using LDpred2-auto to
768 calculate the PGIs in the full trios using the pgs.py script provided in the *snipar* package.

769 GWAS of IQ and cognitive performance

770 We used the linear regression function *lm* in R to conduct GWASs on the IQ measures pre- and
771 post-imputation in ALSPAC and the cognitive performance measure in MCS on unrelated
772 genetically inferred European ancestry individuals, controlling for 10 genetic PCs and sex. We
773 removed variants with MAF < 1% or missingness > 2%.

774 Heritability and genetic correlations

775 We used both GREML-LDMS⁴¹ and LD score regression^{87,88} to estimate SNP heritabilities and
776 genetic correlations from the summary statistics of the GWASs conducted on the IQ measures
777 and external GWASs using the previously described LD reference panel on HapMap3 SNPs.

778 Other exposures impacting IQ in ALSPAC and MCS

779 We considered the following exposures: maternal and paternal educational attainment (EA),
780 weeks born preterm, and a composite variable representing maternal illness in ALSPAC only. For
781 educational attainment, in ALSPAC we used the highest qualification reported by the parent and
782 coded it as done previously reported¹⁶, with a university degree equalling 20 years of education,
783 A levels as 13, O levels/vocational as 10, and Certificate of Secondary Education/no degree as 5
784 (c645a and c666a for the mother and father, respectively) and in MCS we used the equivalent
785 self-reported derived variables in sweep 1 from each parent questionnaire (APLFTE00). Maternal
786 illness was defined as a binary variable indicating whether or not a mother had at least one of the
787 following conditions reported during the pregnancy in her obstetric clinical records: preeclampsia,
788 anemia (DELP_1060), any diabetes (pregnancy_diabetes), or genital herpes, gonorrhoea, syphilis,
789 urinary tract infection, vaginal infection, or hepatitis B noted during delivery (DEL_P1050-1055).
790 Weeks born preterm was coded as 40 minus the gestational age at birth in weeks assessed in a
791 formal pediatric assessment (DEL_B4401). All variables were standardized to have mean zero
792 and standard deviation 1.

793 Exome sequencing data preparation

794 Whole exome sequencing data quality control for ALSPAC and MCS cohorts was carried out by
795 the Human Genetics Informatics team at the Sanger Institute as described in ⁸⁹. Briefly, GATK
796 v4.2 was used to call short variants (SNVs and indels) in 11,994 samples from ALSPAC and
797 11,916 samples from MCS. Sample quality-control (QC) measures were employed to remove
798 outliers on several metrics (e.g., heterozygosity, variant counts) and likely sample mismatches.
799 To identify low quality variants (variant QC), a random forest was trained on pre-defined truth sets
800 in each cohort individually. The random forest filtering was then applied in combination with
801 genotype-level and missingness filters to balance precision, recall, true and false positive rates,
802 and synonymous transmission ratios. Specifically, SNVs were filtered (genotypes set to missing)
803 if they had an allele depth (DP) < 5, a heterozygous allele balance ratio (AB) < 0.2, or a genotype
804 quality (GQ) < 20 (ALSPAC) or < 15 (MCS); Indels were filtered using these thresholds: DP < 10,
805 AB < 0.3, GQ < 10 (ALSPAC) or GQ < 20 (MCS). The variants were excluded if they failed the
806 random forest filtering or if the fraction of missing genotypes (missingness) exceeded 0.5. The
807 final dataset included 8,436 children and 3,215 parents in ALSPAC and 7,667 children and 6,925
808 parents in MCS. Calling and QC of *de novo* mutations is described in the Supplementary Methods.

809
810 In UK Biobank, we performed quality control for whole exome sequencing data on 469,836
811 participants within the UKB research analysis platform. First, we split and left-aligned multi-allelic
812 variants in the population-level Variant Call Format files into separate alleles using *bcftools*
813 *norm*⁹⁰. Next, we performed genotype-level filtering using *bcftools* filter separately for Single

814 Nucleotide Variants (SNVs) and Insertions/Deletions (InDels) using a missingness-based
815 approach. Specifically, SNV genotypes with a depth lower than 7 and genotype quality lower than
816 20, or InDel genotypes with a depth lower than 10 and genotype quality lower than 20, were set
817 to missing. We further tested for an expected alternate allele contribution of 50% for heterozygous
818 SNVs using a binomial test, and SNV genotypes with a binomial test p-value ≤ 0.0001 were set
819 to missing. Finally, we recalculated the proportion of individuals with a missing genotype for each
820 variant and excluded all variants with a missingness value greater than 50%.

821 Classification of deleterious rare variants

822 All variants were annotated using the MANE transcript from Ensembl⁹¹. In MCS and ALSPAC,
823 rare pLoFs were defined as pLoFs annotated as high confidence by LOFTEE³⁵ that had CADD⁹²
824 (if SNVs) > 25 , were not located in the last exon or intron, and had a gnom-AD³⁵ V3 allele
825 frequency of $< 3 \times 10^{-5}$ (up to ~ 5 occurrences in gnomAD r3 genomes & ~ 10 x in exomes) and an
826 in sample allele frequency of $< 0.1\%$ among the unrelated set of children from that cohort⁹¹. Rare
827 damaging missense variants were defined using identical allele frequency and CADD filters to
828 pLoFs but were additionally required to have an MPC score⁹³ ≥ 2 . Rare synonymous variants
829 were defined with identical allele frequency thresholds.

830
831 In UKB individuals with inferred European genetic ancestry, we defined rare variants as those
832 with a within-sample allele frequency $< 0.001\%$. For pLoFs, we retained only those variants
833 defined as high-confidence pLoFs by LOFTEE and CADD > 25 as in MCS and ALSPAC. For
834 missense variants, we defined the damaging missense variants by including variants with REVEL
835 > 0.5 , AlphaMissense⁹⁴ > 0.56 , and MPC > 2 .

837 Calculating rare variant burden scores

838 We calculated RVB using the following formula:

$$839 \quad RVB_{i,c} = \sum_{g \in G} I_{g,c}(i) S_g$$

840 where $I_{g,c}(i)$ is an indicator function for whether an individual i has a variant of consequence c in
841 gene g , S_g is the fitness cost for heterozygous carriers of a pLoF allele in gene i estimated in Sun
842 *et al.*⁴⁶, and G is the set of all autosomal genes.

843
844 For analyses to ascertain the relative contribution of *de novo* versus inherited variants
845 (Supplementary Note 4), we also calculated a separate rare variant burden metric which we call
846 constrained variant count:

$$847 \quad \text{constrained pLoF count}_i = \sum_{g \in \text{constrained genes}} I_{g,pLoF}(i)$$

848 where $I_{g,pLoF}(i)$ is an indicator function for whether an individual i has a pLoF in gene g , and
849 *constrained genes* is the set of genes identified as constrained in ⁴⁶.
850

851 Sampling and non-response weights in MCS

852 Sampling weights were developed by MCS to adjust for the nonrandom sampling scheme devised
853 for the study⁹⁵. We used the full UK sampling weights. Nonresponse inverse probability weights
854 were developed for each sample as described previously²⁶. Briefly, we used logistic regression to
855 predict whether an individual had complete data for a given regression. We then extracted the
856 predicted values for all individuals with complete data from the logistic regression which
857 correspond to their relative probabilities of having complete data, calculated the inverse of that
858 value, and used that as the nonresponse weight. To combine sampling and nonresponse weights,
859 we multiplied the two weights and incorporated them into the regressions.

860 Associations between genetic measures and traits

861 Cross-sectional associations for a genetic score and IQ/cognitive performance measure were
862 conducted using the *lm* function in R, restricting to unrelated samples with genetically-inferred
863 European ancestry as follows:

$$864 IQ_i \sim G_i + PC1_i + \dots + PC10_i + sex_i$$

865 where G_i is the genetic score (PGI or RVB) for child i and $PC1-10_i$ are their genetic PCs.

866
867 Genetic scores, unless otherwise stated, were standardized to mean 0 and standard deviation 1.
868 In trio-based analyses, the parental genetic values were similarly standardized and regression
869 was conducted as follows:

$$870 IQ_i \sim G_i + G_{i,m} + G_{i,p} + PC1_i + \dots + PC10_i + sex_i$$

871 where $G_{i,m}$ and $G_{i,p}$ are the genetic scores for child i 's mother and father, respectively.

872
873 Mixed effect linear models were conducted using the *lme4* package in R⁹⁶. The same covariates
874 were used for the cross-sectional analysis with the addition of age, an age-by-genetic score
875 interaction effect and a child-specific random intercept term as follows:

$$876 IQ_{i,t} \sim age_{i,t} + G_i + age_{i,t} * G_i + PC1_i + \dots + PC10_i + sex_i + C_i$$

877 where $age_{i,t}$ and $IQ_{i,t}$ are child i 's age and IQ at age t and C_i is a child-specific random effect
878 intercept. Age was coded as age-4 such that the genetic predictor intercept estimate was
879 equivalent to the effect at the first age at which IQ data were collected. In the trio-based analyses,
880 an age interaction effect was additionally modeled for the parental genetic values as follows:

$$881 IQ_{i,t} \sim age_{i,t} + G_i + age_{i,t} * G_i + G_{i,m} + age_{i,t} * G_{i,m} + G_{i,p} + age_{i,t} * G_{i,p} + PC1_i + \dots + PC10_i + sex_i + C_i$$

882 For cross-sectional and longitudinal quantile regression modeling, we used the R package
883 *quantreg*⁹⁷ and *rqpd*⁹⁸, respectively, using the same models as above. Mixed-effects model
884 standard errors were determined using bootstrapping with 500 bootstrap replications.

