Sciencexpress Reports

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

All authors with their affiliations appear at the end of this paper.

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. These findings provide promising candidate SNPs for follow-up work, and our effect size estimates can anchor power analyses in social-science genetics.

Twin and family studies suggest that a broad range of psychological traits (*1*), economic preferences (*2–4*), and social and economic outcomes (*5*) are moderately heritable. Discovery of genetic variants associated with such traits leads to insights regarding the biological pathways underlying human behavior. If the predictive power of a set of genetic variants considered jointly is sufficiently large, then a "risk score" that aggregates their effects could be useful to control for genetic factors that are otherwise unobserved, or to identify populations with certain genetic propensities, for example in the context of medical intervention (*6*).

To date, however, few if any robust associations between specific genetic variants and social-scientific outcomes have been identified likely because existing work [for review see (*7*)] has relied on samples that are too small [for discussion, see (*4*, *6*, *8*, *9*)]. In this paper, we apply to a complex behavioral trait—educational attainment—an approach to gene discovery that has been successfully applied to medical and physical phenotypes (*10*), namely meta-analyzing data from multiple samples.

The phenotype of educational attainment is available in many samples with genotyped subjects (*5*). Educational attainment is influenced by many known environmental factors, including public policies. Educational attainment is strongly associated with social outcomes, and there is a well-documented health-education gradient (*5*, *11*). Estimates suggest that around 40% of the variance in educational attainment is explained by genetic factors (*5*). Furthermore, educational attainment is moderately correlated with other heritable characteristics (*1*), including cognitive function (*12*) and personality traits related to persistence and self-discipline (*13*).

To create a harmonized measure of educational attainment, we coded study-specific measures using the International Standard Classification of Education (ISCED 1997) scale (*14*). 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.

A genome-wide association study (GWAS) meta-analysis was performed across 42 cohorts in the discovery phase. The overall discovery sample comprises 101,069 individuals for *EduYears* and 95,427 for *College*. Analyses were performed at the cohort level according to a prespecified analysis plan, which restricted the sample to Caucasians (to help reduce stratification concerns). Educational attainment was measured at an age at which subjects were very likely to have completed their education [over 95% of the sample was at least 30; (*5*)]. On average, subjects have 13.3 years of schooling, and 23.1% have a college degree. To enable pooling of GWAS results, all studies conducted analyses with data imputed to the HapMap 2 CEU (r22.b36) reference set. To guard against population stratification, the first four principal components of the genotypic data were included as controls in all the cohort-level analyses. All study-specific GWAS results were quality controlled, cross-checked, and meta-analyzed using single genomic control and a sample-size weighting scheme at three independent analysis centers.

At the cohort level, there is little evidence of general inflation of *p*-values. As in previous GWA studies of complex traits (*15*), the Q-Q plot of the meta-analysis exhibits strong inflation. This inflation is not driven by specific cohorts and is expected for a highly polygenic phenotype even in the absence of population stratification (*16*).

From the discovery phase, we identified one genome-wide significant locus (rs9320913, $p = 4.2 \times 10^{-9}$) and three suggestive loci (defined as $p < 10^{-6}$) for *EduYears*. For *College*, we identified two genome-wide significant loci (rs11584700, $p = 2.1 \times 10^{-9}$, and rs4851266, $p = 2.2 \times 10^{-9}$) and an additional four suggestive loci (Table 1). We conducted replication analyses in 12 additional, independent cohorts that became available after the completion of the discovery meta-analysis, using the same pre-specified analysis plan. For both *EduYears* and *College*, the replication sample comprises 25,490 individuals.

