NUTRITION AND DIETETICS

Indicators of Intelligence and Genetic Predisposition

Cognitive Development, Intelligence and Genetic Architecture

Department of Nutrition and Dietetics

Harokopio University of Athens

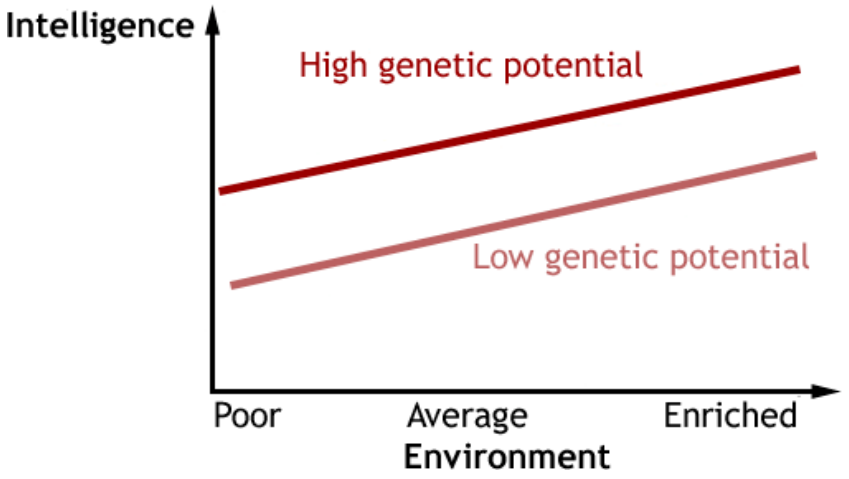
Prof. Georgios Dedoussis

Genetic Architecture

Rare vs Common Effects

Gene-Environment Interaction


TIME
THE FUNNIEST MAN IN AMERICA
What scientists have uncovered about HOW MEMORY WORKS and how to improve it
THE I.Q. GENE?


Intelligence
High genetic potential
Low genetic potential
Poor
Average
Enriched
Environment

Investigating the effects of copy number variants on reading and language performance

Alessandro Gialluisi[1,2], Alessia Visconti[3], Erik G. Willcutt[4,5], Shelley D. Smith[6], Bruce F. Pennington[7], Mario Falchi[3], John C. DeFries[4,5], Richard K. Olson[4,5], Clyde Francks[1,8*] and Simon E. Fisher[1,8*]

In a dataset of children recruited for a history of reading disability (RD, also known as **dyslexia**) or attention deficit hyperactivity disorder (**ADHD**) and their siblings, they investigated the effects of CNVs on **reading and language performance**.

Reading disability (RD or developmental dyslexia) and specific language impairment (SLI) are two of the most prevalent neurodevelopmental disorders, with a prevalence of ≈**5–8 %** among school-aged children

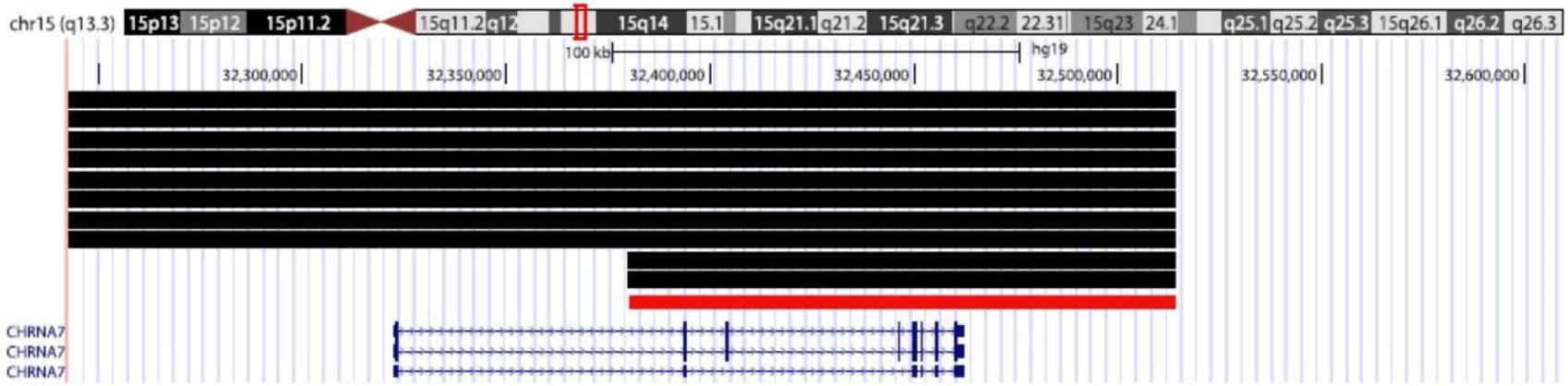


Fig. 2 Candidate region of CNV overlap associated with PC1 in the GWAS with CNV state. The *red line* indicates the associated interval chr15:32,380,064-32,514,341, partially overlapping *CHRNA7* (15q13.3). *Black horizontal lines* represent the individual CNV calls detected in this region (nine heterozygous duplications and one heterozygous deletion; see Additional file 2: Table S2e for details)

Nicotinic cholinergic receptors are ligand-gated ion channels that mediate fast signal transmission at synapses and are ubiquitously expressed in the CNS. Several studies have suggested a possible involvement of **CHRNA7** in language skills.

The 15q13.3 region is also a hotspot of neuropsychiatric CNVs, which have been implicated in several disorders including **SCZ, ASD, ADHD, and epilepsy**

Gialluisi *et al. Journal of Neurodevelopmental Disorders* (2016) 8:17

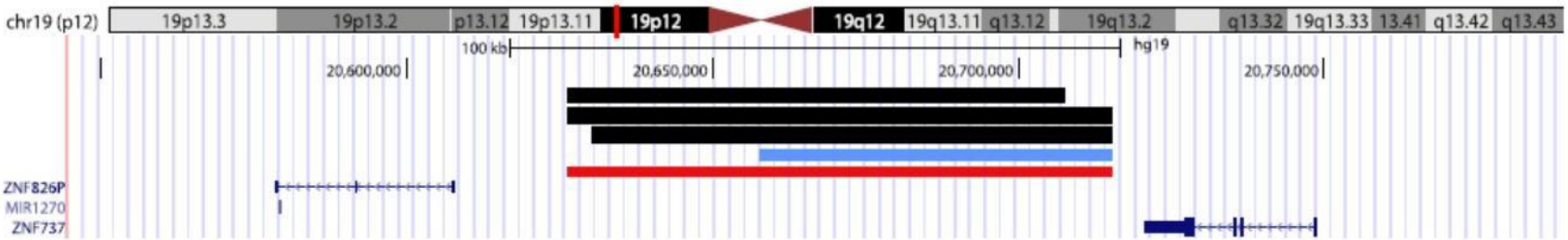


Fig. 3 19p12 region associated with PC traits in the GWAS with probe intensity data. The *blue line* indicates the interval of association with probe intensity data (chr19:20,657,781-20,715,228), while the *red line* indicates the overlapping region of association between CNV state and PC scores (chr19:20,626,179-20,715,228). *Black horizontal lines* represent the three types of deletions detected in this region (reported in Additional file 2: Table S2g)

Zinc finger protein 737 has not been functionally characterized, but the presence of a zinc finger domain suggests a possible involvement in transcriptional regulation. Interestingly, a microdeletion within another zinc finger gene, ZNF277, has been suggested as susceptibility CNV for SLI

Several of the candidate genes are known to have roles in important processes in **central nervous system (CNS) development**, such as neuronal migration, axonal guidance, and neurite outgrowth

Conclusions

These data suggest that CNVs **do not** underlie a substantial proportion of variance in reading and language skills.

Gialluisi *et al. Journal of Neurodevelopmental Disorders* (2016) 8:17

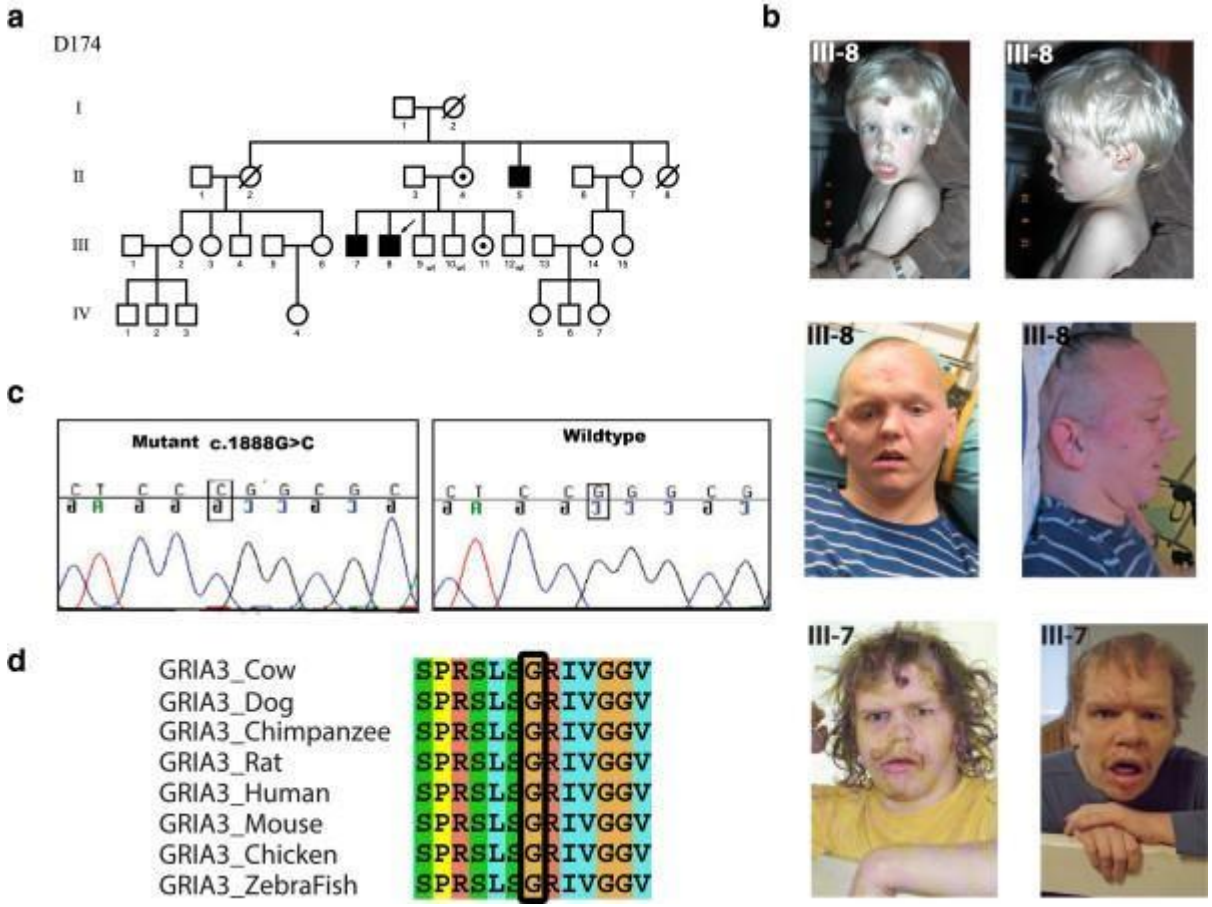
frontiers in Molecular Neuroscience

ORIGINAL RESEARCH
published: 11 January 2016
doi: 10.3389/fnmol.2015.00085

Novel Missense Mutation A789V in *IQSEC2* Underlies X-Linked Intellectual Disability in the MRX78 Family

Vera M. Kalscheuer[1,2], Victoria M. James[3], Miranda L. Himelright[4], Philip Long[3], Renske Oegema[5], Corinna Jensen[1], Melanie Bienek[1], Hao Hu[1], Stefan A. Haas[6], Maya Topf[7], A. Jeannette M. Hoogeboom[5], Kirsten Harvey[3], Randall Walikonis[4] and Robert J. Harvey[3]**

Intellectual disability (ID) is a developmental brain disorder characterized by impaired intellectual and adaptive functions, and can be defined by an **IQ below 70** and limitations in intellectual functioning and adaptive behaviours. As a result of the excess in males affected by ID, and the identification of families where ID shows **clear X-linked** segregation, significant attention has focused on the genetics of X-linked intellectual disability (XLID)—a common, clinically and genetically complex disorder often arising from mutations in one **of >100 genes** on the X chromosome.


a
D174
I
II
III
IV
b
III-8
III-8
III-8
III-8
III-7
III-7
c
Mutant c.1888G>C
Wildtype
d
GRIA3_Cow
GRIA3_Dog
GRIA3_Chimpanzee
GRIA3_Rat
GRIA3_Human
GRIA3_Mouse
GRIA3_Chicken
GRIA3_ZebraFish
SPRSLSGRIVGGV

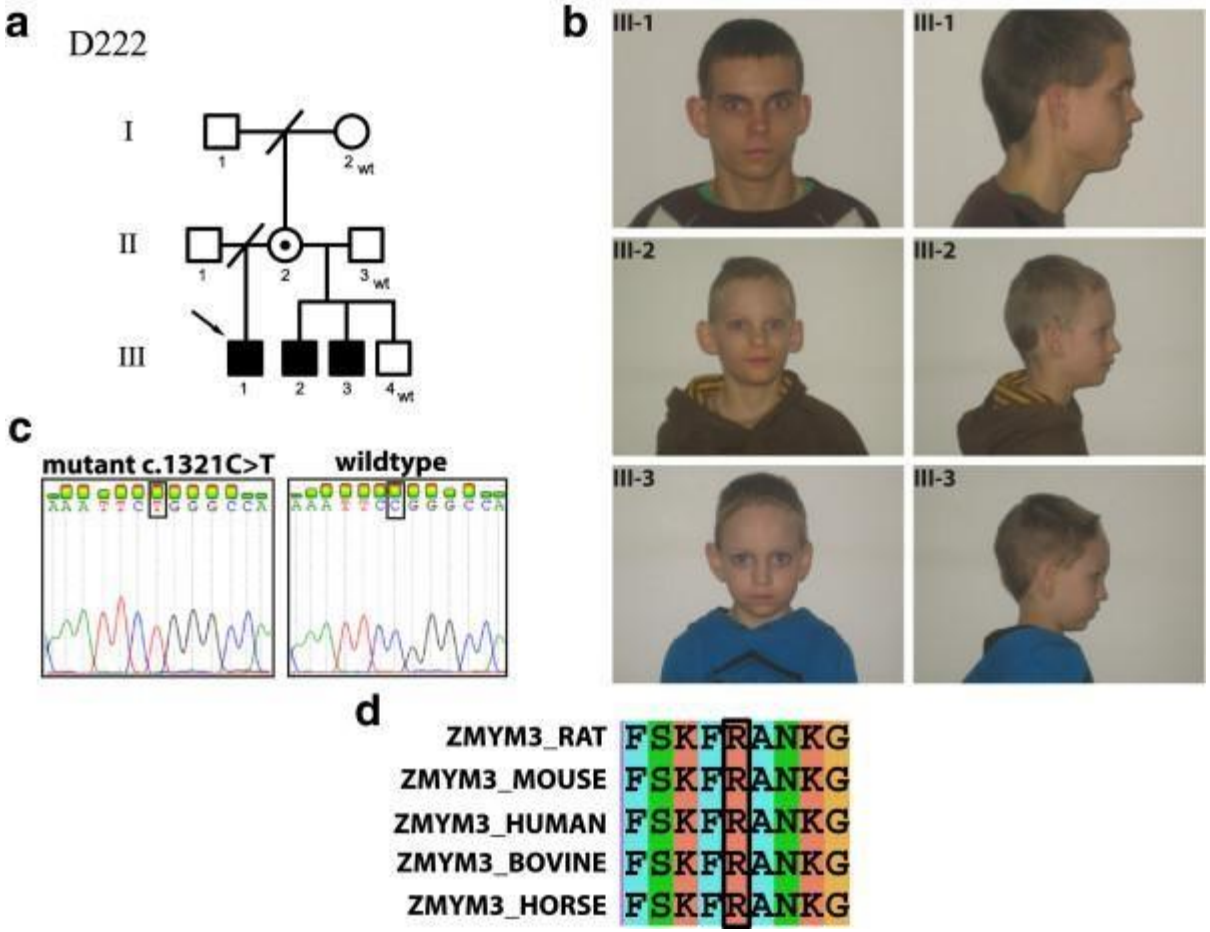

a
D222
I
II
III
1
2 wt
1
2
3 wt
1
2
3
4 wt
b
III-1
III-1
III-2
III-2
III-3
III-3
c
mutant c.1321C>T
wildtype
d
ZMYM3_RAT FSKFRANKG
ZMYM3_MOUSE FSKFRANKG
ZMYM3_HUMAN FSKFRANKG
ZMYM3_BOVINE FSKFRANKG
ZMYM3_HORSE FSKFRANKG

TABLE 1 | Summary of clinical features present in affected males and females in the MRX78 family.