885
886 For the academic performance score, the cross-sectional models were performed in the same
887 way described above, with the addition of age at testing in weeks as a covariate (*ks2age_w* and

888 ks3age_w). As there were only two timepoints, a slightly different longitudinal model was used for
889 the linear modeling (Extended Data Figures 2 and 4) as follows:
890 $(APS_{9,i} - APS_{6,i}) \sim G_i + age_{6,i} + age_{9,i} + PC1_i + \dots + PC10_i + sex_i$
891 where $APS_{6/9,i}$ and $age_{6/9,i}$ are the academic performance scores and ages of individual i at Year
892 6 and 9, respectively. The coefficient estimated for G_i in this regression is equivalent to G 's effect
893 at Year 9 minus that at Year 6 within an individual, i.e. the age interaction effect. To compare the
894 effects for the quantile regression effect size estimates at Years 6 and 9, we conducted z tests
895 between the effect sizes (Extended Data Figure 7A).

896 Marginal and conditional associations between genetic and other factors and 897 IQ

898 To assess the associations between genetic and other factors with IQ (Figure 4), we considered
899 both marginal and conditional models. For marginal models, we conducted the following
900 regressions:

$$901 IQ_{i,t} \sim age_{i,t} + E_i + age_{i,t} * E_i + PC1_i + \dots + PC10_i + sex_i + C_i$$

902 where E_i is the given measure of interest. In Figure 4, the coefficient on the E_i term (i.e. the main
903 effect) is shown in the top panel, and the coefficient on the $age_{i,t} * E_i$ term (i.e. the interaction effect)
904 in the bottom panel.

905
906 As maternal and paternal EA are highly correlated, we modified the model for the effects of
907 parental EA as follows:

$$908 IQ_{i,t} \sim age_{i,t} + EA_{i,m} + EA_{i,p} + age_{i,t} * EA_{i,m} + age_{i,t} * EA_{i,p} + PC1_i + \dots + PC10_i + sex_i + C_i$$

909 where $EA_{i,m}$ and $EA_{i,p}$ are the maternal and paternal EA respectively. Similarly, we fit the PGI_{EA}
910 values for the child, mother, and father jointly in a trio model.

911
912 To calculate the variance explained by a given variable, we took the square of the standardized
913 effect size of that variable's main effect. To calculate the total variance explained by two
914 uncorrelated variables (e.g. RVB_{pLoF} and $RVB_{Missense}$ that have a correlation of -0.01, $p=0.16$), we
915 summed the squares of their standardized effect sizes. To test for difference in effect size between
916 the combination of $RVB_{Missense}$ and RVB_{pLoF} versus other factors, we used the square root of the
917 previous variance explained estimate as the effect size estimate and determined the standard
918 error for that effect size by summing the square of the standard error for each effect estimate and
919 then taking the square root of the new estimate. We were then able to use z tests to compare the
920 effect size estimates as done previously.

921
922 For the full joint models, we included all genetic and other measures in a single model, with a
923 main and age interaction effect for each measure. Since only a subset of probands had
924 information on gestational age, we fitted the full joint model both with and without the 'weeks born
925 preterm' variable.

926 References

927 1. Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A. & Goldberg, L. R. The Power of

- 928 Personality: The Comparative Validity of Personality Traits, Socioeconomic Status, and
929 Cognitive Ability for Predicting Important Life Outcomes. *Perspect. Psychol. Sci.* **2**, 313–
930 345 (2007).
- 931 2. Johnson, W., Corley, J., Starr, J. M. & Deary, I. J. Psychological and physical health at age
932 70 in the Lothian Birth Cohort 1936: links with early life IQ, SES, and current cognitive
933 function and neighborhood environment. *Health Psychol.* **30**, 1–11 (2011).
- 934 3. Strenze, T. Intelligence and socioeconomic success: A meta-analytic review of longitudinal
935 research. *Intelligence* **35**, 401–426 (2007).
- 936 4. Schmidt, F. L. & Hunter, J. General mental ability in the world of work: occupational
937 attainment and job performance. *J. Pers. Soc. Psychol.* **86**, 162–173 (2004).
- 938 5. Deary, I. J., Whalley, L. J. & Starr, J. M. *A Lifetime of Intelligence: Follow-up Studies of the*
939 *Scottish Mental Surveys of 1932 and 1947*. (American Psychological Association, 2009).
- 940 6. Christensen, G. T., Mortensen, E. L., Christensen, K. & Osler, M. Intelligence in young
941 adulthood and cause-specific mortality in the Danish Conscription Database – A cohort
942 study of 728,160 men. *Intelligence* **59**, 64–71 (2016).
- 943 7. Deary, I. J., Cox, S. R. & Hill, W. D. Genetic variation, brain, and intelligence differences.
944 *Mol. Psychiatry* **27**, 335–353 (2022).
- 945 8. Husby, A., Wohlfahrt, J. & Melbye, M. Gestational age at birth and cognitive outcomes in
946 adolescence: population based full sibling cohort study. *BMJ* **380**, e072779 (2023).
- 947 9. Makharia, A. *et al.* Effect of environmental factors on intelligence quotient of children. *Ind.*
948 *Psychiatry J.* **25**, 189–194 (2016).
- 949 10. Brynne, M. *et al.* Maternal infection during pregnancy and likelihood of autism and
950 intellectual disability in children in Sweden: a negative control and sibling comparison
951 cohort study. *Lancet Psychiatry* **9**, 782–791 (2022).
- 952 11. Plomin, R. & Deary, I. J. Genetics and intelligence differences: five special findings. *Mol.*

- 953 *Psychiatry* **20**, 98–108 (2015).
- 954 12. Savage, J. E. *et al.* Genome-wide association meta-analysis in 269,867 individuals
955 identifies new genetic and functional links to intelligence. *Nat. Genet.* **50**, 912–919 (2018).
- 956 13. Howe, L. J. *et al.* Within-sibship genome-wide association analyses decrease bias in
957 estimates of direct genetic effects. *Nat. Genet.* **54**, 581–592 (2022).
- 958 14. Kong, A. *et al.* The nature of nurture: Effects of parental genotypes. *Science* **359**, 424–428
959 (2018).
- 960 15. Young, A. I. *et al.* Relatedness disequilibrium regression estimates heritability without
961 environmental bias. *Nat. Genet.* **50**, 1304–1310 (2018).
- 962 16. Okbay, A. *et al.* Polygenic prediction of educational attainment within and between families
963 from genome-wide association analyses in 3 million individuals. *Nat. Genet.* **54**, 437–449
964 (2022).
- 965 17. Nivard, M. G. *et al.* More than nature and nurture, indirect genetic effects on children’s
966 academic achievement are consequences of dynastic social processes. *Nat Hum Behav* **8**,
967 771–778 (2024).
- 968 18. Wang, B. *et al.* Robust genetic nurture effects on education: A systematic review and meta-
969 analysis based on 38,654 families across 8 cohorts. *Am. J. Hum. Genet.* **108**, 1780–1791
970 (2021).
- 971 19. Demange, P. A. *et al.* Estimating effects of parents’ cognitive and non-cognitive skills on
972 offspring education using polygenic scores. *Nat. Commun.* **13**, 4801 (2022).
- 973 20. Young, A. I. *et al.* Mendelian imputation of parental genotypes improves estimates of direct
974 genetic effects. *Nat. Genet.* **54**, 897–905 (2022).
- 975 21. Tan, T. *et al.* Family-GWAS reveals effects of environment and mating on genetic
976 associations. *Genetic and Genomic Medicine* (2024).
- 977 22. Deciphering Developmental Disorders Study. Prevalence and architecture of de novo

- 978 mutations in developmental disorders. *Nature* **542**, 433–438 (2017).
- 979 23. Martin, H. C. *et al.* Quantifying the contribution of recessive coding variation to
980 developmental disorders. *Science* **362**, 1161–1164 (2018).
- 981 24. Martin, H. C. *et al.* The contribution of X-linked coding variation to severe developmental
982 disorders. *Nat. Commun.* **12**, 627 (2021).
- 983 25. Niemi, M. E. K. *et al.* Common genetic variants contribute to risk of rare severe
984 neurodevelopmental disorders. *Nature* **562**, 268–271 (2018).
- 985 26. Huang, Q. Q. *et al.* Dissecting the contribution of common variants to risk of rare
986 neurodevelopmental conditions. *bioRxiv* (2024) doi:10.1101/2024.03.05.24303772.
- 987 27. Chen, C.-Y. *et al.* The impact of rare protein coding genetic variation on adult cognitive
988 function. *Nat. Genet.* **55**, 927–938 (2023).
- 989 28. Gardner, E. J. *et al.* Reduced reproductive success is associated with selective constraint
990 on human genes. *Nature* **603**, 858–863 (2022).
- 991 29. Kingdom, R., Beaumont, R. N., Wood, A. R., Weedon, M. N. & Wright, C. F. Genetic
992 modifiers of rare variants in monogenic developmental disorder loci. *Nat. Genet.* **56**, 861–
993 868 (2024).
- 994 30. Weiner, D. J. *et al.* Polygenic architecture of rare coding variation across 394,783 exomes.
995 *Nature* **614**, 492–499 (2023).
- 996 31. Wright, C. F. *et al.* Evaluating variants classified as pathogenic in ClinVar in the DDD
997 Study. *Genet. Med.* **23**, 571–575 (2021).
- 998 32. Wolstencroft, J. *et al.* Neuropsychiatric risk in children with intellectual disability of genetic
999 origin: IMAGINE, a UK national cohort study. *Lancet Psychiatry* **9**, 715–724 (2022).
- 1000 33. Samocha, K. E. *et al.* Substantial role of rare inherited variation in individuals with
1001 developmental disorders. *bioRxiv* (2024) doi:10.1101/2024.08.28.24312746.
- 1002 34. Wright, C. F. *et al.* Assessing the Pathogenicity, Penetrance, and Expressivity of Putative