For each of the ten loci that reached at least suggestive significance, we brought forward for replication the SNP with the lowest *p*-value. The three genome-wide significant SNPs replicate at the Bonferroni-adjusted 5% level, with point estimates of the same sign and similar magnitude (Fig. 1 and Table 1). The seven loci that did not reach genome-wide significance did not replicate (the effect went in the anticipated direction in 5 out of 7 cases). The meta-analytic findings are not driven by extreme results in a small number of cohorts (see p_{het} in Table 1), by cohorts from a specific geographic region (figs. S7 to S15), or by a single sex (figs. S3 to S6). Given the high correlation between *EduYears* and *College* (*5*), it is unsurprising that the set of SNPs with low *p*-values exhibit considerable overlap in the two analyses (tables S8 and S9).

The observed effect sizes of the three replicated individual SNPs are small [see (*5*) for discussion]. For *EduYears*, the strongest effect identified (rs9320913) explains 0.022% of phenotypic variance in the replication sample. This R^2 corresponds to a difference of ~1 months of schooling per allele. For college completion, the SNP with the strongest estimated effect (rs11584700) has an odds ratio of 0.912 in the replication sample, equivalent to a 1.8 percentage-point difference per allele in the frequency of completing college.

We subsequently conducted a "combined stage" meta-analysis, including both the discovery and replication samples. This analysis revealed additional genome-wide significant SNPs: four for *EduYears* and three for *College*. Three of these newly genome-wide significant SNPs (rs1487441, rs11584700, rs4851264) are in linkage disequilibrium with the replicated SNPs. The remaining four are located in different loci and warrant replication attempts in future research: rs7309, a 3′UTR variant

in *TANK*; rs11687170, close to *GBX2*; rs1056667, a 3′UTR variant in *BTN1A1*; and rs13401104 in *ASB18*.

Using the results of the combined meta-analyses of discovery and replication cohorts, we conducted a series of complementary and exploratory supplemental analyses to aid in interpreting and contextualizing the results: gene-based association tests; eQTL analyses of brain and blood tissue data; pathway analysis; functional annotation searches; enrichment analysis for cell-type-specific overlap with H3K4me3 chromatin marks; and predictions of likely gene function using gene-expression data. Table S20 summarizes promising candidate loci identified through follow-up analyses (*5*). Two regions in particular showed convergent evidence from functional annotation, blood cis-eQTL analyses, and gene-based tests: chromosome 1q32 (including *LRRN2*, *MDM4*, and *PIK3C2B*) and chromosome 6 near the Major Histocompatibility Complex (MHC). We also find evidence that in anterior caudate cells, there is enrichment of H3K4me3 chromatin marks (believed to be more common in active regulatory regions) in the genomic regions implicated by our analyses (fig. S20). Many of the implicated genes have previously been associated with health, central nervous system, or cognitive-process phenotypes in either human-GWAS or model-animal studies (table S22). Gene co-expression analysis revealed that several implicated genes (including *BSN*, *GBX2*, *LRRN2*, and *PIK3C2B*) are likely involved in pathways related to cognitive processes (such as learning and long-term memory) and neuronal development or function (table S21).

Although the effects of individual SNPs on educational attainment are small, many of their potential uses in social science depend on their combined explanatory power. To evaluate the combined explanatory power, we constructed a linear polygenic score (*5*) for each of our two education measures using the meta-analysis results (combining discovery and replication), excluding one cohort. We tested these scores for association with educational attainment in the excluded cohort. We constructed the scores using SNPs whose nominal *p*-values fall below a certain threshold, ranging from 5×10^{-8} (only the genome-wide significant SNPs were included) to 1 (all SNPs were included).

We replicated this procedure with two of the largest cohorts in the study, both of which are family-based samples (QIMR and STR). The results suggest that educational attainment is a highly polygenic trait (Fig. 2 and table S23): the amount of variance accounted for increases as the *p*-value threshold becomes less conservative (i.e., includes more SNPs). The linear polygenic score from all measured SNPs accounts for $\approx 2\%$ ($p = 1.0 \times 10^{-29}$) of the variance in *EduYears* in the STR sample and $\approx 3\%$ ($p = 7.1 \times 10^{-24}$) in the QIMR sample.