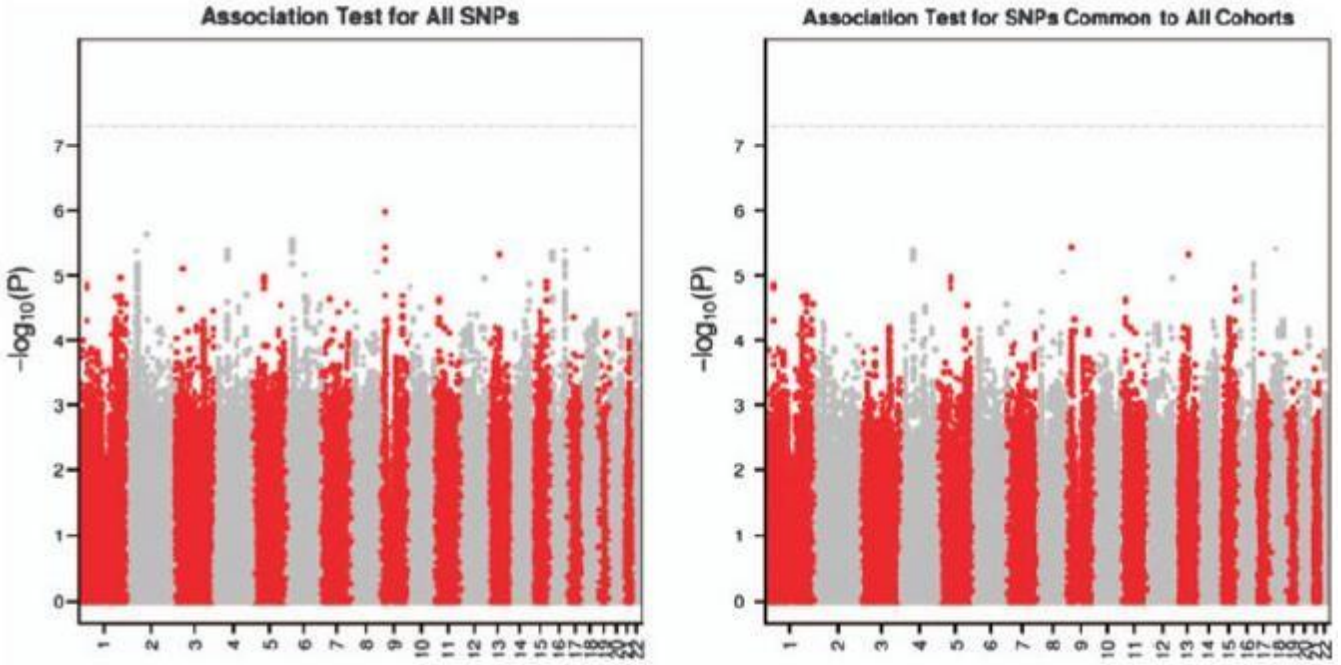
Subject	Gender	Intellectual disability	Epilepsy	Behavioral problems
I-2	F	Mild, illiterate	–	–
II-2	F	Mild	–	–
II-6	F	Present	–	–
II-7	M	Learning difficulties*	–	–
II-8	F	Mild	–	–
II-9	M	Severe, does not speak	Present	Does not interact socially
III-1	M	Severe, does not speak	One seizure as a teenager	Aggressive against others
III-2	M	Severe, does not speak	–	Aggressive against others
III-3	M	Moderate	–	–
III-4	F	Learning difficulties	–	–
III-5	M	Moderate-severe	Diffuse encephalopathy without epileptic discharge	Occasional verbal aggression, suspected of ASD but not formally tested
III-6	F	Learning difficulties	–	–
III-8	M	Severe	–	Pervasive developmental disorder, ASD, severe aggressive outbursts against objects
IV-1	F	Learning difficulties	–	–

Although originally described as non- syndromic XLID, newly acquired clinical data from the MRX78 family suggests that additional features might be associated with the IQSEC2p.A789V mutation ,including variable seizures in males, which is consistent with other reports, and behavioral disturbances in five out of six affected males. By contrast, heterozygous female have learning disabilities. As IQSEC2 is one of the few genes that escapeX-inactivation in females with an expression level similar in males and females

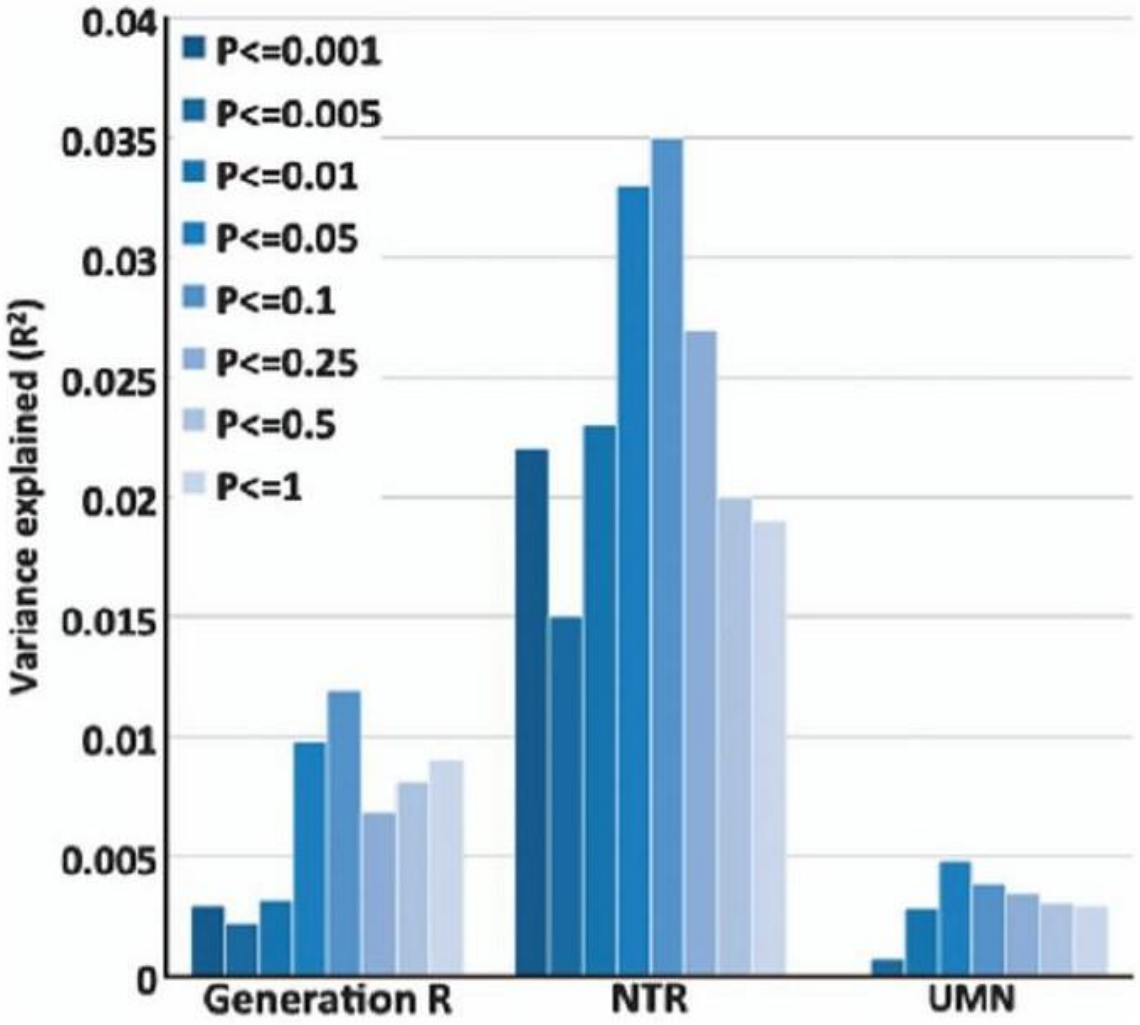
Mol Psychiatry. 2014 February ; 19(2): 253–258. doi:10.1038/mp.2012.184.

Childhood intelligence is heritable, highly polygenic and associated with *FNBP1L*

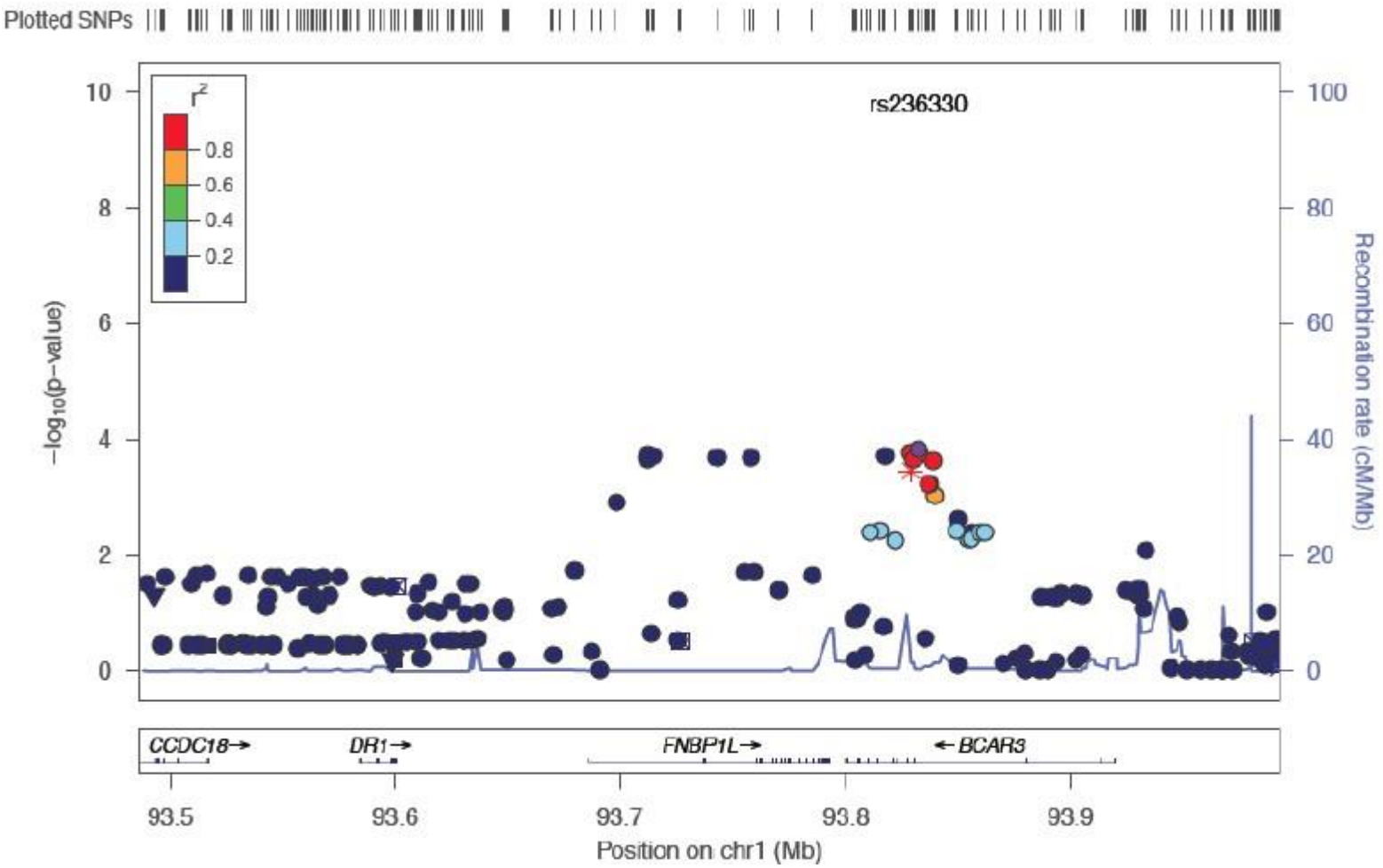
B Benyamin[1,2,26], **BSt Pourcain**[3,26], **OS Davis**[4,26], **G Davies**[5,26], **NK Hansell**[2], **M-JA Brion**[3,6], **RM Kirkpatrick**[7], **RAM Cents**[8,9], **S Franić**[10], **MB Miller**[7], **CMA Haworth**[4], **E Meaburn**[11], **TS Price**[4], **DM Evans**[3], **N Timpson**[3], **J Kemp**[3], **S Ring**[3], **W McArdle**[3], **SE Medland**[2], **J Yang**[12], **SE Harris**[13,14], **DC Liewald**[5,14], **P Scheet**[10], **X Xiao**[15], **JJ Hudziak**[16], **EJC de Geus**[10], **Wellcome Trust Case Control Consortium 2 (WTCCC2)**, **VWV Jaddoe**[8,17,18], **JM Starr**[14,19], **FC Verhulst**[9], **C Pennell**[6], **H Tiemeier**[9,17,20], **WG Iacono**[7], **LJ Palmer**[21,22], **GW Montgomery**[2], **NG Martin**[2], **DI Boomsma**[10], **D Posthuma**[9,23,24], **M McGue**[7,25], **MJ Wright**[2], **G Davey Smith**[3,27], **IJ Deary**[5,14,27], **R Plomin**[4,27], and **PM Visscher**[1,2,12,14,27]



Cohort	N	h^2 *(s.e.)*	P-*value*
ALSPAC	5517	0.46 (0.06)	3.9×10^{-15}
TEDS	2794	0.22 (0.10)	0.014
UMN	1736	0.40 (0.21)	0.028



Supplementary Figure 6. Association P-values around *FNBP1L*, the most significantly associated gene from gene-based analysis



OPEN

Molecular Psychiatry (2016) 21, 1145–1151

www.nature.com/mp

ORIGINAL ARTICLE

A genome-wide analysis of putative functional and exonic variation associated with extremely high intelligence

SL Spain[1,8], I Pedroso[1,8], N Kadeva[1], MB Miller[2], WG Iacono[2], M McGue[2], E Stergiakouli[3], GD Smith[3], M Putallaz[4], D Lubinski[5], EL Meaburn[6], R Plomin[7] and MA Simpson[1]