- 1003 Disease-Causing Variants in a Population Setting. *Am. J. Hum. Genet.* **104**, 275–286
1004 (2019).
- 1005 35. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in
1006 141,456 humans. *Nature* **581**, 434–443 (2020).
- 1007 36. Falconer, D. S. The inheritance of liability to certain diseases, estimated from the incidence
1008 among relatives. *Ann. Hum. Genet.* **29**, 51–76 (1965).
- 1009 37. Fraser, A. *et al.* Cohort Profile: the Avon Longitudinal Study of Parents and Children:
1010 ALSPAC mothers cohort. *Int. J. Epidemiol.* **42**, 97–110 (2013).
- 1011 38. Connelly, R. & Platt, L. Cohort profile: UK Millennium Cohort Study (MCS). *Int. J.*
1012 *Epidemiol.* **43**, 1719–1725 (2014).
- 1013 39. Wechsler, D. Wechsler Adult Intelligence Scale--Third Edition. *Front. Psychol.*
1014 doi:10.1037/t49755-000.
- 1015 40. Niileksela, C. R. & Reynolds, M. R. Enduring the tests of age and time: Wechsler constructs
1016 across versions and revisions. *Intelligence* **77**, 101403 (2019).
- 1017 41. Yang, J. *et al.* Genetic variance estimation with imputed variants finds negligible missing
1018 heritability for human height and body mass index. *Nat. Genet.* **47**, 1114–1120 (2015).
- 1019 42. Trzaskowski, M., Yang, J., Visscher, P. M. & Plomin, R. DNA evidence for strong genetic
1020 stability and increasing heritability of intelligence from age 7 to 12. *Mol. Psychiatry* **19**, 380–
1021 384 (2014).
- 1022 43. Mollon, J. *et al.* Genetic influence on cognitive development between childhood and
1023 adulthood. *Mol. Psychiatry* **26**, 656–665 (2021).
- 1024 44. Demange, P. A. *et al.* Investigating the genetic architecture of noncognitive skills using
1025 GWAS-by-subtraction. *Nat. Genet.* **53**, 35–44 (2021).
- 1026 45. Benjamin, D. J., Cesarini, D., Turley, P. & Young, A. S. Social-Science Genomics:
1027 Progress, Challenges, and Future Directions. (2024) doi:10.3386/w32404.

- 1028 46. Sun, K. Y. *et al.* A deep catalogue of protein-coding variation in 983,578 individuals. *Nature*
1029 (2024) doi:10.1038/s41586-024-07556-0.
- 1030 47. Lee, J. J. *et al.* Gene discovery and polygenic prediction from a genome-wide association
1031 study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121
1032 (2018).
- 1033 48. Li, M. *et al.* Integrative functional genomic analysis of human brain development and
1034 neuropsychiatric risks. *Science* **362**, (2018).
- 1035 49. Odd, D., Evans, D. & Emond, A. M. Prediction of school outcome after preterm birth: a
1036 cohort study. *Arch. Dis. Child.* **104**, 348–353 (2019).
- 1037 50. Alterman, N. *et al.* Gestational age at birth and academic attainment in primary and
1038 secondary school in England: Evidence from a national cohort study. *PLoS One* **17**,
1039 e0271952 (2022).
- 1040 51. Malanchini, M. *et al.* Genetic associations between non-cognitive skills and academic
1041 achievement over development. *Nat Hum Behav* (2024) doi:10.1038/s41562-024-01967-9.
- 1042 52. McCall, R. B. Childhood IQ's as Predictors of Adult Educational and Occupational Status.
1043 *Science* **197**, 482–483 (1977).
- 1044 53. Keage, H. A. d. *et al.* Age 7 intelligence and paternal education appear best predictors of
1045 educational attainment: The Port Pirie Cohort Study. *Aust. J. Psychol.* **68**, 61–69 (2016).
- 1046 54. Wang, H. *et al.* Genotype-by-environment interactions inferred from genetic effects on
1047 phenotypic variability in the UK Biobank. *Sci Adv* **5**, eaaw3538 (2019).
- 1048 55. Paré, G., Cook, N. R., Ridker, P. M. & Chasman, D. I. On the use of variance per genotype
1049 as a tool to identify quantitative trait interaction effects: a report from the Women's Genome
1050 Health Study. *PLoS Genet.* **6**, e1000981 (2010).
- 1051 56. Cao, Y., Wei, P., Bailey, M., Kauwe, J. S. K. & Maxwell, T. J. A versatile omnibus test for
1052 detecting mean and variance heterogeneity. *Genet. Epidemiol.* **38**, 51–59 (2014).

- 1053 57. Ek, W. E. *et al.* Genetic variants influencing phenotypic variance heterogeneity. *Hum. Mol.*
1054 *Genet.* **27**, 799–810 (2018).
- 1055 58. Pershad, Y., Poisner, H., Corty, R. W., Hellwege, J. N. & Bick, A. G. Variance quantitative
1056 trait loci reveal gene-gene interactions which alter blood traits. *Genetic and Genomic*
1057 *Medicine* (2024).
- 1058 59. Wang, C. *et al.* Genome-wide discovery for biomarkers using quantile regression at
1059 biobank scale. *Nat. Commun.* **15**, 6460 (2024).
- 1060 60. Pasmant, E., Vidaud, M., Vidaud, D. & Wolkenstein, P. Neurofibromatosis type 1: from
1061 genotype to phenotype. *J. Med. Genet.* **49**, 483–489 (2012).
- 1062 61. Tye, C. *et al.* Long-term cognitive outcomes in tuberous sclerosis complex. *Dev. Med. Child*
1063 *Neurol.* **62**, 322–329 (2020).
- 1064 62. Morris, D. M., Ritchie, S. J. & Plomin, R. Mechanisms of developmental change in genetic
1065 and environmental influences on intelligence. *PsyArXiv* (2024) doi:10.31234/osf.io/uqybg.
- 1066 63. Bast, J. & Reitsma, P. Analyzing the development of individual differences in terms of
1067 Matthew effects in reading: results from a Dutch Longitudinal study. *Dev. Psychol.* **34**,
1068 1373–1399 (1998).
- 1069 64. Shaywitz, B. A. *et al.* A Matthew Effect for IQ but Not for Reading: Results from a
1070 Longitudinal Study. *Read. Res. Q.* **30**, 894–906 (1995).
- 1071 65. Stanovich, K. E. Matthew Effects in Reading: Some Consequences of Individual
1072 Differences in the Acquisition of Literacy. *Journal of Education* **189**, 23–55 (2009).
- 1073 66. Duff, D., Tomblin, J. B. & Catts, H. The Influence of Reading on Vocabulary Growth: A
1074 Case for a Matthew Effect. *J. Speech Lang. Hear. Res.* **58**, 853–864 (2015).
- 1075 67. Tucker-Drob, E. M. & Harden, K. P. Early childhood cognitive development and parental
1076 cognitive stimulation: evidence for reciprocal gene-environment transactions. *Dev. Sci.* **15**,
1077 250–259 (2012).

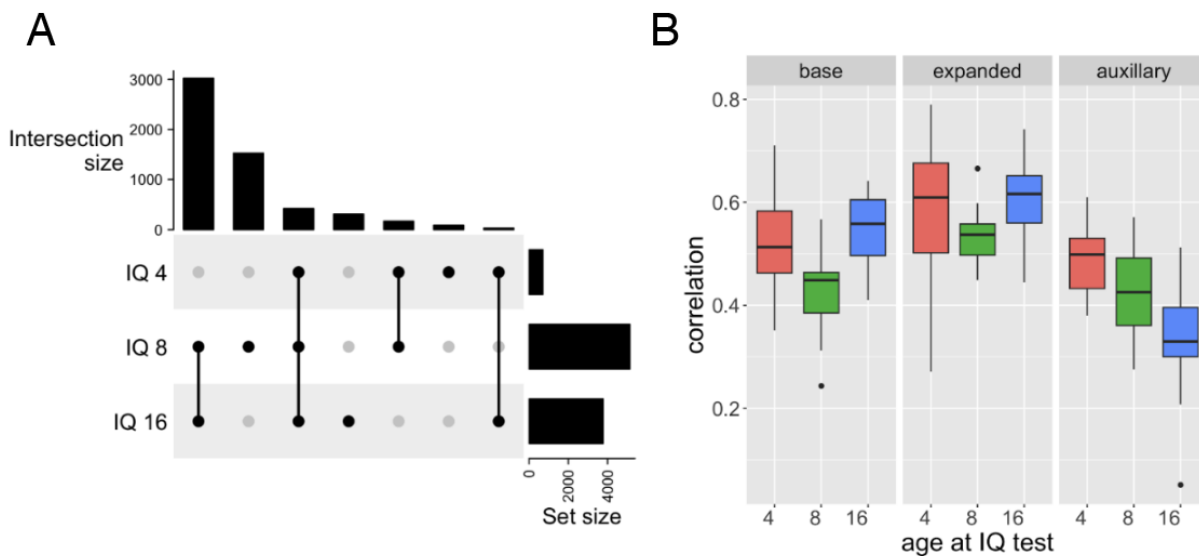
- 1078 68. Schuerger, J. M. & Witt, A. C. The temporal stability of individually tested intelligence. *J.*
1079 *Clin. Psychol.* **45**, 294–302 (1989).
- 1080 69. Whitaker, S. Error in the estimation of intellectual ability in the low range using the WISC-IV
1081 and WAIS-III. *Pers. Individ. Dif.* **48**, 517–521 (2010).
- 1082 70. Mostafa, T. & Ploubidis, G. Millennium Cohort Study.
1083 https://discovery.ucl.ac.uk/id/eprint/10060140/1/mcs6_report_on_response.pdf.
- 1084 71. Boyd, A. *et al.* Cohort Profile: the ‘children of the 90s’--the index offspring of the Avon
1085 Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **42**, 111–127 (2013).
- 1086 72. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data.
1087 *Nature* **562**, 203–209 (2018).
- 1088 73. Backman, J. D. *et al.* Exome sequencing and analysis of 454,787 UK Biobank participants.
1089 *Nature* **599**, 628–634 (2021).
- 1090 74. Davies, G. *et al.* Study of 300,486 individuals identifies 148 independent genetic loci
1091 influencing general cognitive function. *Nat. Commun.* **9**, 2098 (2018).
- 1092 75. Hastie, T., Mazumder, R., Lee, J. D. & Zadeh, R. Matrix Completion and Low-Rank SVD via
1093 Fast Alternating Least Squares. *J. Mach. Learn. Res.* **16**, 3367–3402 (2015).
- 1094 76. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies.
1095 *Bioinformatics* **26**, 2867–2873 (2010).
- 1096 77. 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation.
1097 *Nature* **526**, 68–74 (2015).
- 1098 78. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.*
1099 **2**, e190 (2006).
- 1100 79. Price, A. L. *et al.* Long-range LD can confound genome scans in admixed populations. *Am.*
1101 *J. Hum. Genet.* **83**, 132–5; author reply 135–9 (2008).
- 1102 80. Taliun, D. *et al.* Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program.

- 1103 *Nature* **590**, 290–299 (2021).
- 1104 81. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat. Genet.* **48**,
1105 1284–1287 (2016).
- 1106 82. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat.*
1107 *Genet.* **48**, 1279–1283 (2016).
- 1108 83. Fitzsimons, E. *et al.* Collection of genetic data at scale for a nationally representative
1109 population: the UK Millennium Cohort Study. *Longit. Life Course Stud.* **13**, 169–187 (2021).
- 1110 84. Privé, F., Arbel, J. & Vilhjálmsón, B. J. LDpred2: better, faster, stronger. *Bioinformatics* **36**,
1111 5424–5431 (2021).
- 1112 85. Vilhjálmsón, B. J. *et al.* Modeling Linkage Disequilibrium Increases Accuracy of Polygenic
1113 Risk Scores. *Am. J. Hum. Genet.* **97**, 576–592 (2015).
- 1114 86. Privé, F., Albiñana, C., Arbel, J., Pasaniuc, B. & Vilhjálmsón, B. J. Inferring disease
1115 architecture and predictive ability with LDpred2-auto. *Am. J. Hum. Genet.* **110**, 2042–2055
1116 (2023).
- 1117 87. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits.
1118 *Nat. Genet.* **47**, 1236–1241 (2015).
- 1119 88. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity
1120 in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
- 1121 89. Koko, M. *et al.* Exome sequencing of UK birth cohorts. *Wellcome Open Research* (2024).
- 1122 90. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, (2021).
- 1123 91. Kersey, P. J. *et al.* Ensembl Genomes 2016: more genomes, more complexity. *Nucleic*
1124 *Acids Res.* **44**, D574–80 (2016).
- 1125 92. Rentzsch, P., Schubach, M., Shendure, J. & Kircher, M. CADD-Splice-improving genome-
1126 wide variant effect prediction using deep learning-derived splice scores. *Genome Med.* **13**,
1127 31 (2021).

- 1128 93. Samocha, K. E. *et al.* Regional missense constraint improves variant deleteriousness
1129 prediction. *bioRxiv* 148353 (2017) doi:10.1101/148353.
- 1130 94. Cheng, J. *et al.* Accurate proteome-wide missense variant effect prediction with
1131 AlphaMissense. *Science* **381**, eadg7492 (2023).
- 1132 95. Plewis, I. Millennium cohort study first survey: Technical report on sampling (3rd. Edition).
1133 (2004).
- 1134 96. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models using
1135 lme4. *arXiv [stat.CO]* (2014).
- 1136 97. Wan, C. & Zhong, W. Estimation and Inference for Multi-Kink Quantile Regression [R
1137 package MultiKink version 0.1.0]. (2020).
- 1138 98. Koenker, R. & Bache, S. H. rqpd: Regression quantiles for panel data. *R package version*
1139 *0.6/r10* (2014).
- 1140 99. DDG2P (Version 3.79). <https://panelapp.genomicsengland.co.uk/panels/484/>.
- 1141 100. Pardiñas, A. F. *et al.* Common schizophrenia alleles are enriched in mutation-intolerant
1142 genes and in regions under strong background selection. *Nat. Genet.* **50**, 381–389 (2018).
- 1143

1144 Extended Data Figures

1145 Extended Data Figure 1



1146
1147 Assessing quality of IQ imputation across ages and sets of variables in ALSPAC. A) Upset plot indicating
1148 number of individuals with a given set of measured IQ values. B) Correlations between masked measured
1149 and imputed values for 100 random individuals for 100 trials for three different combinations of variables
1150 (base, expanded and auxiliary variables, as described in Methods) used in the SoftImpute imputation for
1151 IQ measured at age 4, 8, and 16.