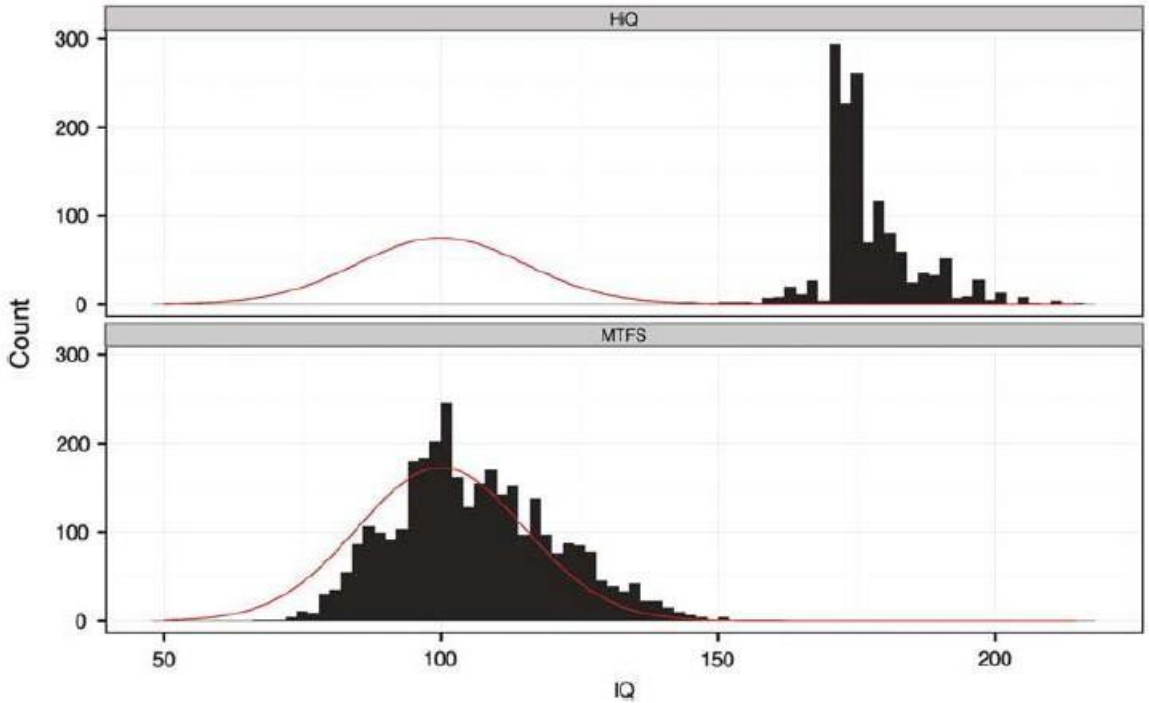
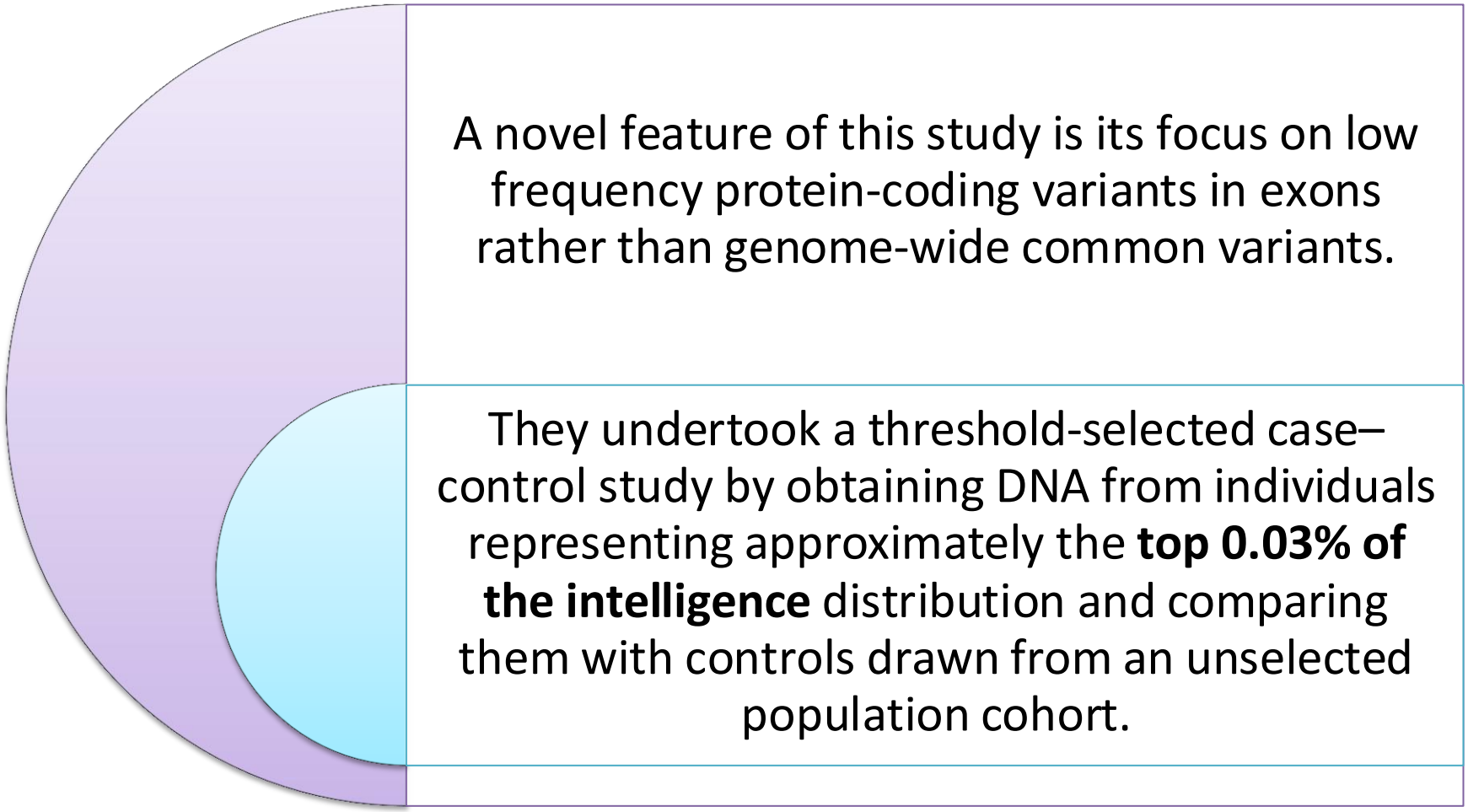
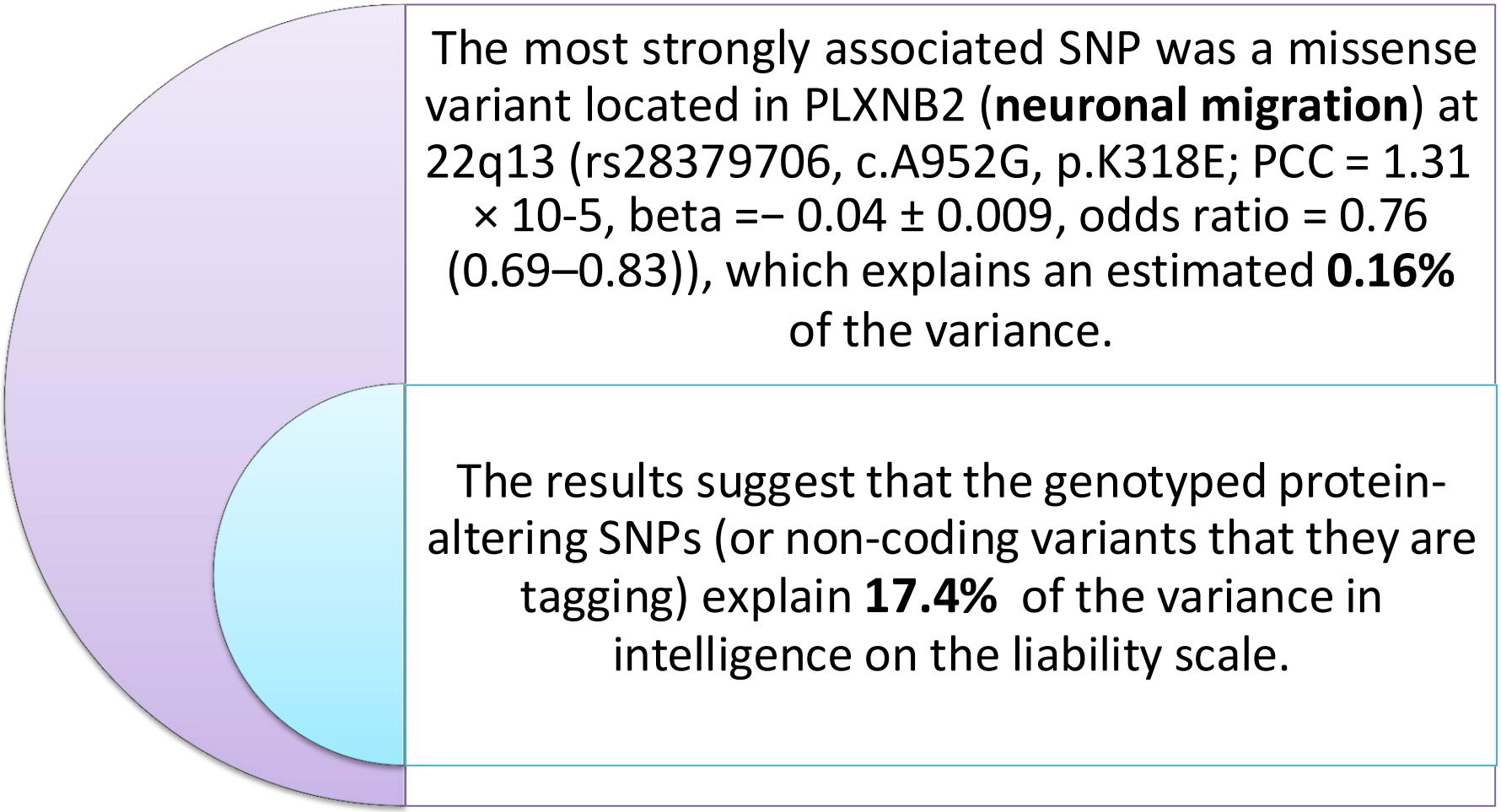


Figure 1. Extreme sampling of the high cognitive ability cohort. Distribution of general cognitive ability, measured on a standardised IQ scale (mean 100, s.d. 15), of the high cognitive ability cohort (HiQ, $n = 1409$, top panel) and the control cohort (MTFS, $n = 3253$, lower panel). The red lines illustrate the expected distribution of IQ scores with unbiased sampling.


A novel feature of this study is its focus on low frequency protein-coding variants in exons rather than genome-wide common variants.
They undertook a threshold-selected case–control study by obtaining DNA from individuals representing approximately the **top 0.03% of the intelligence** distribution and comparing them with controls drawn from an unselected population cohort.

High-intelligence cases (HiQ). Individuals were recruited from the Duke University Talent Identification Program (TIP), a non-profit organisation established in 1980 and dedicated to identifying and fostering the development of academically gifted children35 (http://www.tip.duke.edu)


The most strongly associated SNP was a missense variant located in PLXNB2 (**neuronal migration**) at 22q13 (rs28379706, c.A952G, p.K318E; PCC = 1.31 × 10-5, beta =– 0.04 ± 0.009, odds ratio = 0.76 (0.69–0.83)), which explains an estimated **0.16%** of the variance.
The results suggest that the genotyped protein-altering SNPs (or non-coding variants that they are tagging) explain **17.4%** of the variance in intelligence on the liability scale.

Definition

Educational attainment refers to the highest level of schooling that a person has reached. At the primary and secondary school level, **educational attainment** refers to the number of grades completed. At the postsecondary level, it refers to institutions attended and certificates, degrees or diplomas obtained.

Educational attainment is the visible output of education systems and a measure of their success. For individuals, achievement levels have a major impact, both personally and professionally, on the quality of life and job opportunities.

International Standard Classification of Education (ISCED)

The International Standard Classification of Education (ISCED) was developed by **UNESCO** to facilitate comparisons of education statistics and indicators across countries on the basis of uniform and internationally agreed definitions.

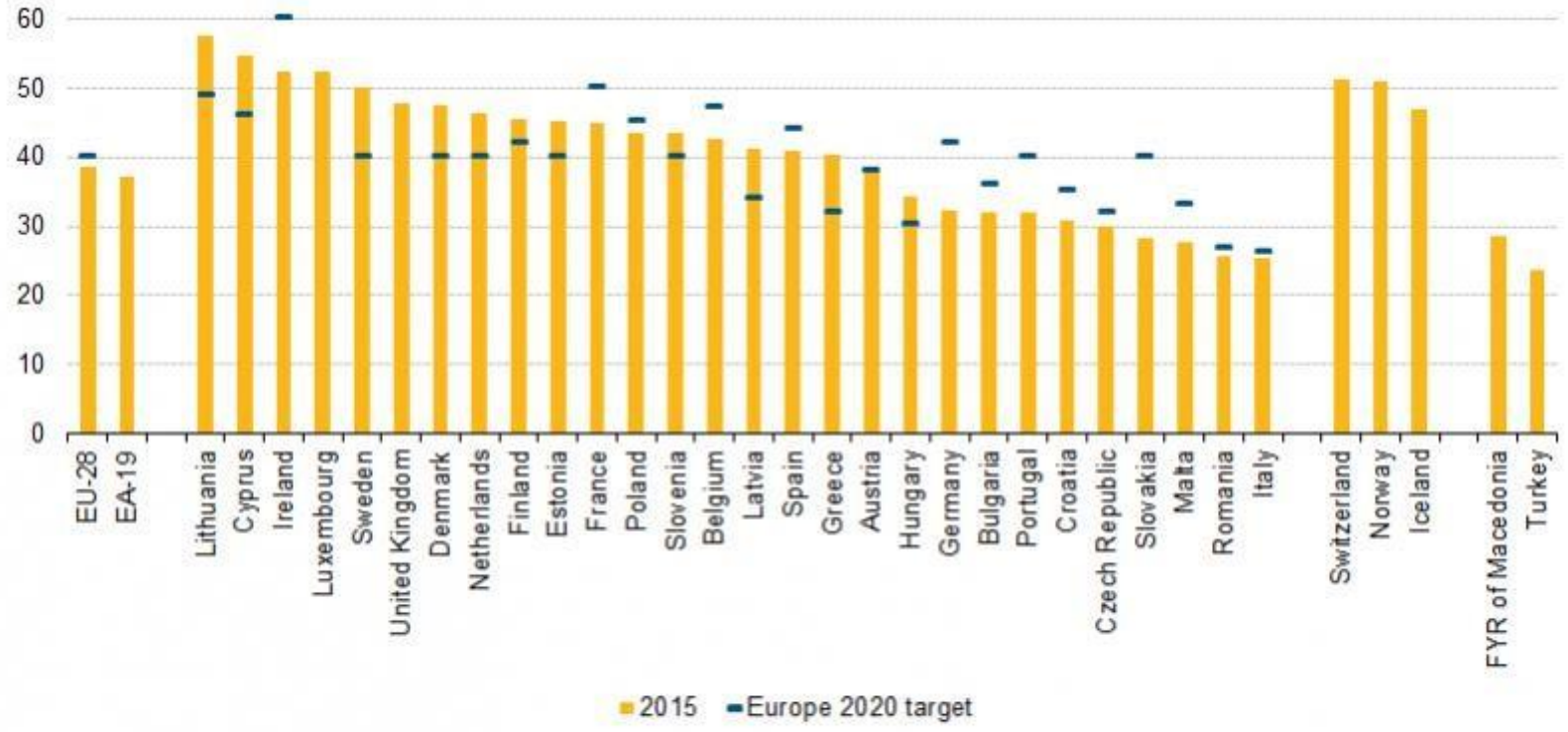
ISCED levels

ISCED 1997	ISCED 2011
0 Pre-primary (designed for children aged 3 years and above)	0 Early childhood education 01 Early childhood educational development (designed for children aged under 3 years) 02 Pre-primary (designed for children aged 3 years and above)
1 Primary (or 1st stage of basic education)	1 Primary
2 Lower secondary (or 2nd stage of basic education)	2 Lower secondary
3 Upper secondary	3 Upper secondary
4 Post-secondary non-tertiary	4 Post-secondary non-tertiary
5 First stage of tertiary	5 Short cycle tertiary 6 Bachelor's or equivalent level 7 Master's or equivalent level
6 Second stage of tertiary	8 Doctoral or equivalent level

Educational attainment levels of the population have improved significantly over the last thirty years. In 2015, 79.1 % of people aged 25–54 in the EU-28 had at least attained an upper secondary level of education, compared with 62.6 % of those aged 55–74. Those with tertiary educational attainment amounted to 32.6 % of those aged 25–54 and 20 % of those aged 55–74

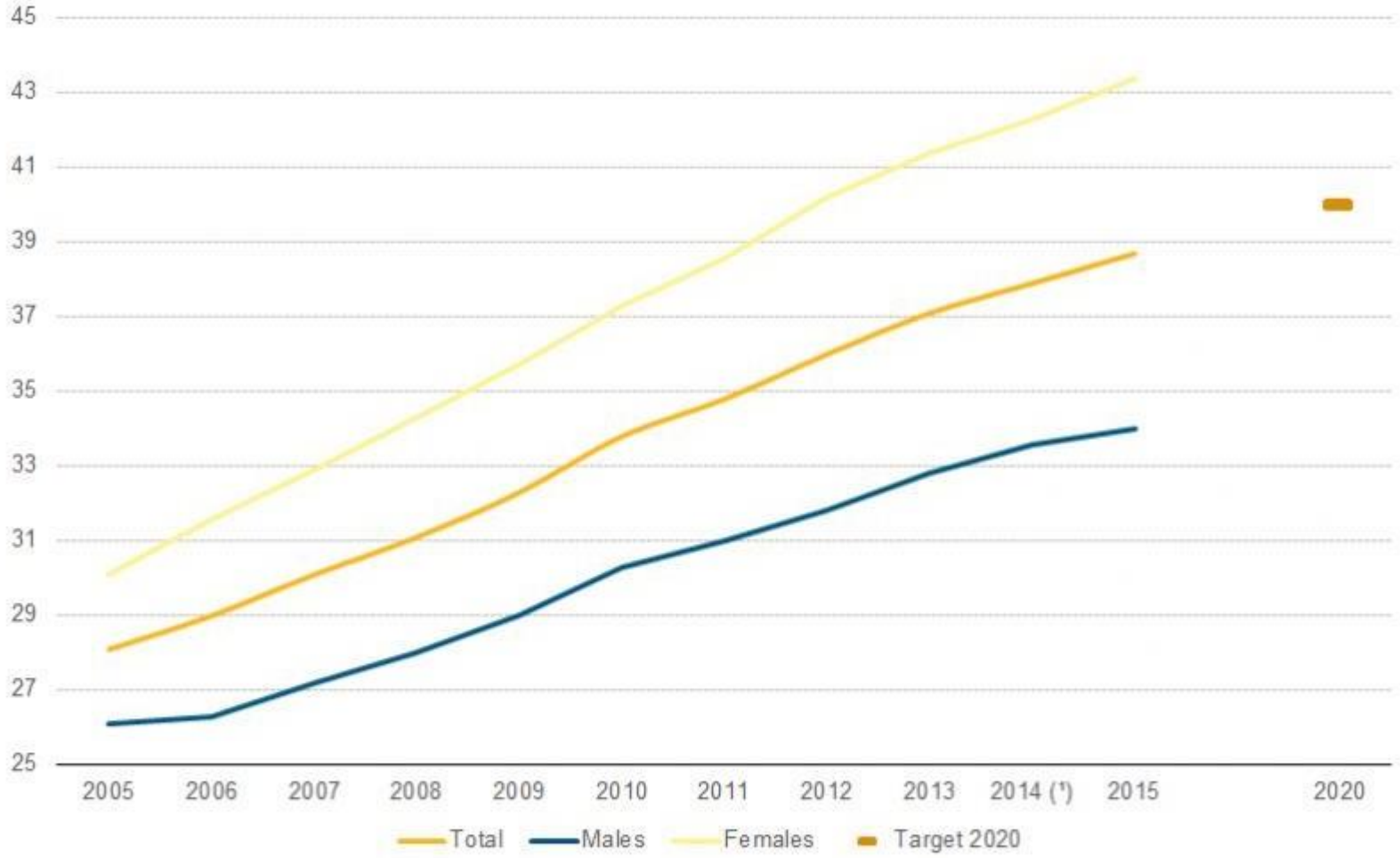
	25–54 years			55–74 years		
	Low (ISCED 0–2)	Medium (ISCED 3–4)	High (ISCED 5–8)	Low (ISCED 0–2)	Medium (ISCED 3–4)	High (ISCED 5–8)
EU-28	20.8	46.5	32.6	37.4	42.6	20.0
EA-19	23.8	44.6	31.6	41.8	38.5	19.7
Belgium	20.9	39.1	40.0	44.4	31.0	24.7
Bulgaria	16.9	54.2	28.9	27.8	51.4	20.8
Czech Republic	5.2	70.3	24.5	13.7	72.9	13.4
Denmark	17.1	43.1	39.8	30.6	43.1	26.3
Germany	12.9	58.7	28.3	16.2	59.3	24.5
Estonia	9.2	51.9	38.9	12.1	52.5	35.3
Ireland	15.6	37.6	46.8	45.9	30.7	23.3
Greece	24.2	44.2	31.7	56.8	26.4	16.8
Spain	38.1	23.5	38.4	66.7	14.0	19.3
France	18.3	43.7	38.0	39.8	39.5	20.6
Croatia	12.5	62.5	25.0	33.1	49.7	17.2
Italy	36.3	44.6	19.1	61.1	28.5	10.4
Cyprus	17.4	37.8	44.8	47.5	31.8	20.6
Latvia	10.3	55.9	33.8	13.9	63.3	22.8
Lithuania	6.8	51.5	41.7	13.8	60.3	25.9
Luxembourg	21.9	33.4	44.7	34.3	39.8	25.9
Hungary	14.9	58.6	26.5	28.1	55.2	16.7
Malta	50.4	26.6	23.0	77.8	13.8	8.5
Netherlands	19.9	42.1	38.1	40.0	35.5	24.5
Austria	13.1	53.9	33.1	26.2	53.7	20.1
Poland	7.2	60.0	32.7	20.3	65.8	13.9
Portugal	48.4	25.6	26.0	80.7	8.8	10.5
Romania	21.9	58.5	19.6	45.3	46.6	8.1
Slovenia	9.9	56.2	33.9	25.1	56.5	18.4
Slovakia	6.9	69.6	23.5	17.3	69.3	13.4
Finland	9.6	45.4	45.0	27.8	39.0	33.2
Sweden	13.1	44.2	42.7	29.0	42.6	28.4
United Kingdom	18.5	37.7	43.8	29.4	38.7	31.9
Iceland	22.9	35.1	42.0	37.1	38.0	24.9
Norway	16.8	37.2	46.1	20.6	49.5	29.9
Switzerland	10.8	44.8	44.4	17.2	53.5	29.3
FYR of Macedonia	31.0	48.8	20.2	47.4	38.1	14.5
Turkey	62.7	19.1	18.2	84.3	8.2	7.5

http://ec.europa.eu/eurostat/statistics-explained/index.php/Main_Page



(1) In the cases where the national target has been set within a range between two possible values, the lower level has been taken. The United Kingdom did not set a specific Europe 2020 target.
(2)The national target for Luxembourg is 66%.

Population aged 30-34 with tertiary educational attainment (ISCED 5-8), by sex, EU-28, 2005-15

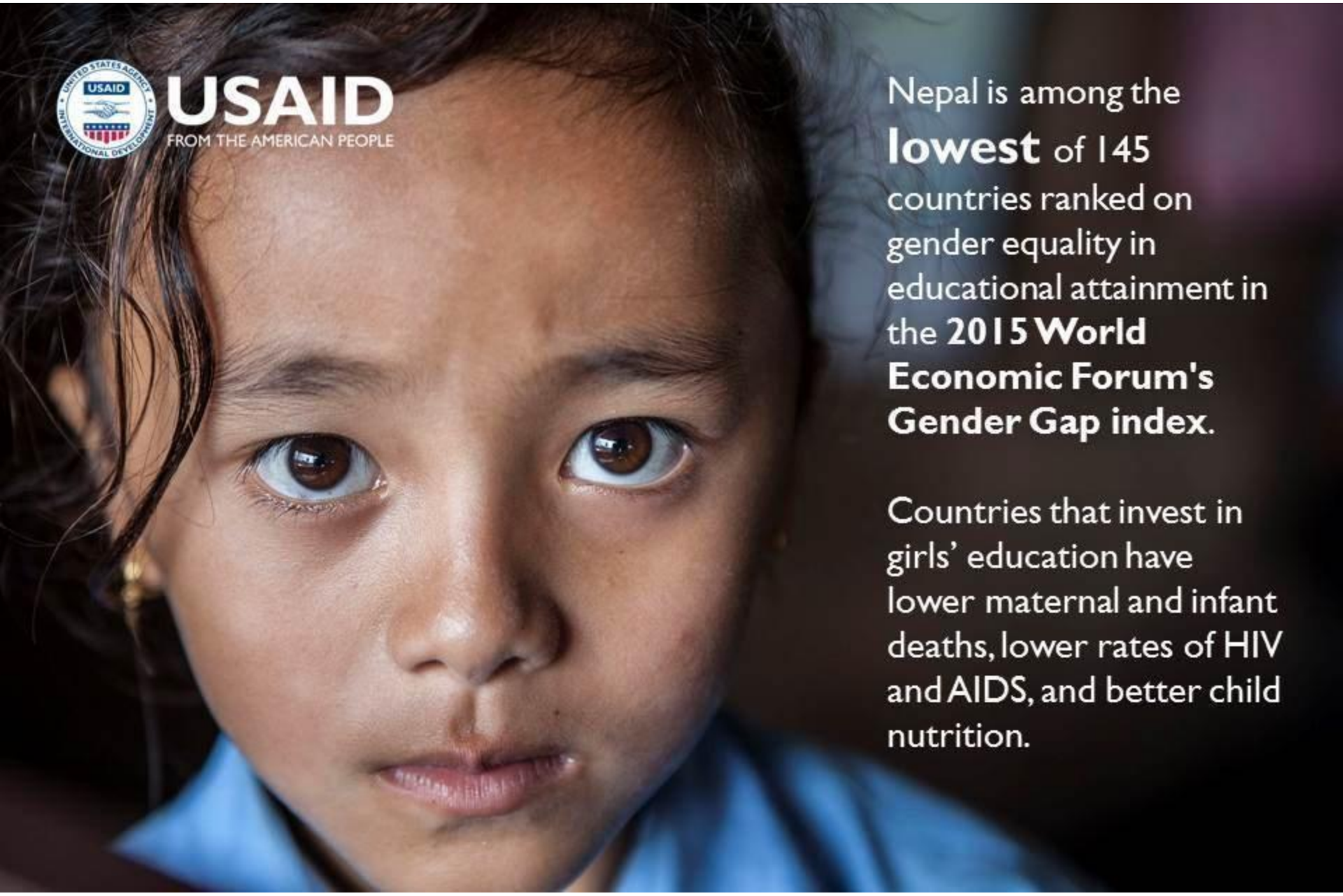


(1) Break in the series.

The first 9 months can shape the rest of your life…….

Critical periods before and during pregnancy when specific nutrients are needed for optimal development.

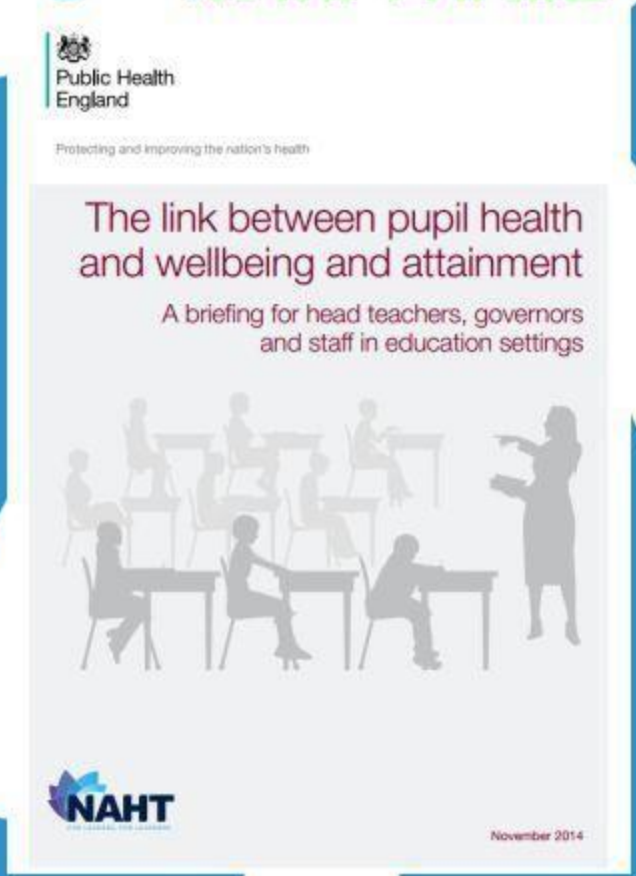
Risks of a number of **chronic diseases** in adulthood such as hypertension, diabetes, heart disease may have their **origins before birth**.


USAID
FROM THE AMERICAN PEOPLE
Nepal is among the **lowest** of 145 countries ranked on gender equality in educational attainment in the **2015 World Economic Forum's Gender Gap index**.
Countries that invest in girls' education have lower maternal and infant deaths, lower rates of HIV and AIDS, and better child nutrition.

The link between pupil health and wellbeing and attainment

Research evidence shows that education and health are closely linked.

So promoting the health and wellbeing of pupils and students within schools and colleges has the potential to improve their educational outcomes and their health and wellbeing outcomes.



https://www.g nment/publications/the-link-betw alth-and-wellbeing-and-attainm

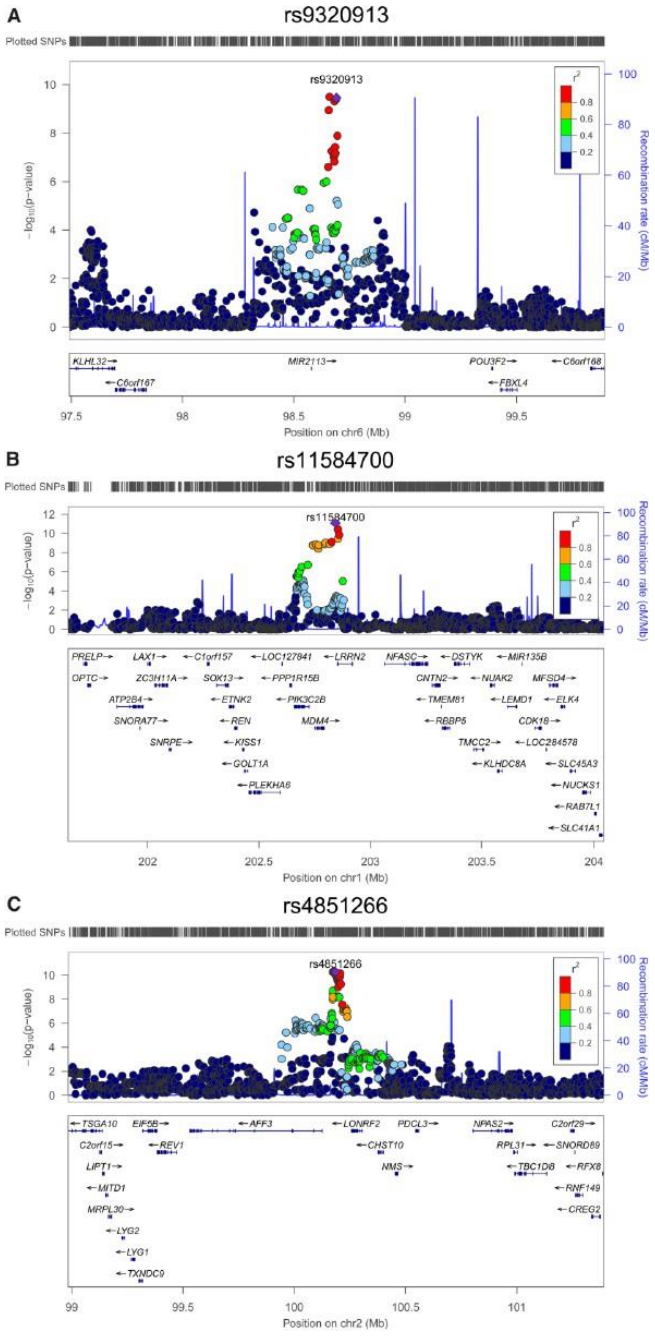
- Estimates suggest that around **40%** of the variance in educational attainment is explained by genetic factors. Furthermore, educational attainment is moderately correlated with other heritable characteristics, including cognitive function and personality traits related to persistence and self-discipline

GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment

To create a harmonized measure of educational attainment, we coded study-specific measures using the International Standard Classification of Education (**ISCED 1997**) scale. We analyzed a quantitative variable defined as an individual's years of schooling (**EduYears**) and a binary variable for college completion (**College**). College may be more comparable across countries, whereas EduYears contains more information about individual differences within countries.

Table 1. The results of the GWAS meta-analysis for the independent signals reaching $p < 10^{-6}$ in the discovery stage. The rows in bold are the independent signals reaching $p < 5 \times 10^{-8}$ in the discovery stage. "Frequency" refers to allele-frequency in the combined-stage meta-analysis. "Beta/OR" refers to the effect size in the *EduYears* analysis and to the Odds Ratio in the *College* analysis. All *p*-values are from the sample-size-weighted meta-analysis (fixed effects). The *p*-value in the replication stage meta-analysis was calculated from a one-sided test. I^2 represents the % heterogeneity of effect size between the discovery stage studies. p_{het} is the heterogeneity *p*-value.

						Discovery stage				Replication stage		Combined stage			Combined stage – sex-specific			
SNP	**Chr**	**Position (bp)**	**Nearest gene**	**Effective allele**	**Frequency**	**Beta/OR**	***P*-value**	I^2	P_{het}	**Beta/OR**	***P*-value**	**Beta/OR**	***P*-value**	P_{het}	**Beta/OR (Males)**	***P*-value (Males)**	**Beta/OR (Females)**	***P*-value (Females)**
									EduYears									
rs9320913	**6**	**98691454**	**LOC100129158**	**A**	**0.483**	**0.106**	**4.19×10^{-9}**	**18.3**	**0.097**	**0.077**	**0.012**	**0.101**	**$3.50\times10^{-}$**	**0.350**	**0.095**	**1.87×10^{-4}**	**0.100**	**1.43×10^{-6}**
rs3783006	13	97909210	STK24	C	0.454	0.096	2.29×10^{-7}	0	0.982	0.056	0.055	0.088	$8.45\times10^{-}$	0.959	0.064	1.44×10^{-2}	0.108	3.35×10^{-7}
rs8049439	16	28745016	ATXN2L	T	0.581	0.090	7.12×10^{-7}	10.7	0.229	0.065	0.026	0.086	$1.15\times10^{-}$	0.205	0.097	1.43×10^{-4}	0.078	1.90×10^{-4}
rs13188378	5	101958587	SLCO6A1	A	0.878	-0.136	7.49×10^{-7}	0	0.791	0.091	0.914	-0.097	$1.37\times10^{-}$	0.646	-0.134	8.21×10^{-3}	-0.080	5.92×10^{-3}
									College									
rs11584700	**1**	**202843606**	**LRRN2**	**A**	**0.780**	**0.921**	**2.07×10^{-9}**	**13.8**	**0.179**	**0.912**	**4.86×10^{-4}**	**0.919**	**$8.24\times10^{-}$**	**0.221**	**0.934**	**6.11×10^{-4}**	**0.911**	**2.12×10^{-9}**
rs4851266	**2**	**100184911**	**LOC150577**	**T**	**0.396**	**1.050**	**2.20×10^{-9}**	**23.7**	**0.049**	**1.049**	**0.003**	**1.050**	**$5.33\times10^{-}$**	**0.072**	**1.054**	**1.55×10^{-5}**	**1.052**	**6.74×10^{-8}**
rs2054125	2	199093966	PLCL1	T	0.064	1.468	5.55×10^{-8}	7	0.325	1.098	0.225	1.376	$2.12\times10^{-}$	0.268	1.264	1.74×10^{-2}	1.503	1.95×10^{-7}
rs3227	6	33770273	ITPR3	C	0.498	1.043	6.02×10^{-8}	5	0.363	1.010	0.280	1.037	$3.24\times10^{-}$	0.415	1.046	9.44×10^{-5}	1.029	1.37×10^{-3}
rs4073894	7	104254200	LHFPL3	A	0.207	1.076	4.41×10^{-7}	0	0.765	1.003	0.467	1.062	$5.55\times10^{-}$	0.513	1.050	2.18×10^{-2}	1.073	1.74×10^{-5}
rs12640626	4	176863266	GPM6A	A	0.580	1.041	4.94×10^{-7}	10.9	0.234	1.000	0.495	1.034	$7.48\times10^{-}$	0.420	1.038	1.59×10^{-3}	1.031	7.61×10^{-4}


A
rs9320913
Plotted SNPs
rs9320913
r²
0.8
0.6
0.4
0.2
−log10(p-value)
Recombination rate (cM/Mb)
KLHL32
MIR2113
POU3F2
C6orf168
C6orf167
FBXL4
97.5
98
98.5
99
99.5
Position on chr6 (Mb)
B
rs11584700
Plotted SNPs
rs11584700
PRELP
LAX1
C1orf157
LOC127841
LRRN2
NFASC
DSTYK
MIR135B
OPTC
ZC3H11A
SOX13
PPP1R15B
CNTN2
NUAK2
MFSD4
ATP2B4
ETNK2
PIK3C2B
TMEM81
LEMD1
ELK4
SNORA77
REN
MDM4
RBBP5
CDK18
SNRPE
KISS1
TMCC2
LOC284578
GOLT1A
KLHDC8A
SLC45A3
PLEKHA6
NUCKS1
RAB7L1
SLC41A1
202
202.5
203
203.5
204
Position on chr1 (Mb)
C
rs4851266
Plotted SNPs
rs4851266
TSGA10
EIF5B
AFF3
LONRF2
PDCL3
NPAS2
C2orf29
C2orf15
REV1
CHST10
RPL31
SNORD89
LIPT1
NMS
TBC1D6
RFX8
MITD1
RNF149
MRPL30
CREG2
LYG2
LYG1
TXNDC9
99
99.5
100
100.5
101
Position on chr2 (Mb)