1152

1153

1154

1155

1156

1157

1158

1159

1160

1161

1162

1163

1164

1165

1166

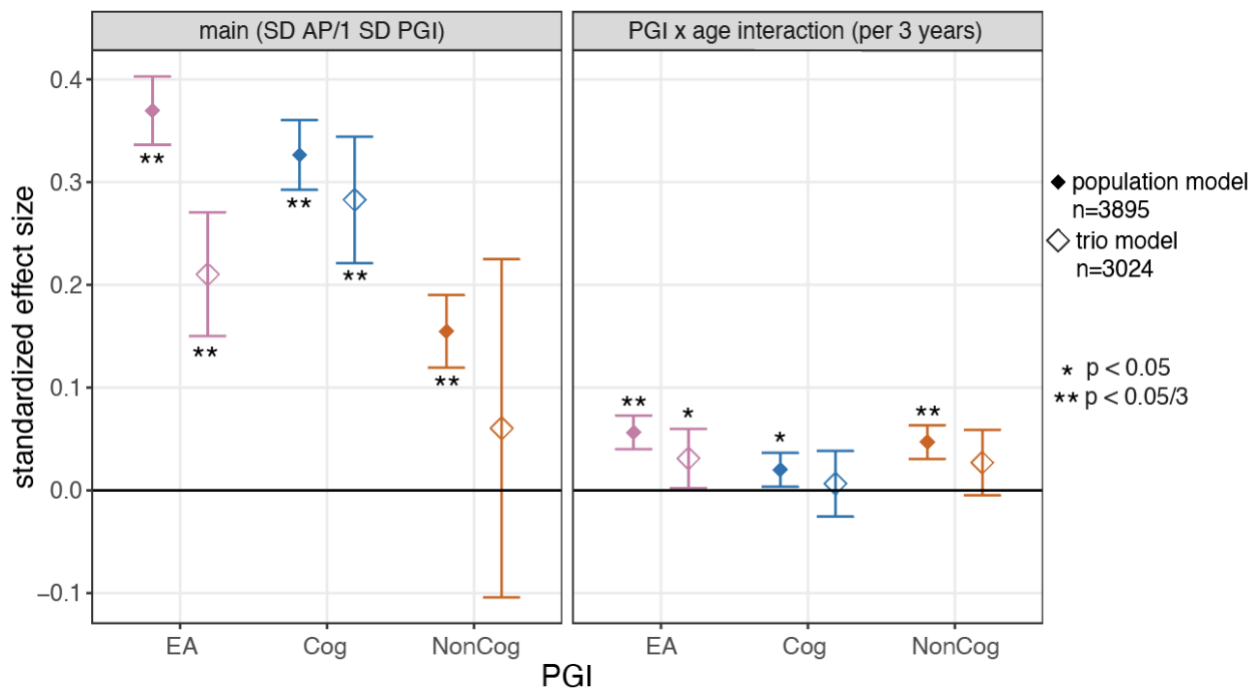
1167

1168

1169

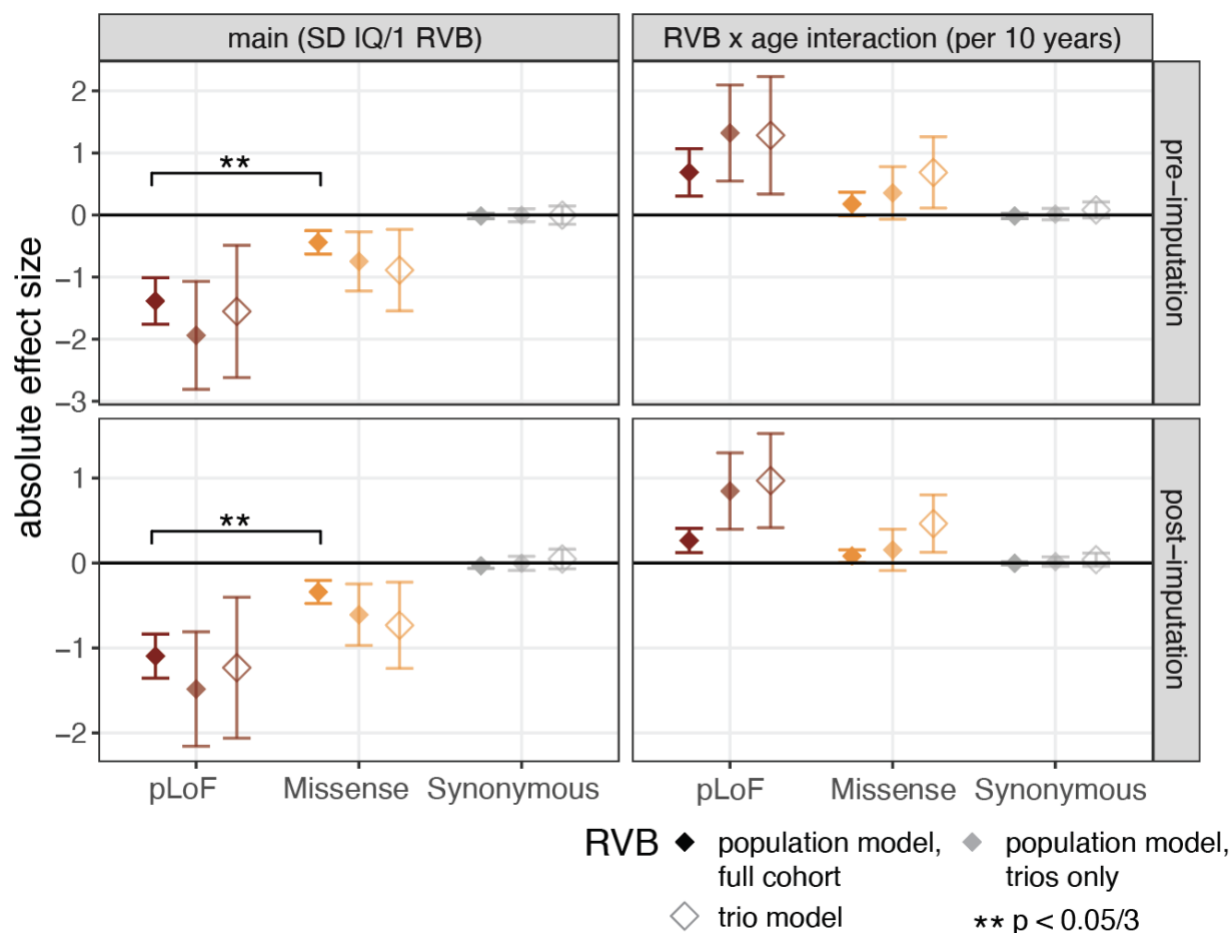
1170

1171 Extended Data Figure 2



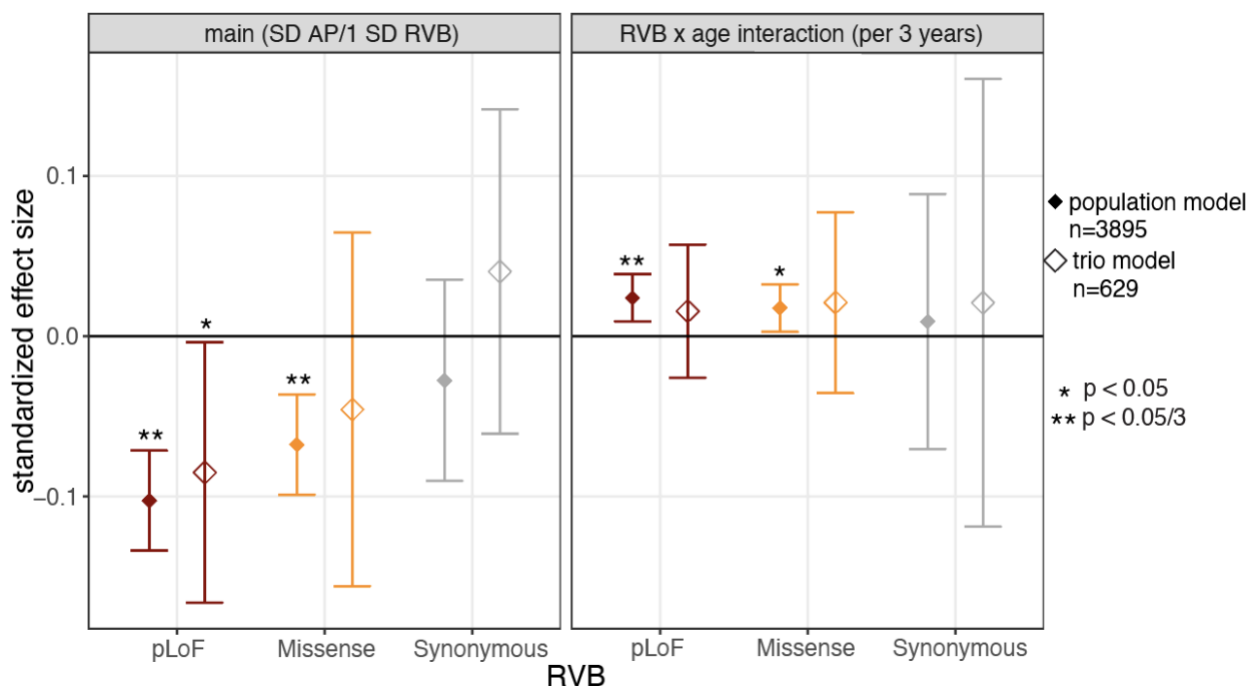
1172 Association between PGIs and academic performance in ALSPAC. The points show the effect at age 11
 1173 (main effect, left panel) and the difference in effects between age 14 and 11 (PGI x age interaction, right
 1174 panel) and error bars show 95% confidence intervals. Results are shown for the children's PGIs not
 1175 controlling (i.e. 'population model') and controlling (right) for parental PGIs. Population effect sizes are
 1176 shown estimated in the full sample.
 1177
 1178
 1179

1180 Extended Data Figure 3



1181
 1182 Association between unstandardized rare variant burden (RVB) and IQ across ages in ALSPAC, estimated
 1183 with a mixed-effects linear model. (i.e. as for Figure 2 but with unstandardized rather than standardized
 1184 RVB scores.) Absolute effects and 95% confidence intervals estimated for the main effects and RVB-by-
 1185 age-interaction effects for RVB calculated with three different consequence classes (pLoF, missense or
 1186 synonymous), either pre- (top) and post-imputation (bottom). Results for the children's RVBs not controlling
 1187 (i.e. 'population model') and controlling (right) for parental RVBs (i.e. trio model). Population effect sizes
 1188 are shown estimated in the full sample (opaque) and in the subset of children with parental RVBs
 1189 (translucent). The square brackets indicate significant comparisons highlighted in the text (z tests).
 1190
 1191