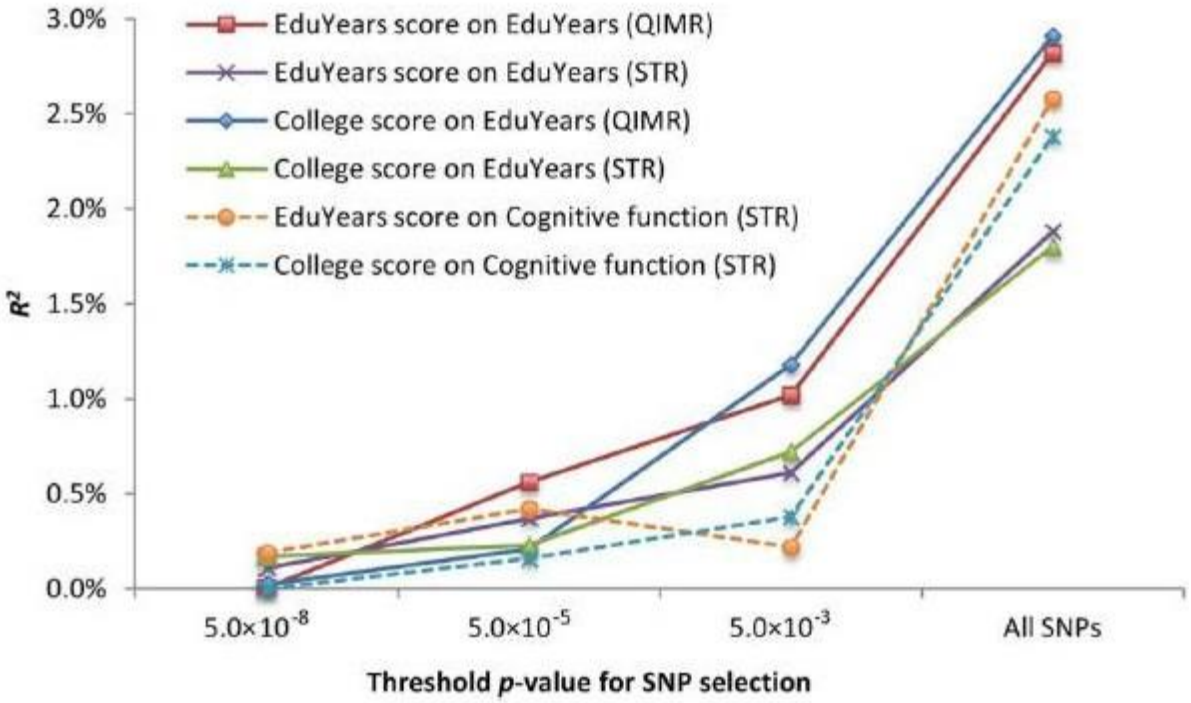
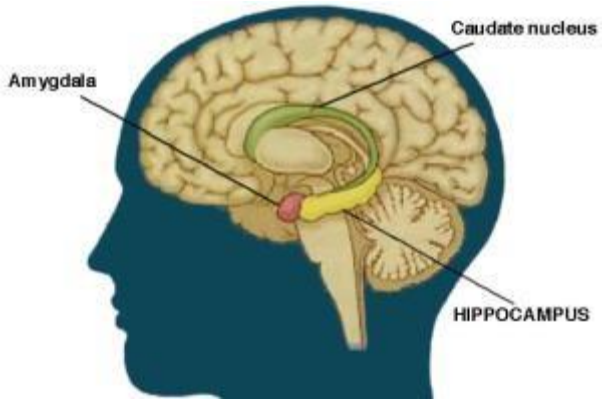


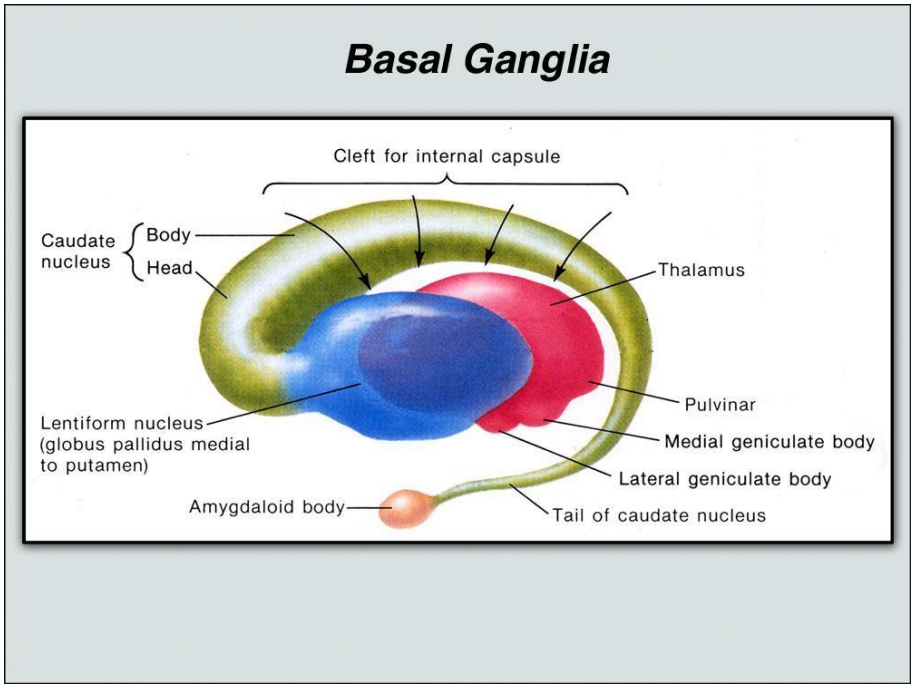
Fig. 2. Solid lines show results from regressions of *EduYears* on linear polygenic scores in a set of unrelated individuals from the QIMR (N = 3526) and STR (N = 6770) cohorts. Dashed lines show results from regressions of *Cognitive function* on linear polygenic scores in a sample from STR (N = 1419). The scores are constructed from the meta-analysis for either *EduYears* or *College*, excluding the QIMR and STR cohorts.

A genome-wide association study of educational attainment was conducted in a discovery sample of **101,069** individuals and a replication sample of **25,490. Three independent SNPs** are genome-wide significant (rs9320913, rs11584700, rs4851266), and all three replicate.

Estimated effects sizes are small **(*R*2 ≈ 0.02%),** approximately **1 month of schooling per allele**. A linear polygenic score from all measured SNPs accounts for **≈ 2% of the variance** in both educational attainment and cognitive function.

Genes in the region of the loci have previously been associated with health, cognitive, and central nervous system phenotypes, and bioinformatics analyses suggest the involvement of the **anterior caudate nucleus**.





While our linear polygenic score for education achieves an *R*2 of **2%** estimated from a sample of 120,000, a score for height reached **10%** estimated from a sample of 180,000, and a score for BMI using only the top 32 SNPs reached **1.4%**. Taken together, our findings suggest that the genetic architecture of complex behavioral traits is far more diffuse than that of complex physical traits.

LETTER

doi:10.1038/nature17671

Genome-wide association study identifies 74 loci associated with educational attainment

Here we report the results of a genome-wide association study (GWAS) for educational attainment that extends our earlier discovery sample of 101,069 individuals to 293,723 individuals, and a replication study in an independent sample of 111,349 individuals from the UK Biobank.

Figure 1 | Manhattan plot for *EduYears* associations (N = 293,723).
The x-axis is chromosomal position, and the y-axis is the significance on a $-\log_{10}$ scale. The black line shows the genome-wide significance level (5×10^{-8}). The red x's are the 74 approximately independent genome-wide significant associations ("lead SNPs"). The black dots labeled with rs numbers are the 3 Rietveld et al.[9] SNPs.

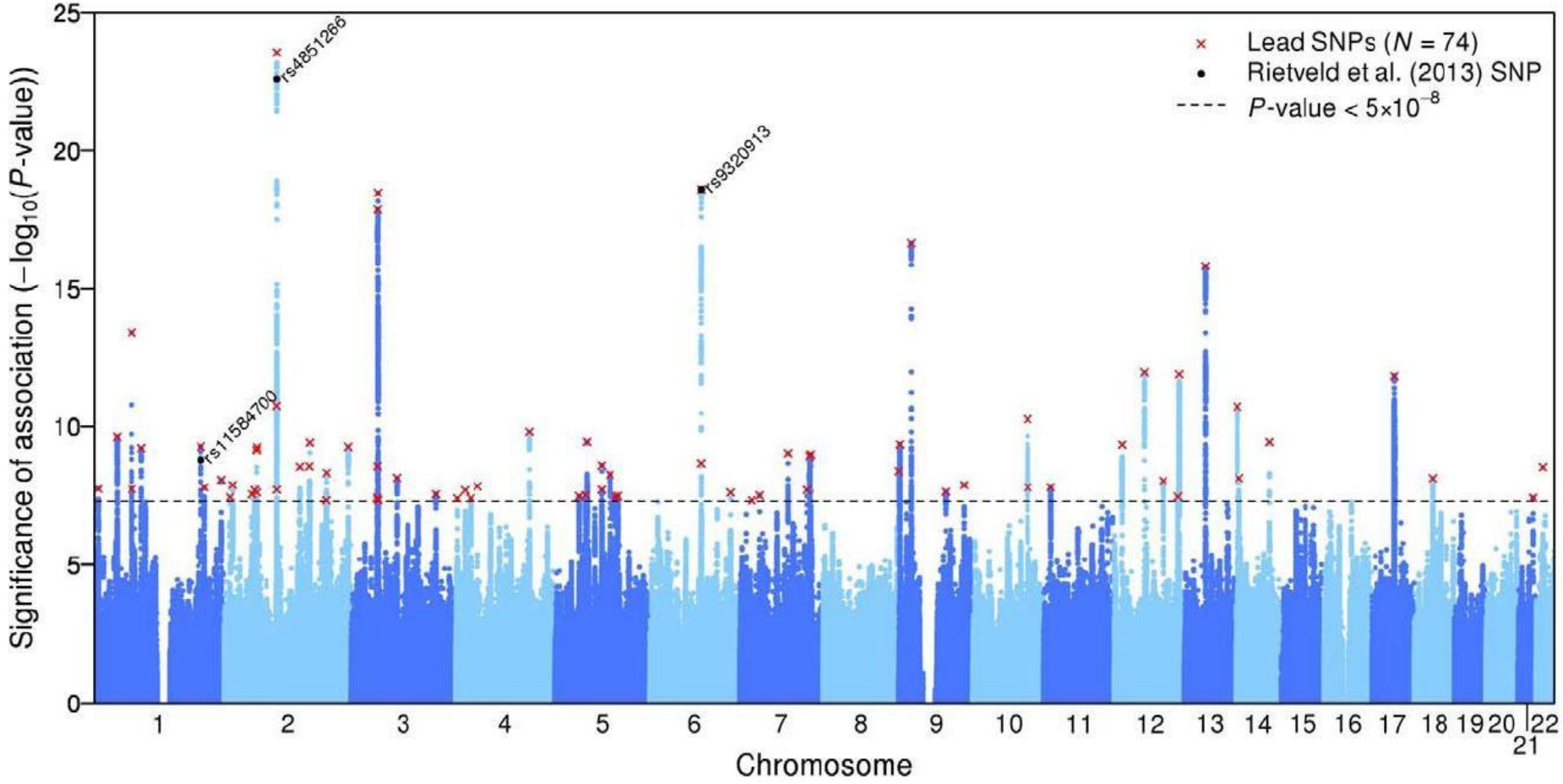
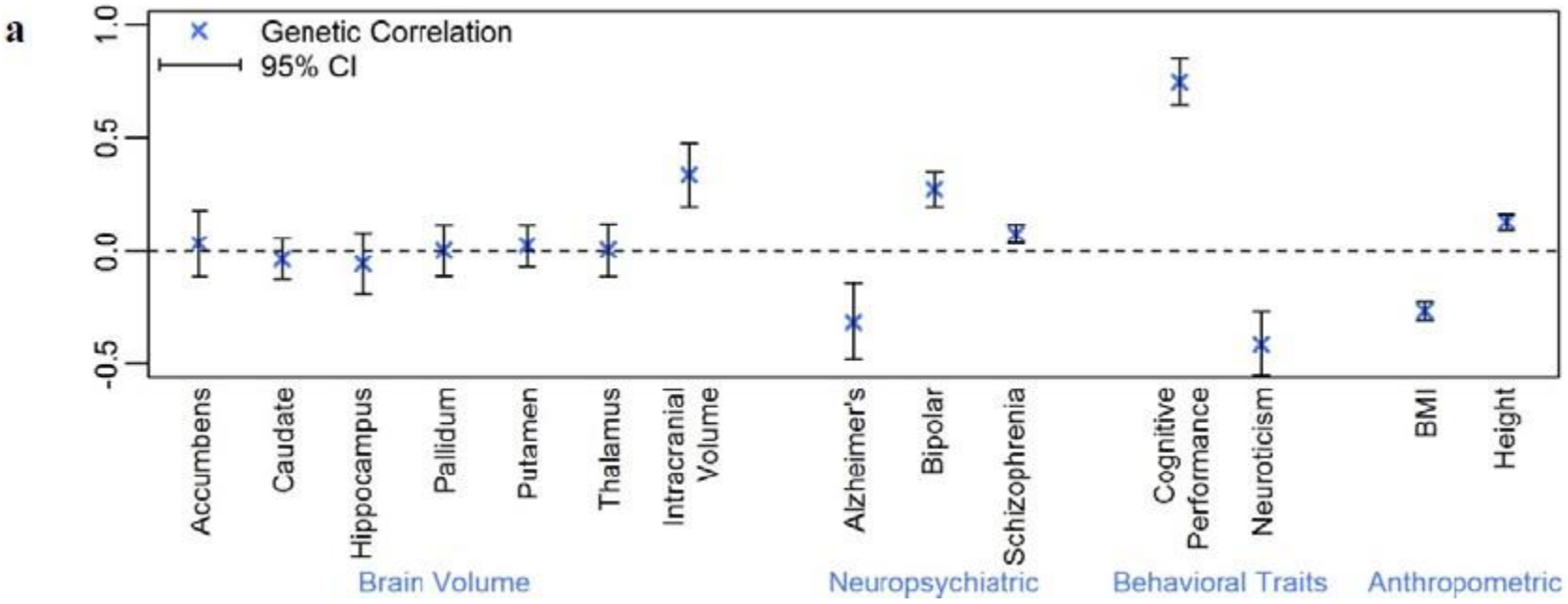


Figure 2 | Genetic overlap between *EduYears* and other traits.
a, Results from bivariate Linkage-Disequilibrium (LD) Score regressions: estimates of genetic correlation with brain volume, neuropsychiatric, behavioral, and anthropometric phenotypes for which GWAS summary statistics were available in the public domain. **b,** Q-Q plots for the 74 lead *EduYears* SNPs in selected phenotypes (for other phenotypes, see Extended Data Fig. 4). SNPs with concordant effects on both phenotypes are pink, and SNPs with discordant effects are blue. SNPs outside the gray area pass Bonferroni-corrected significance thresholds that correct for the total number of SNPs we tested ($P < 0.05/74 = 6.8\times10^{-4}$) and are labeled with their rs numbers. For the sign concordance test: * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$.



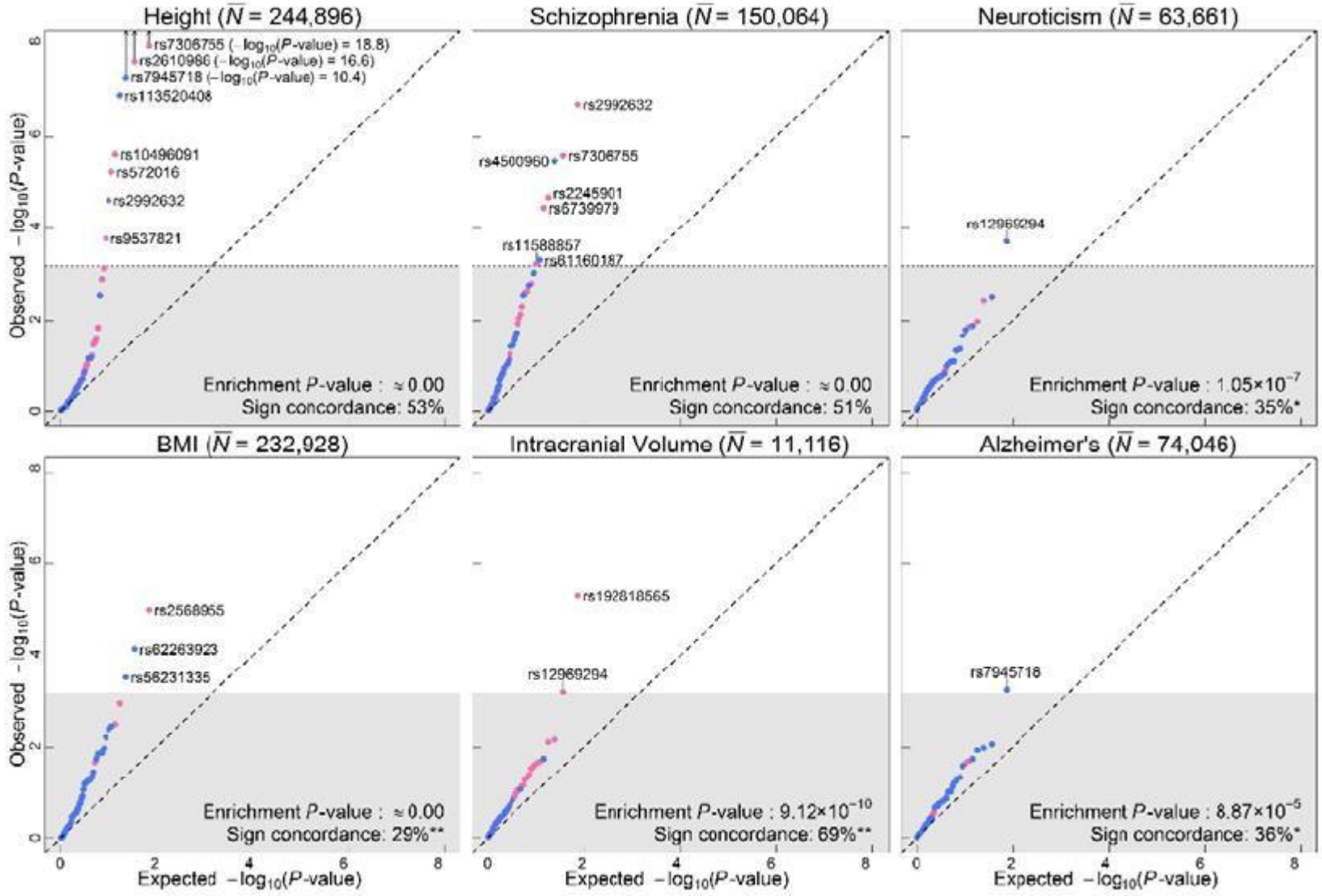
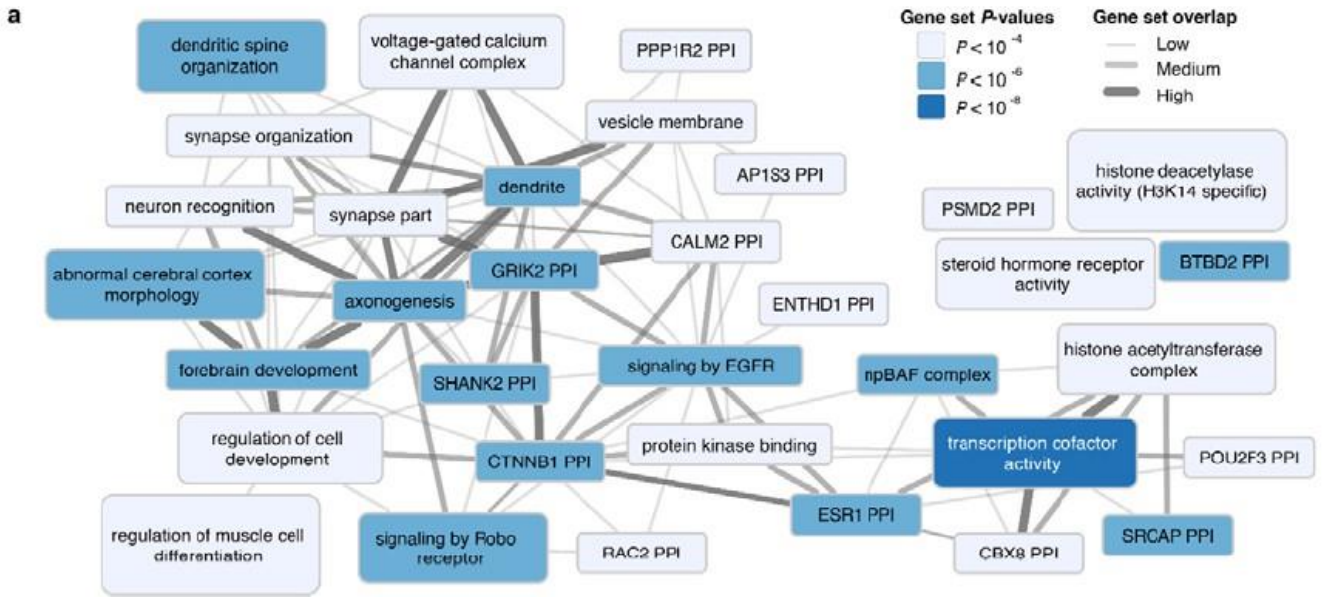

b
Height ($\bar{N}$ = 244,896)
rs7306755 (−log10(P-value) = 18.8)
rs2610986 (−log10(P-value) = 16.6)
rs7945718 (−log10(P-value) = 10.4)
rs113520408
rs10496091
rs572016
rs2992632
rs9537821
Enrichment P-value : ≈ 0.00
Sign concordance: 53%
Schizophrenia ($\bar{N}$ = 150,064)
rs2992632
rs4500960
rs7306755
rs2245901
rs6739979
rs11588857
rs61160187
Enrichment P-value : ≈ 0.00
Sign concordance: 51%
Neuroticism ($\bar{N}$ = 63,661)
rs12969294
Enrichment P-value : 1.05×10−7
Sign concordance: 35%*
BMI ($\bar{N}$ = 232,928)
rs2568955
rs62263923
rs56231335
Enrichment P-value : ≈ 0.00
Sign concordance: 29%**
Intracranial Volume ($\bar{N}$ = 11,116)
rs192818565
rs12969294
Enrichment P-value : 9.12×10−10
Sign concordance: 69%**
Alzheimer's ($\bar{N}$ = 74,046)
rs7945718
Enrichment P-value : 8.87×10−5
Sign concordance: 36%*
Observed −log10(P-value)
Expected −log10(P-value)
0
2
4
6
8

Figure 3 | Overview of biological annotation.

a, We identified 283 significantly (FDR < 0.05) enriched reconstituted gene sets, which we grouped using the Affinity Propagation method into the 34 clusters displayed in the graph. Each cluster is named after one of its member gene sets. The color represents the *P*-value of the member set exhibiting the most statistically significant enrichment. Overlap between pairs of clusters is represented by an edge. Edge width represents the Pearson correlation ρ between the two vectors of gene membership scores ($\rho < 0.3$, no edge; $0.3 \leq \rho < 0.5$, thin edge; $0.5 \leq \rho < 0.7$, intermediate edge; $\rho \geq 0.7$, thick edge), where each cluster's vector is the vector for the gene set after which the cluster is named. **b,** For each of the plotted disorders, we calculated the odds ratio corresponding to the overlap between genes where *de novo* mutations have been implicated in the disorder and genes prioritized by DEPICT as affecting the given GWAS phenotype. *EduYears*, years of education; BMI, body mass index; WHR, waist-to-hip ratio adjusted for BMI. **c,** The DEPICT-prioritized genes for *EduYears* measured in the BrainSpan Developmental Transcriptome data (red curve) exhibit higher expression in the brain during prenatal relative to postnatal stages. The DEPICT-prioritized genes exhibit similar median gene-expression levels across different brain areas (gray lines). Analyses were based on median $\log_2$-transformed RNA-Seq data. Error bars represent 95% confidence intervals.



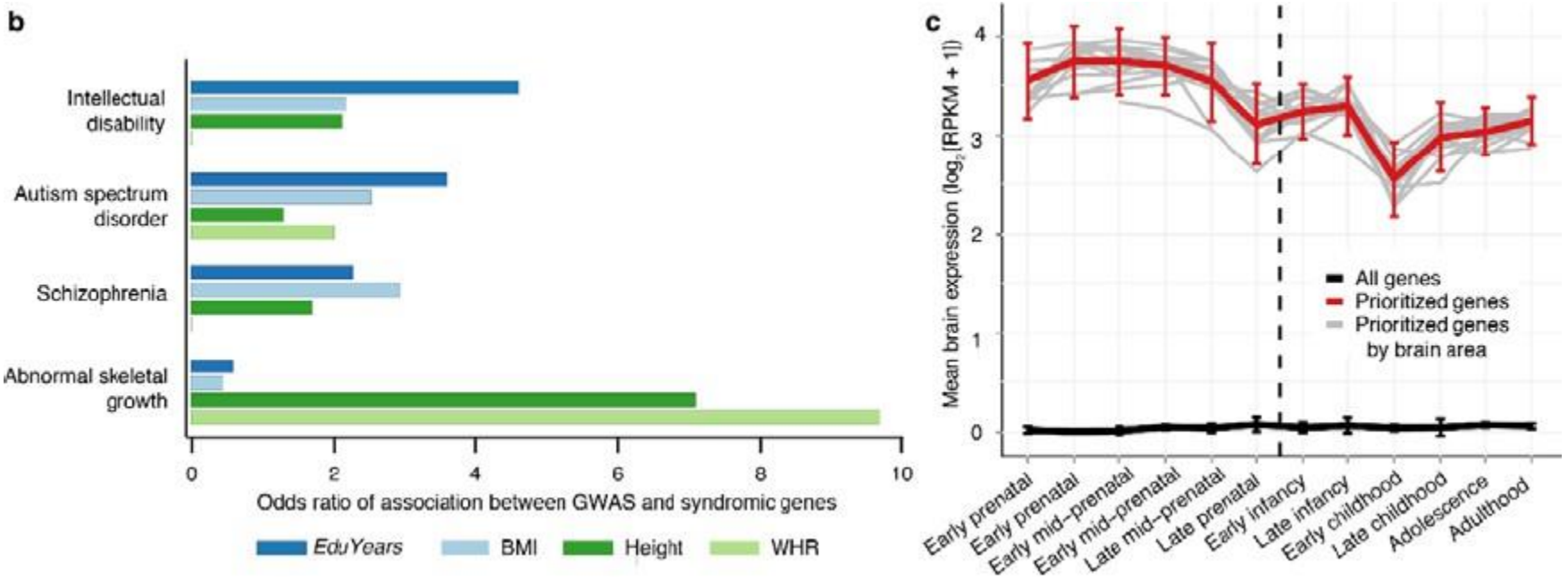
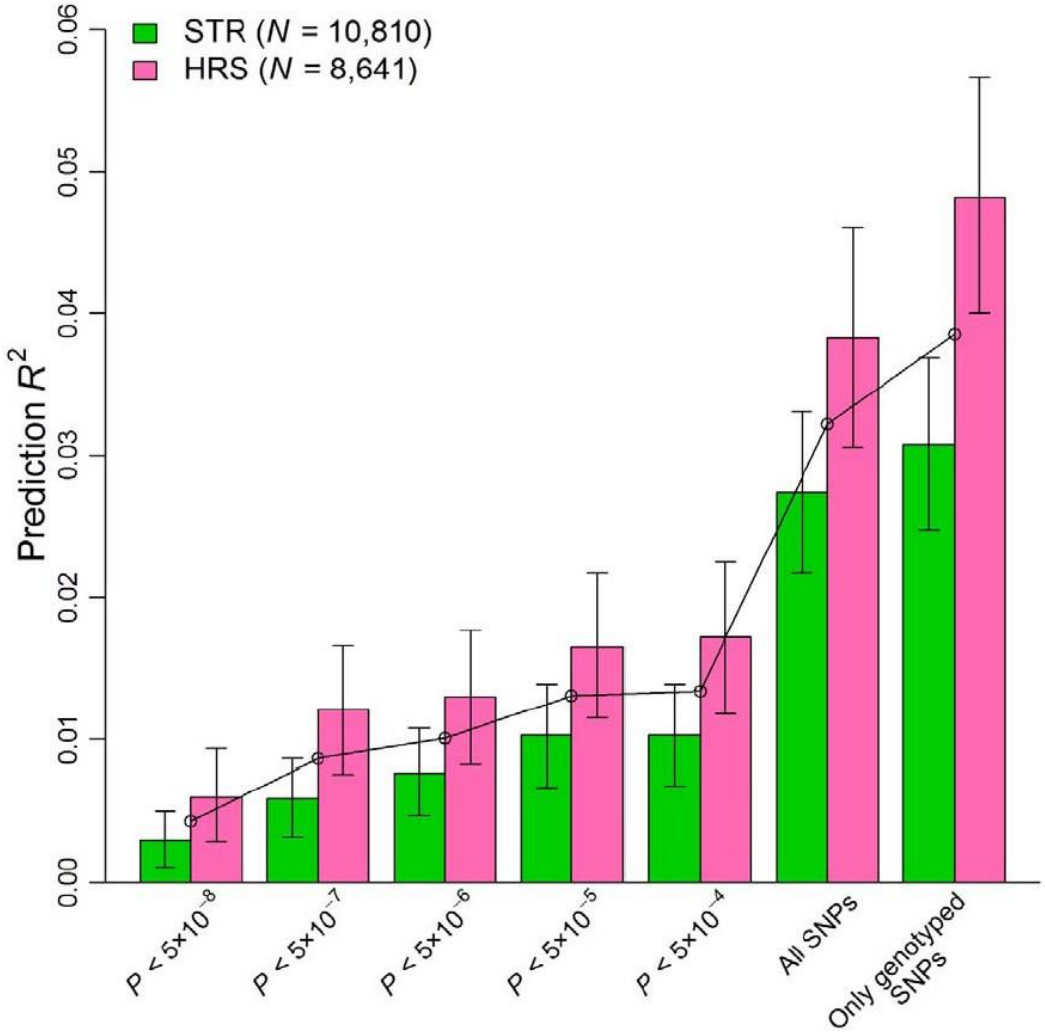
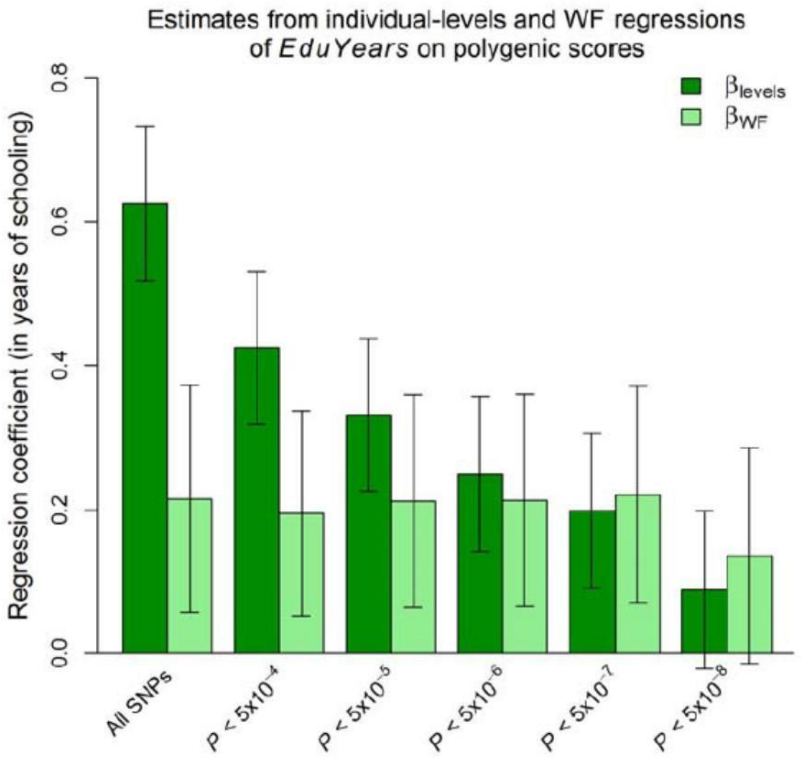
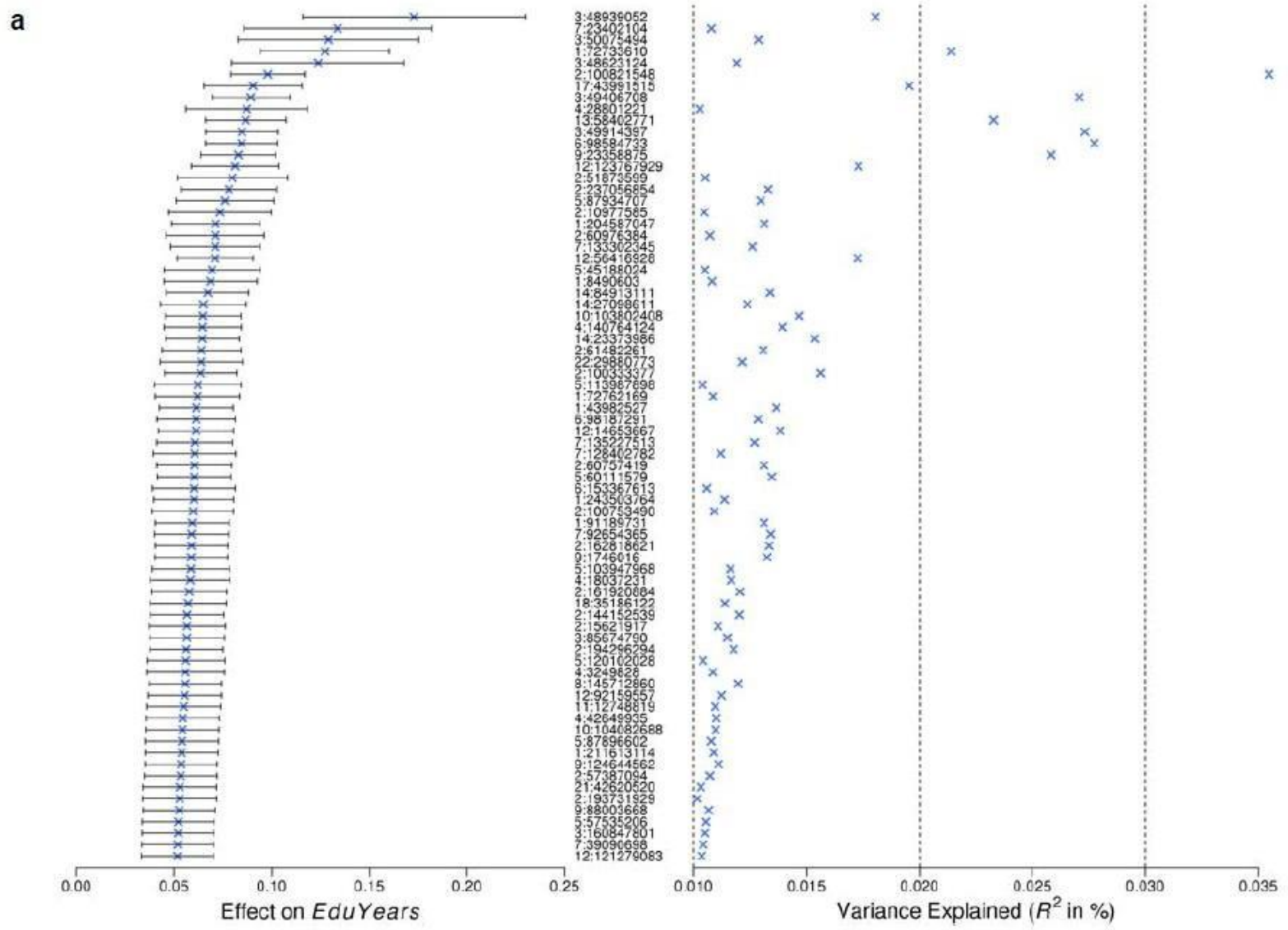

b
Intellectual disability
Autism spectrum disorder
Schizophrenia
Abnormal skeletal growth
0
2
4
6
8
10
Odds ratio of association between GWAS and syndromic genes
EduYears
BMI
Height
WHR
c
Mean brain expression (log2[RPKM + 1])
0
1
2
3
4
All genes
Prioritized genes
Prioritized genes by brain area
Early prenatal
Early prenatal
Early mid-prenatal
Early mid-prenatal
Late mid-prenatal
Late prenatal
Early infancy
Late infancy
Early childhood
Late childhood
Adolescence
Adulthood