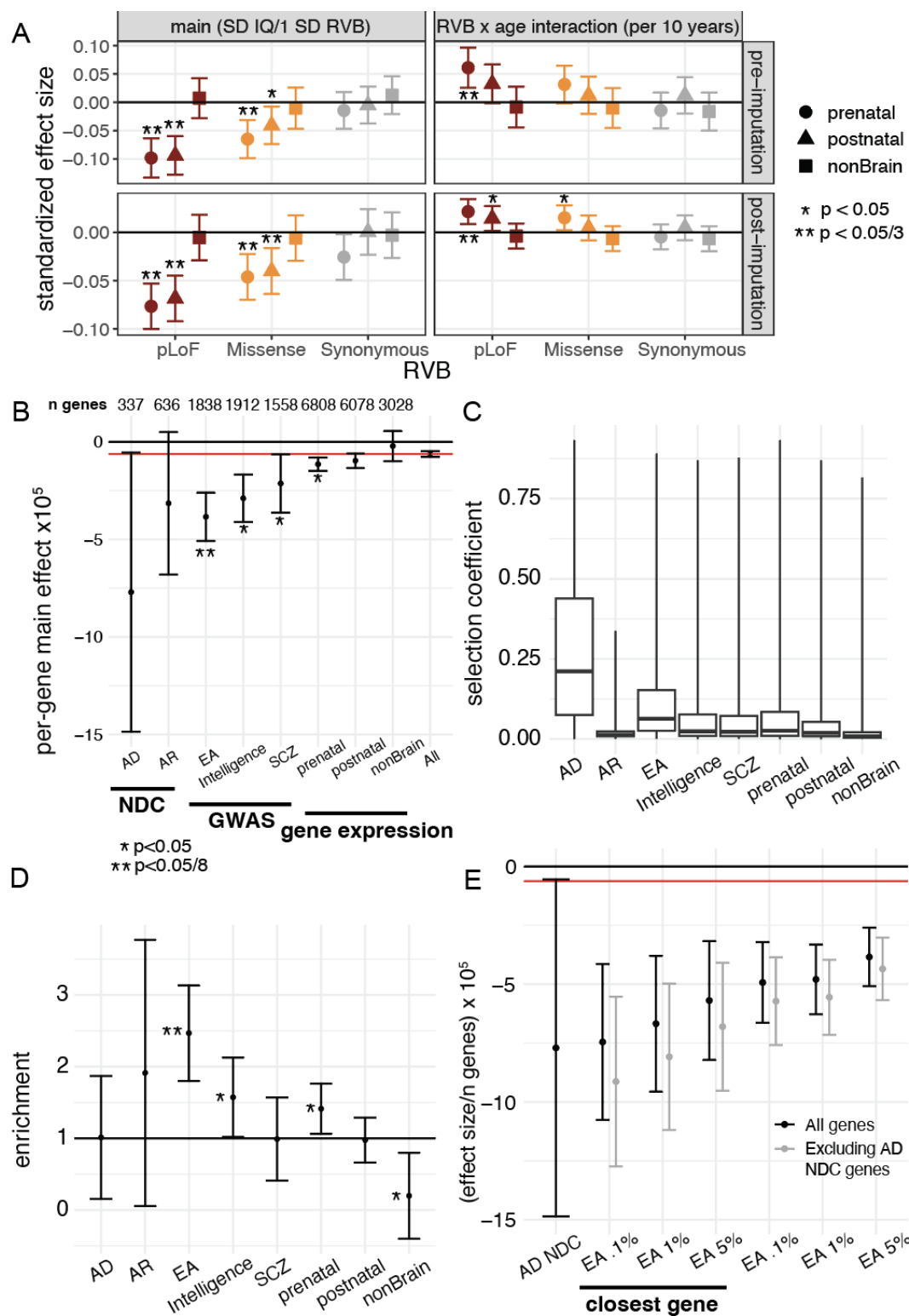
1192 Extended Data Figure 4



1193 Association between RVBs and academic performance in ALSPAC. The points show the effect at age 11
 1194 (main effect, left panel) and the difference in effects between age 14 and 11 (RVB x age interaction, right
 1195 panel) and error bars show 95% confidence intervals. Results are shown for the children's RVB not
 1196 controlling (i.e. 'population model') and controlling (right) for parental RVB. Population effect sizes are
 1197 shown estimated in the full sample.
 1198
 1199

1200
 1201
 1202
 1203
 1204
 1205
 1206
 1207
 1208
 1209
 1210
 1211
 1212
 1213
 1214
 1215
 1216
 1217
 1218

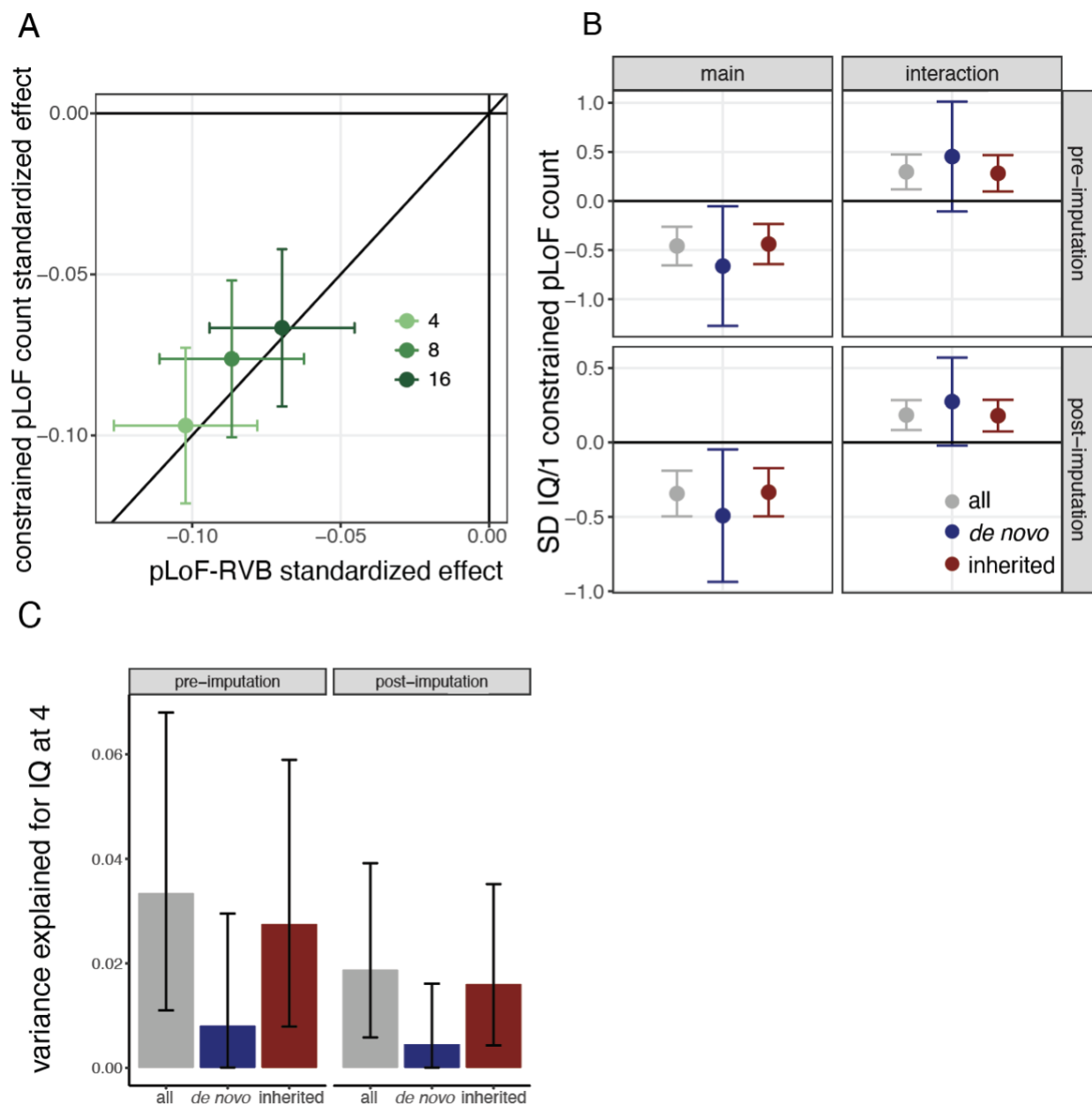
1219 Extended Data Figure 5



1220 Associations and enrichment of gene set-specific RVB across development. A) Standardized effects and
 1221 95% confidence intervals estimated for main effects and RVB-by-age-interaction effects for each RVB in
 1222 three mutually exclusive gene sets from Li et al⁴⁸. These comprise genes in co-expressed clusters that are
 1223

1224 more highly expressed in prenatal or postnatal brain or that are not detected in the study (non-brain). B)
1225 Main effects of RVB_{pLoF} on IQ (post-imputation) stratified by gene set, divided by the number of genes in
1226 the given gene set, with 95% confidence intervals. Red horizontal line indicates the average effect for
1227 RVB_{pLoF} across all genes. Asterisks indicate the p value for difference in per-gene effects between a given
1228 gene set and all genes using a z test, with * indication nominal significance and ** indicating bonferroni
1229 significance for 8 tests. C) Boxplots of the distribution of s_{het} (selection coefficient against heterozygous
1230 pLoFs in that gene⁴⁶) per gene set, where a coefficient of 0 indicates no selection against heterozygous
1231 pLoFs in a given gene and 1 indicates a 100% reduction in fitness for heterozygous pLoF carriers relative
1232 to non-carriers. D) Ratio of the main effect for RVB_{pLoF} for the indicated gene set relative to randomly
1233 sampled gene sets with matching underlying s_{het} distributions (enrichment). E) As in (B) for gene sets
1234 defined using different FDR threshold cutoffs based on gene prioritization in ⁴⁷ (5%, 1%, 0.1%) and by
1235 restricting to prioritized genes at a given cutoff that are also the closest genes to the prioritizing SNP in the
1236 Lee et al. GWAS. Results are shown before (black) or after (grey) excluding overlapping genes from the
1237 set of autosomal dominant neurodevelopmental condition genes with a loss-of-function mechanism from
1238 DDG2P (AD NDC). AD/AR NDC: Autosomal dominant/recessive neurodevelopmental disorder genes with
1239 loss-of-function mechanism from DDG2P⁹⁹, EA: educational attainment GWAS prioritized genes by Lee et
1240 al.⁴⁷ at 5% FDR threshold, Intelligence: intelligence GWAS prioritized genes from Savage et al.¹², SCZ:
1241 schizophrenia GWAS prioritized genes from Pardiñas et al.¹⁰⁰