Figure 4 | Variance explained by *EduYears*-associated SNPs at different levels of significance in the Swedish Twin Registry (STR) and Health and Retirement Study (HRS). Predictive power of the polygenic score in unrelated individuals. Cohorts are HRS (pink) and STR (green). The *y*-axis is the variance explained (incremental R^2 from including the score in the regression). The line is the sample-size-weighted mean R^2. The "All SNPs" score was constructed using all SNPs (genotyped and imputed), and the "Only genotyped SNPs" score was constructed using all genotyped SNPs. For the other scores, SNPs were selected from a GCTA-COJO analysis. For all scores, weights were estimated with the validation cohort excluded from the meta-analysis. All scores are residualized on the first 10 principal components.

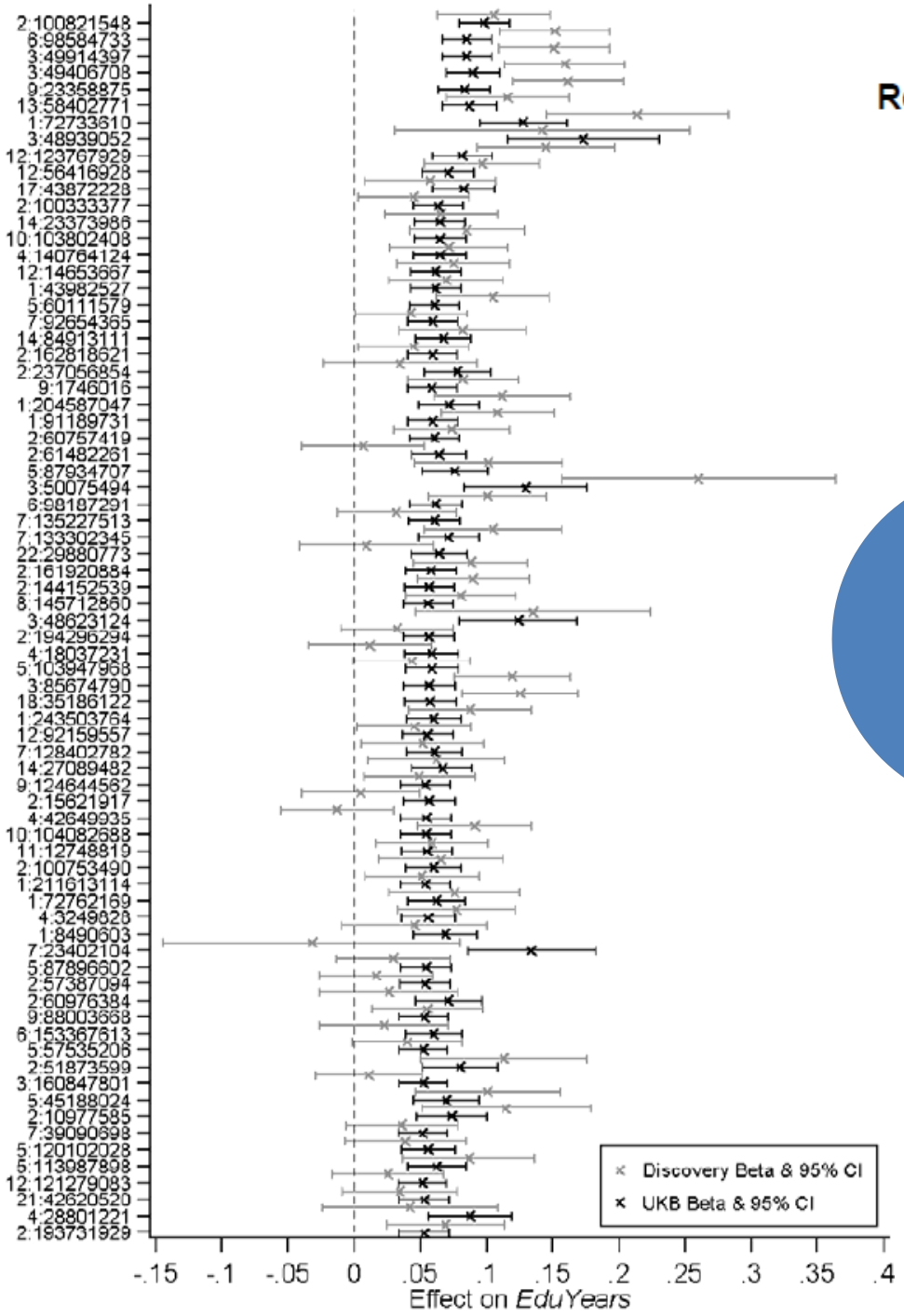




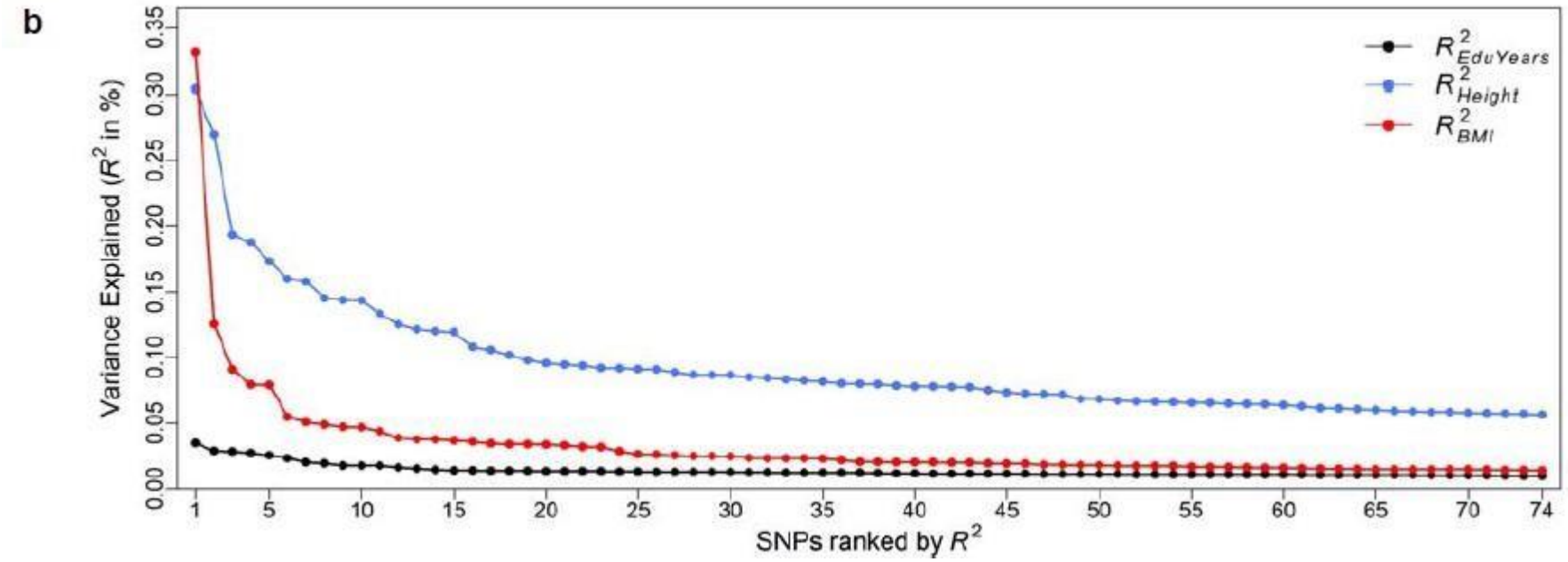
The distribution of effect sizes of the 74 lead SNPs



Replication of 74 lead SNPs in the UK Biobank data



Extended Data Fig. 2 shows the estimated effect sizes of the lead SNPs. The estimates range from **0.014 to 0.048** standard deviations per allele (**2.7 to 9.0 weeks of schooling**), with incremental R2 in the range **0.01% to 0.035%.**

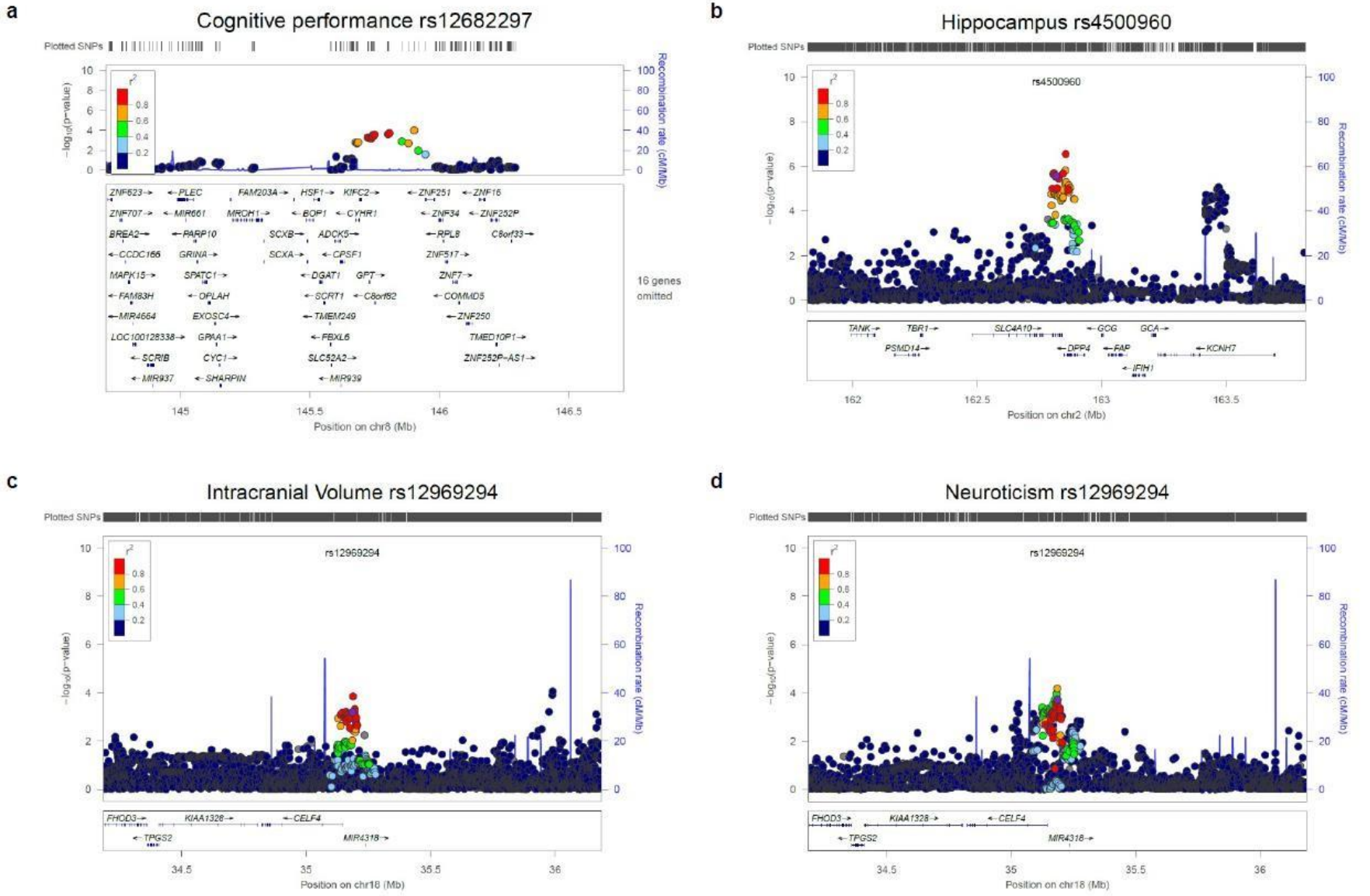


To further test the robustness of our findings, we examined the within sample and out-of-sample replicability of SNPs reaching genomewide significance. We found that SNPs identified in the previous educational attainment meta-analysis replicated in the new cohorts included here, and conversely, that SNPs reaching genome-wide significance in the new cohorts replicated in the old cohorts.

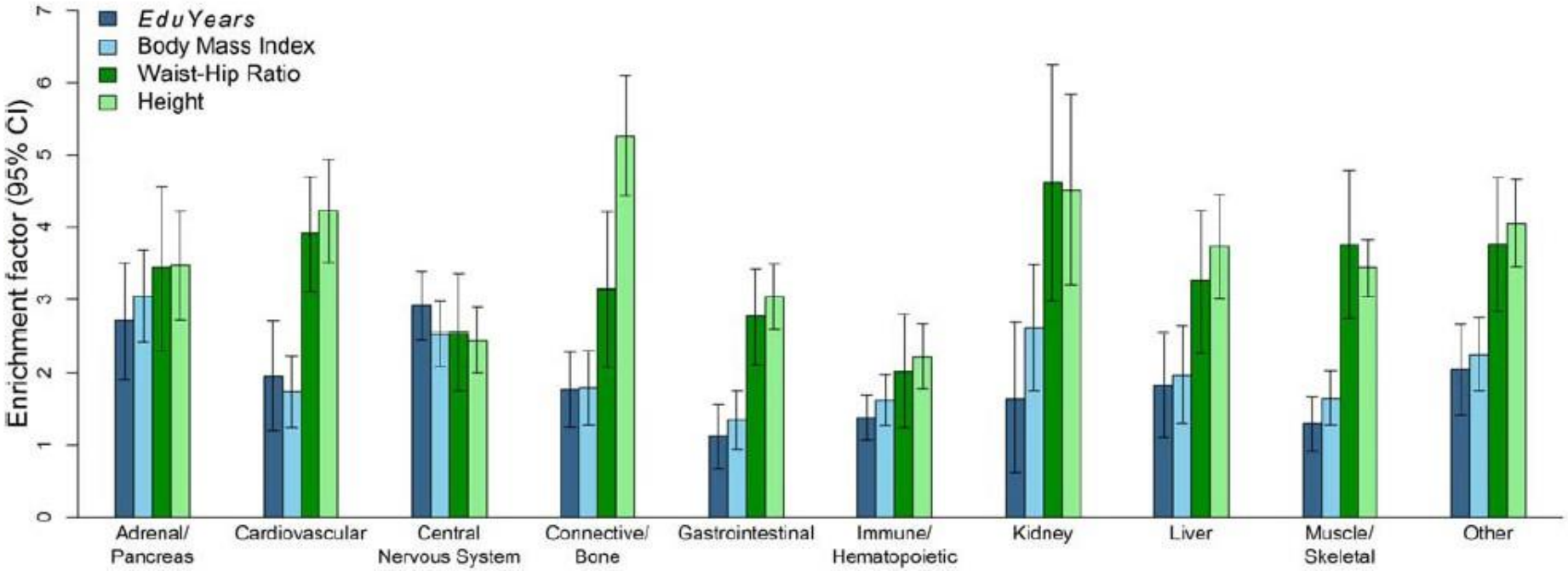
Table 1 | Top genes implicated by bioinformatics analyses

SNP	Gene symbol	Syndromic gene	Prioritization *P*-value	Lines of evidence	Top-ranking gene sets
Top 10 genes, based on DEPICT prioritization *P*-value					
rs34328009	*CAMTA1*	–	4.7×10^{-13}	2	abnormal dorsal root ganglion morphology, G alpha (12/13) signaling events
rs58852793	*PPP6R2*	–	5.2×10^{-12}	3	abnormal striatum morphology, histone deacetylase activity (H3K9 specific), NAD-dependent histone deacetylase activity
rs35436312	*FOXO6*	–	1.2×10^{-11}	3	dendrite development, regulation of cell development
rs72771875	*MEF2C*	ID, ASD	1.4×10^{-8}	5	ErbB signaling pathway, abnormal sternum ossification, regulation of muscle cell differentiation
rs9544418	*NBEA*	SCZ	1.8×10^{-8}	4	developmental biology, signaling by Robo receptor, dendritic shaft
rs12754946	*USP33*	–	6.1×10^{-8}	3	protein serine/threonine phosphatase complex
rs111730030	*NACC1*	–	6.8×10^{-8}	2	transcription factor binding, kinase binding, regulation of cell development
rs11191193	*MGEA5*	–	1.5×10^{-7}	2	protein kinase binding, kinase binding
rs12987662	*LONRF2*	–	4.4×10^{-7}	3	ectopic cerebellar granule cells, dendritic spine organization, synapse organization
rs11712056	*IP6K2*	–	6.0×10^{-7}	2	transcription coactivator activity, SWI/SNF-type complex, nBAF complex
Top 15 genes overlapping the DEPICT-defined loci of the lead SNPs, based on multiple, convergent lines of evidence					
rs4500960	*TBR1*	ID, ASD	6.3×10^{-4}	6	developmental biology, decreased brain size, abnormal cerebral cortex morphology
rs61160187	*ZSWIM6*	–	1.3×10^{-4}	5	transcription factor binding, negative regulation of signal transduction, PI3K events in ErbB4 signaling
rs2457660	*BCL11A*	ASD	7.6×10^{-4}	5	dendritic spine organization, abnormal hippocampal mossy fiber morphology, SWI/SNF-type complex
rs11712056	*CELSR3*	SCZ	8.9×10^{-4}	5	dendrite morphogenesis, dendrite development, abnormal hippocampal mossy fiber morphology
rs192818565	*MAPT*	ID	9.5×10^{-4}	5	dendrite morphogenesis, abnormal hippocampal mossy fiber morphology, abnormal axon guidance
rs7306755	*SBNO1*	SCZ	7.2×10^{-3}	5	protein serine/threonine phosphatase complex
rs12987662	*NBAS*	–	0.05	5	–
rs1871109	*SMARCA2*	ID	8.8×10^{-6}	4	–
rs11712056	*MAP4*	ASD	2.6×10^{-5}	4	developmental biology, signaling by Robo receptor, SWI/SNF-type complex
rs10061788	*LINC00461*	–	6.8×10^{-5}	4	decreased brain size, abnormal cerebral cortex morphology, abnormal hippocampal mossy fiber morphology
rs9320913	*POU3F2*	–	3.6×10^{-4}	4	dendrite morphogenesis, developmental biology, decreased brain size
rs11712056	*RAD54L2*	SCZ	4.8×10^{-4}	4	decreased brain size, SWI/SNF-type complex, nBAF complex
rs2964197	*PLK2*	–	8.1×10^{-4}	4	negative regulation of signal transduction, PI3K events in ErbB4 signaling
rs9537821	*PCDH17*	–	1.6×10^{-3}	4	activation of RAC, abnormal axon guidance, axon guidance (Reactome)

Extended Data Figure 6 | Regional association plots for four of the ten prioritized SNPs for MHBA phenotypes identified using *EduYears* as a proxy phenotype: a, cognitive performance; **b,** hippocampus; **c,** intracranial volume; **d,** neuroticism. The four were selected because very few genome-wide significant SNPs have been previously reported for these traits. Data sources and methods are described in Supplementary Information section 3. The R^2 values are from the hg19 / 1000 Genomes Nov 2014 EUR references samples. The figures were created with LocusZoom (http://csg.sph.umich.edu/locuszoom/). Mb, megabases.

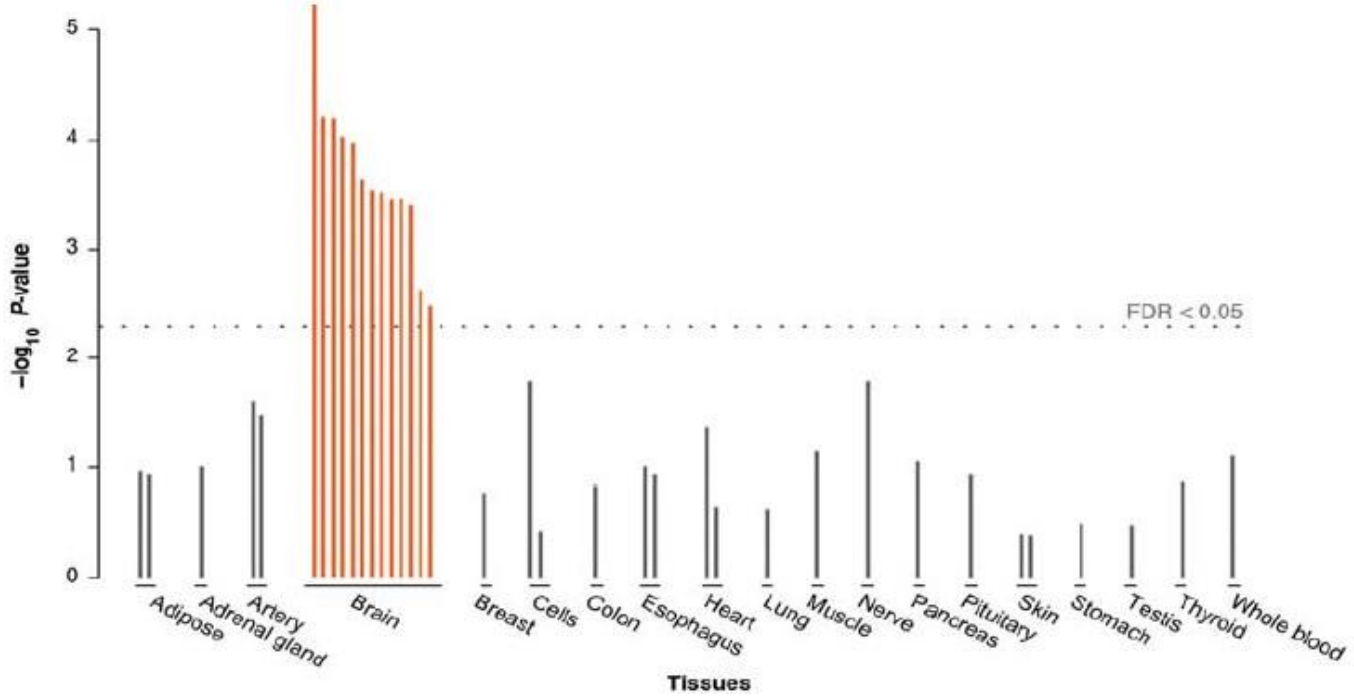


Extended Data Figure 8 | Estimates of enrichment factors by cell-type groups using the method of Funucane et al. (2015). The enrichment factor for a given cell-type group is the ratio of variance explained by SNPs in that group to the overall fraction of SNPs in that group. To benchmark the estimates for *EduYears*, we compare the enrichment factors to those obtained when we use the largest GWAS conducted to date on body mass index, height, and waist-to-hip ratio adjusted for BMI. The estimates were produced with the LDSC python software, using the LD scores and functional partitions introduced in Finucane et al. (2015) and the HapMap3 SNPs with MAF > 0.05. Each of the 10 enrichment calculations for a particular cell type is performed independently, while each controlling for the 52 functional annotation categories in the full baseline model. The 95% confidence intervals are not adjusted for multiple hypothesis testing.

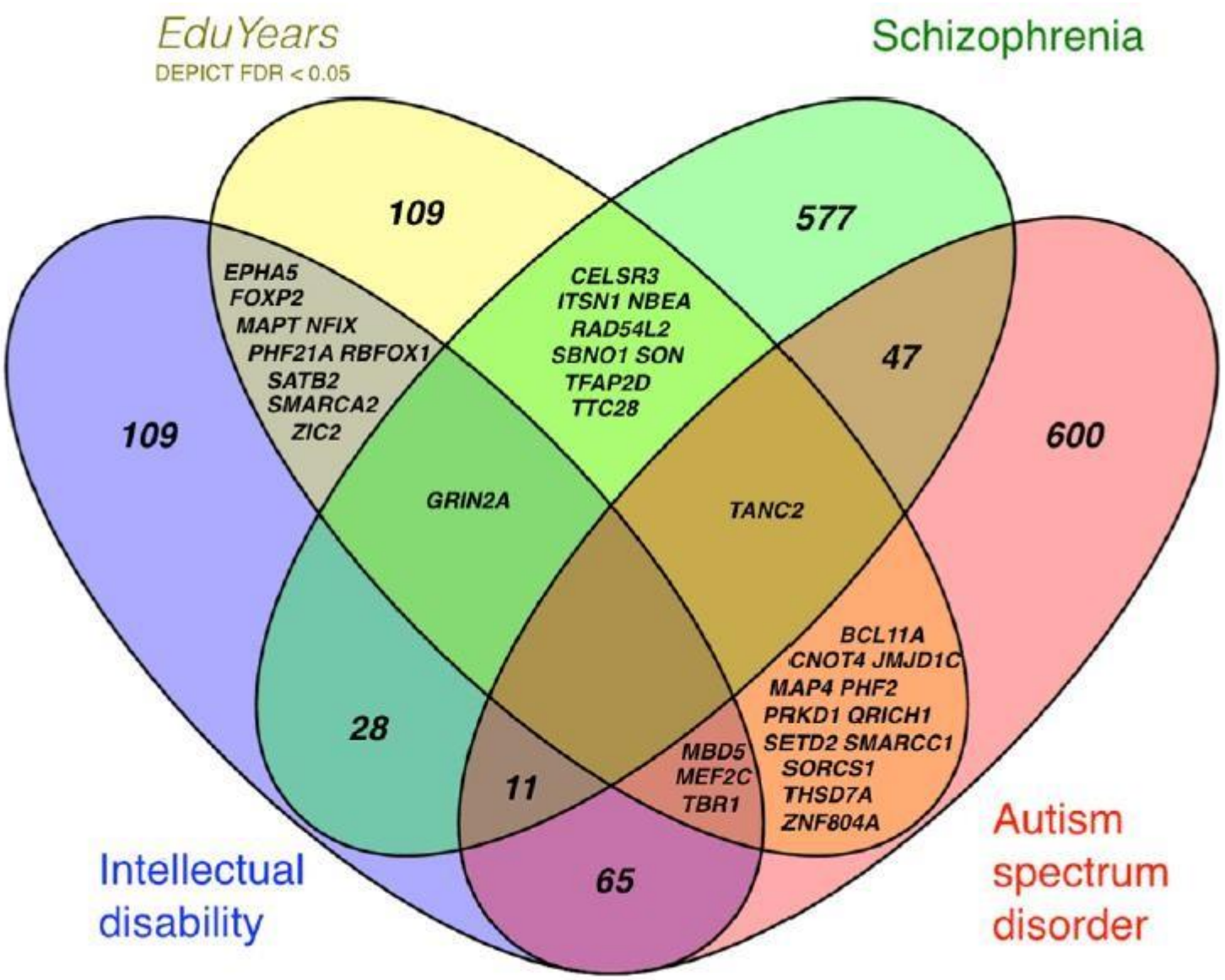


Extended Data Figure 9 | Additional biological annotation. a, Enrichment of tissue types. We took measurements of gene expression by the Genotype-Tissue Expression (GTEx) Consortium and tested whether the genes overlapping *EduYears*-associated loci are significantly overexpressed (relative to genes in random sets of loci matched by gene density) in each of 37 tissue types. The tissue types are grouped by organ. The *P*-value corresponding to FDR < 0.05 is represented by the black dotted line, and tissues passing this significance threshold are shown in red. **b,** Significantly (FDR < 0.05) DEPICT-prioritized genes in *EduYears*-associated loci exhibit substantial overlap with genes previously reported to harbor sites where mutations increase risk of intellectual disability and autism spectrum disorder.

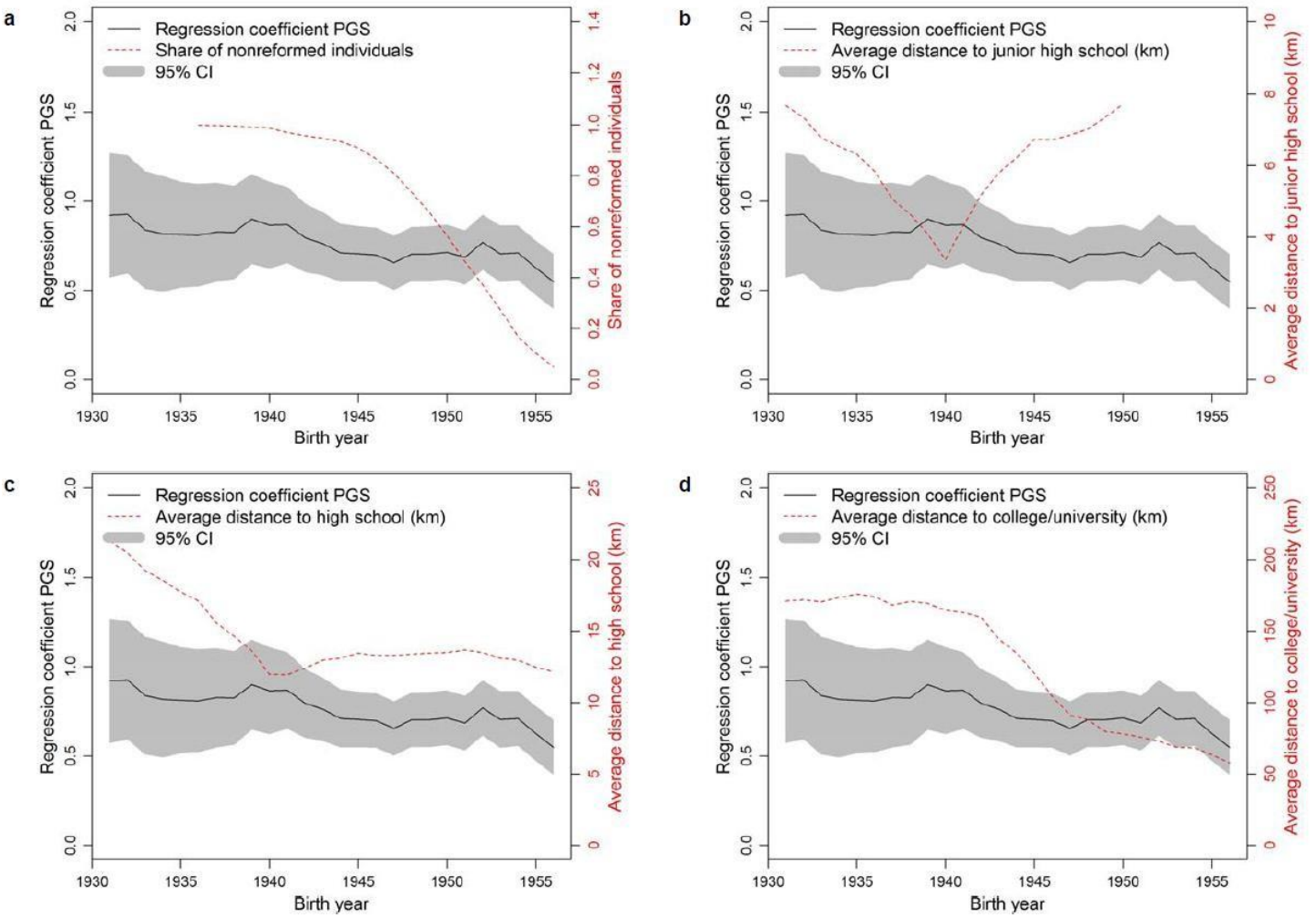
a



b


EduYears
DEPICT FDR < 0.05
Schizophrenia
109
577
EPHA5
FOXP2
MAPT NFIX
PHF21A RBFOX1
SATB2
SMARCA2
ZIC2
CELSR3
ITSN1 NBEA
RAD54L2
SBNO1 SON
TFAP2D
TTC28
47
109
600
GRIN2A
TANC2
BCL11A
CNOT4 JMJD1C
MAP4 PHF2
PRKD1 QRICH1
SETD2 SMARCC1
SORCS1
THSD7A
ZNF804A
28
MBD5
MEF2C
TBR1
11
65
Intellectual disability
Autism spectrum disorder

Extended Data Figure 10 | The predictive power of the all-SNPs polygenic score (PGS) varies in Sweden by birth cohort. Five-year roll
regressions of years of education on the PGS (left axis in all four panels), share of individuals not affected by the comprehensive school reform
right axis), and average distance to nearest junior high school (**b**, right axis), nearest high school (**c**, right axis) and nearest college/university (
right axis). The shaded area displays the 95% confidence interval for the PGS effect.



Across our two holdout samples, the mean predictive power of a polygenic score constructed from all measured SNPs is **3.2%**