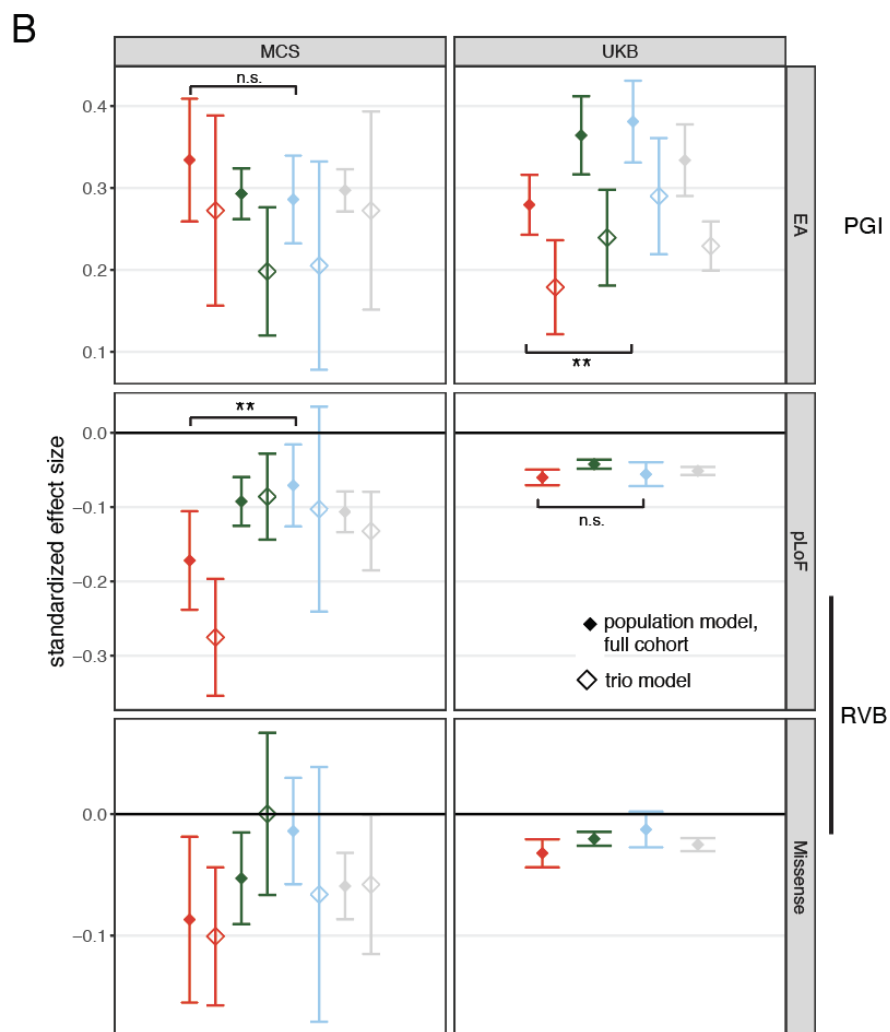
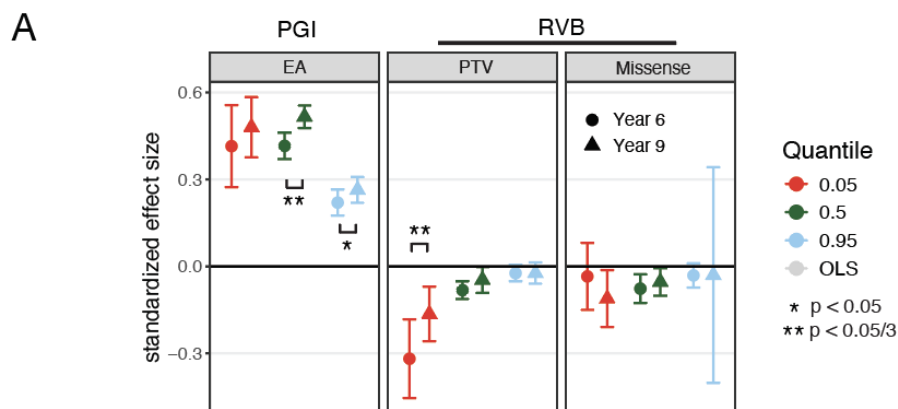
1242 Extended Data Figure 6



1243 Association between inherited and *de novo* pLoF variant counts in constrained genes and IQ in ALSPAC.
 1244 A) Standardized effects and 95% confidence intervals of RVB_{pLoF} and constrained pLoF count (see
 1245 **Supplementary Methods**) on IQ across ages post-imputation. Black line indicates the $y=x$ line. B)
 1246 Standardized effects for the main and RVB-by-age interaction effect from a longitudinal mixed-effects model
 1247 of constrained pLoF counts on standardized IQ when considering all variants or *de novo* mutations and
 1248 inherited variants separately, both pre- and post-imputation. C) Variance explained by constrained pLoF
 1249 counts on standardized IQ when considering all variants or *de novo* and inherited separately, both pre- and
 1250 post-imputation. See Figure S1 for results from a similar analysis of cognitive performance in MCS.
 1251
 1252
 1253

1254 Extended Data Figure 7

1255



1256

1257

1258

1259

Influence of genetic measures on the tails of the phenotypic distribution using academic performance in ALSPAC and cognitive performance measures in MCS and UK Biobank. A) Standardized effects and 95% confidence intervals for quantile regression of the 5th (red), 50th (green), and 95th (blue) percentiles of

1260 genetic measures on academic performance in Year 6 and Year 9 in the full ALSPAC cohort showing
1261 population effects. The square brackets with asterisks indicate significant age interactions. B) Standardized
1262 effects and 95% confidence intervals for quantile regression of the 5th (red), 50th (green), and 95th (blue)
1263 percentiles and linear regression (OLS; gray) estimated from cross-sectional associations with cognitive
1264 performance measures in MCS and UK Biobank. The PGI_{EA} used in UK Biobank was constructed from an
1265 EA GWAS in 23andMe only ¹⁶, as the GWAS excluding 23andMe includes UK Biobank. We only considered
1266 PGI_{EA} as we could not construct PGI_{Cog} to apply in UK Biobank since we did not have an independent
1267 GWAS for cognitive ability in adults that excluded this cohort. The square brackets indicate significant
1268 comparisons highlighted in the text (z tests).

1269

1270

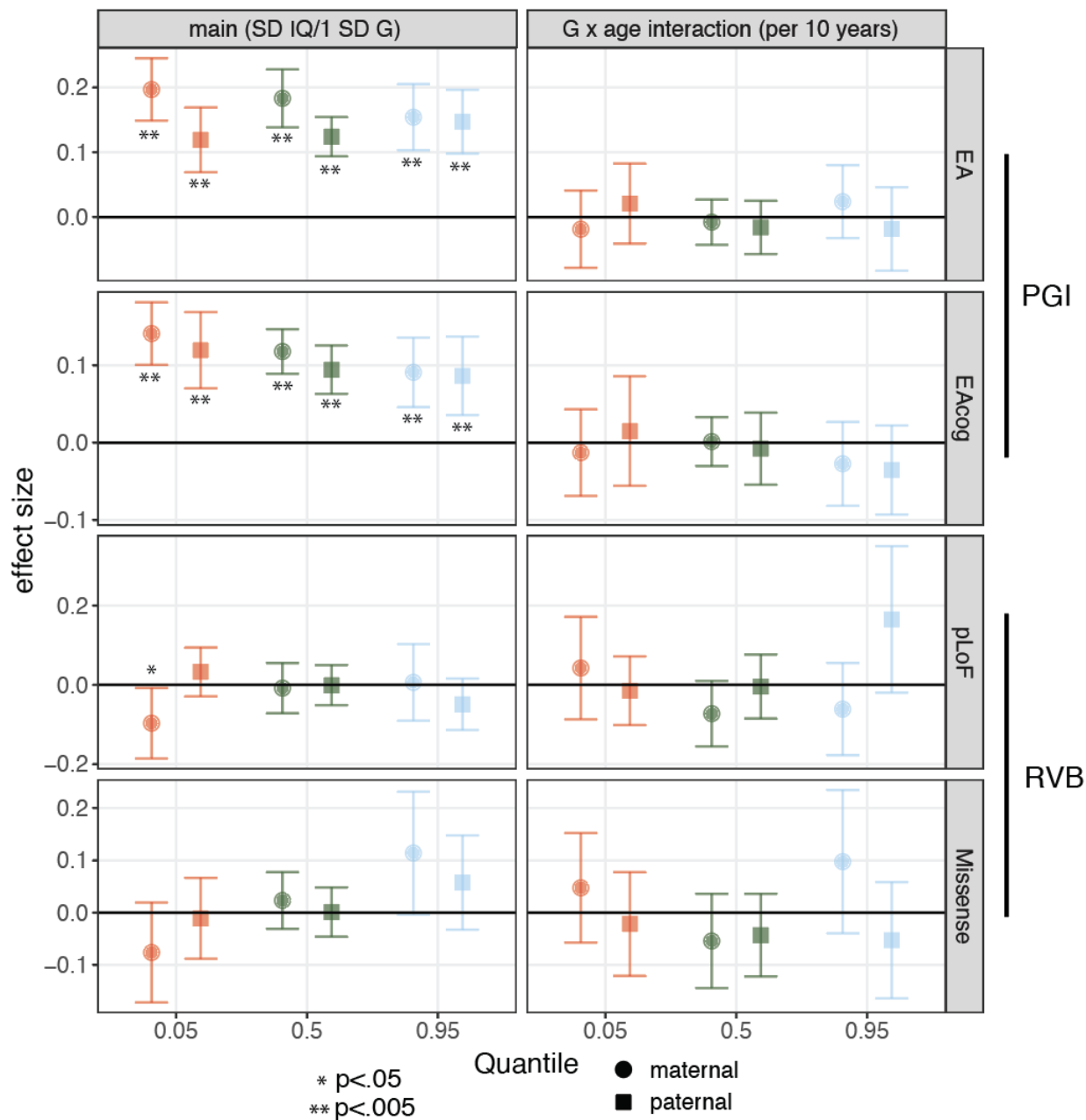
1271

1272

1273

1274

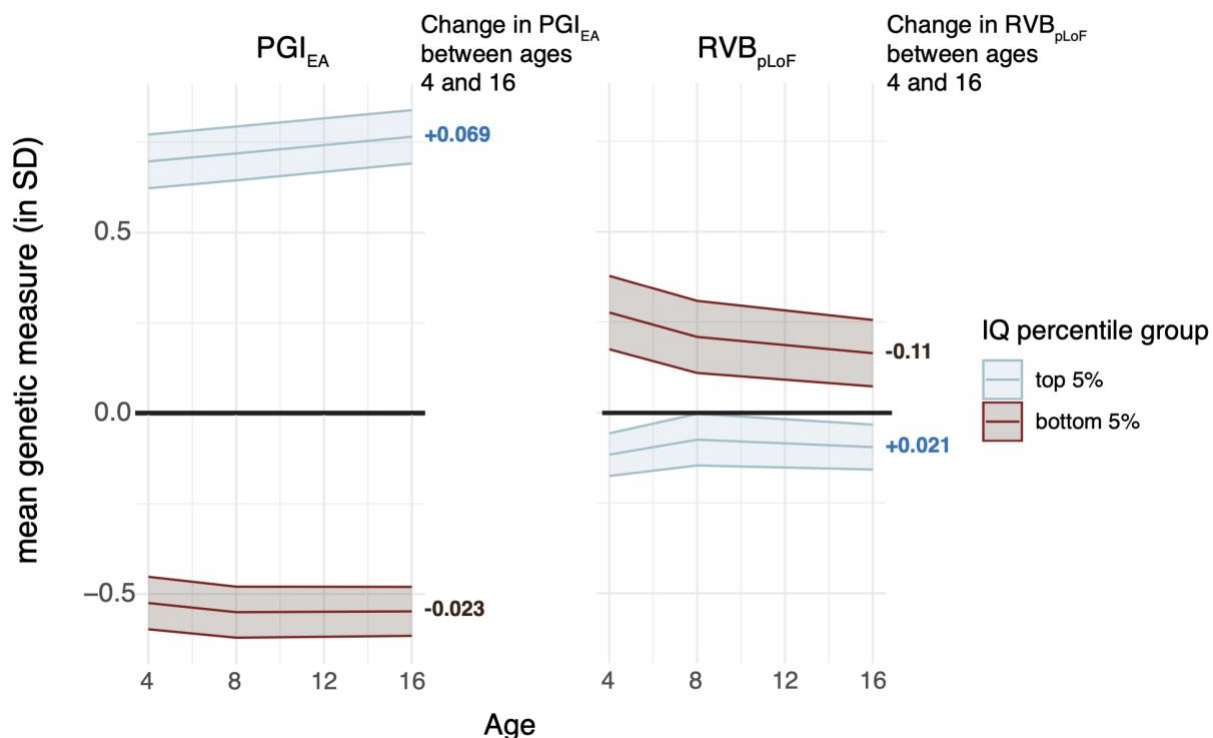
1275 Extended Data Figure 8



1276
 1277 Influence of parental common and rare variants on different quantiles of the IQ distribution post-imputation.
 1278 Standardized effects and 95% confidence intervals for quantile regression of the 5th, 50th and 95th
 1279 percentiles estimated from mixed-effects modeling with post-imputation IQ at ages 4, 8, and 16 for parental
 1280 PGI_{EA}, PGI_{Cog}, RVB_{pLoF} and RVB_{Missense} estimated in a trio model.

1281
 1282
 1283

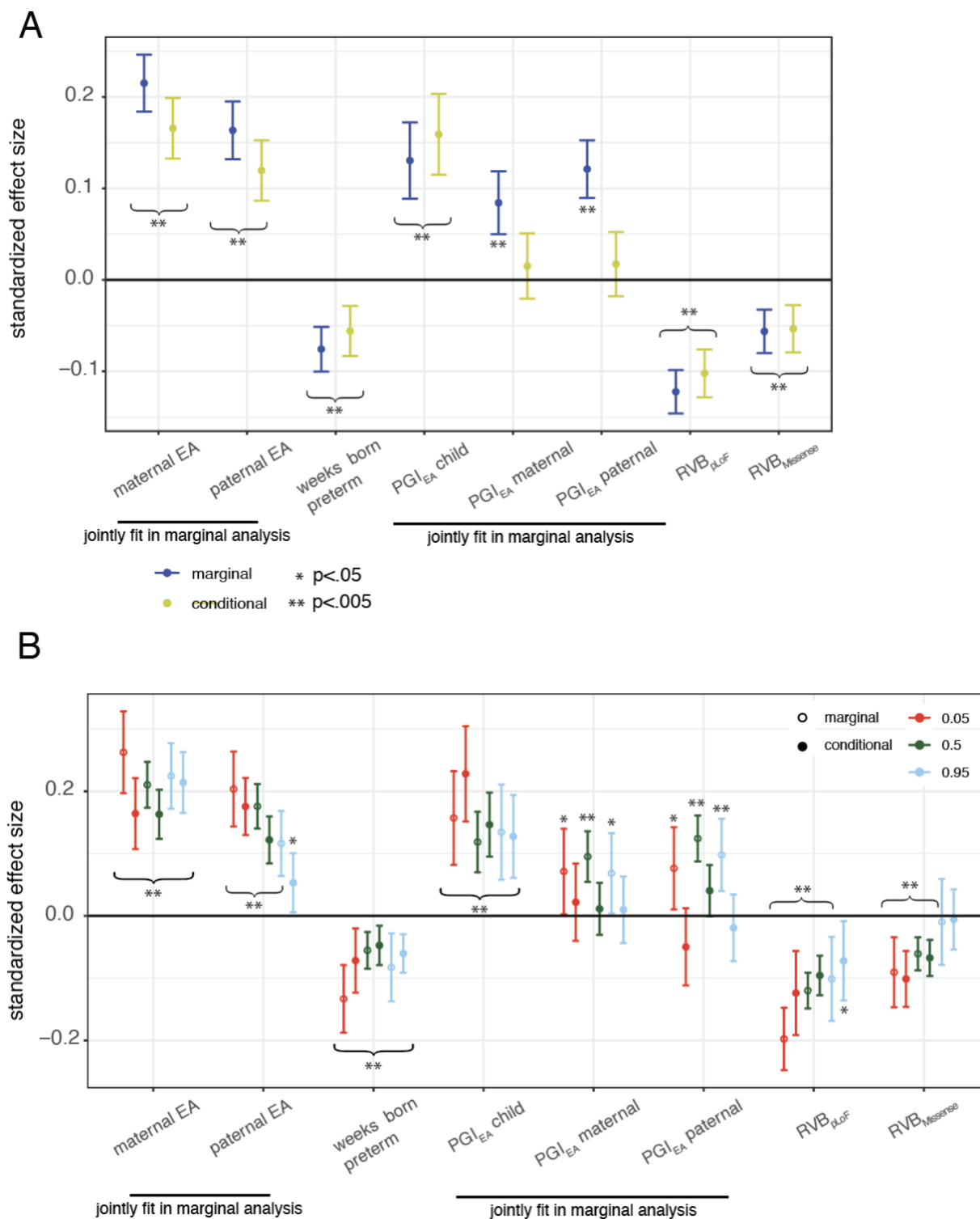
1284 Extended Data Figure 9



1285 Summary of genetic trajectories across development, as inferred in this work. Mean PGI_{EA} (left) and RVB_{pLoF}
1286 (right) for individuals in the top (light blue) or bottom (brown) 5% percentile of IQ at a given age (inferred
1287 post-imputation of IQ). Note that the individuals in the top and bottom 5th percentile groups vary across
1288 development. Children at the top 5th percentile of IQ at age 16 have a higher mean PGI_{EA} than those at the
1289 top of the distribution at age 4, while those at the bottom 5th percentile have a relatively steady mean PGI_{EA}.
1290 In contrast, the quantile-specific age interactions for RVB_{pLoF} suggest that children in the bottom 5th
1291 percentile of IQ at age 4 have a much higher burden of damaging rare variants than those at age 16, while
1292 those at the top 5th percentile have a stable and low rare variant burden.
1293

1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309

1310 Extended Data Figure 10



1311
 1312 Associations between genetic and other factors and cognitive performance across ages in MCS. A)
 1313 Standardized effects and 95% confidence intervals from obtained from two linear models, as indicated in
 1314 the key: marginal associations from models in which only the indicated variable(s) were included in the
 1315 model (in addition to standard covariates), and conditional associations from models in which all the

1316 variables were included in the same model. B) As in (A) using a quantile regression model for the quantiles
1317 as indicated in the figure. Asterisks indicate whether the estimate is significantly different from 0; note that
1318 curly brackets with asterisks indicate that the estimates spanned by the bracket are all significant.