To explore one of the many potential mediating endophenotypes, we examined how much the same polygenic scores (constructed to explain *EduYears* or *College*) could explain individual differences in cognitive function. While it would have been preferable to explore a richer set of mediators, this variable was available in STR, a dataset where we had access to the individual-level genotypic data. Cognitive function had been measured in a subset of males using the Swedish Enlistment Battery (used for conscription) (*5*, *17*). The estimated $R^2 \approx 2.5\%$ ($p < 1.0 \times 10^{-8}$) for cognitive function is actually slightly larger than the fraction of variance in educational attainment captured by the score in the STR sample. One possible interpretation is that some of the SNPs used to construct the score matter for education through their stronger, more direct effects on cognitive function (*5*). A mediation analysis (table S24) provides tentative evidence consistent with this interpretation.

The polygenic score remains associated with educational attainment and cognitive function in within-family analyses (table S25). Thus, these results appear robust to possible population stratification.

If the size of the training sample used to estimate the linear polygenic score increased, the explanatory power of the score in the prediction sample would be larger because the coefficients used for constructing the score would be estimated with less error. In (*5*), we report projections of this increase. We also assess, at various levels of explanatory power, the benefits from using the score as a control variable in a randomized educational intervention (*5*). An asymptotic upper bound for the explanatory power of a linear polygenic score is the additive genetic variance across individuals captured by current SNP microarrays. Using combined data from STR and QIMR, we estimate that this upper bound is 22.4% (*S.E.* = 4.2%) in these samples (*5*) (table S12).

Placed in the context of the GWAS literature (*10*), our largest estimated SNP effect size of 0.02% is over an order of magnitude smaller than those observed for height and BMI: 0.4% (*15*) and 0.3% (*18*) respectively. 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 (*15*), and a score for BMI using only the top 32 SNPs reached 1.4% (*18*). Taken together, our findings suggest that the genetic architecture of complex behavioral traits is far more diffuse than that of complex physical traits.

Existing claims of "candidate gene" associations with complex social-science traits have reported widely varying effect sizes—many with R^2 values more than one hundred times larger than those we find (*4*, *6*). For complex social-science phenotypes that are likely to have a genetic architecture similar to educational attainment, our estimate of 0.02% can serve as a benchmark for conducting power analyses and evaluating the plausibility of existing findings in the literature.

The few GWAS studies conducted to date in social-science genetics have not found genome-wide significant SNPs that replicate consistently (*19*, *20*). One commonly proposed solution is to gather better measures of the phenotypes in more environmentally homogenous samples. Our findings demonstrate the feasibility of a complementary approach: identify a phenotype that, although more distal from genetic influences, is available in a much larger sample [see (*5*) for a simple theoretical framework and power analysis]. The genetic variants uncovered by this "proxy-phenotype" methodology can then serve as a set of empirically-based candidate genes in follow-up work, such as tests for associations with well-measured endophenotypes (e.g., personality, cognitive function), research on gene-environment interactions, or explorations of biological pathways.

In social-science genetics, researchers must be especially vigilant to avoid misinterpretations. One of the many concerns is that a genetic association will be mischaracterized as "the gene for X," encouraging misperceptions that genetically influenced phenotypes are immune to environmental intervention [for rebuttals, see (*21*, *22*)] and misperceptions that individual SNPs have large effects (which our evidence contradicts). If properly interpreted, identifying SNPs and constructing polygenic scores are steps toward usefully incorporating genetic data into social-science research.

References and Notes

1. R. Plomin, J. DeFries, V. Knopik, J. Neiderhiser, *Behavioral Genetics* (Worth Publishers, ed. 6, 2013), pp. 560.
2. D. Cesarini, C. T. Dawes, M. Johannesson, P. Lichtenstein, B. Wallace, Genetic variation in preferences for giving and risk taking. *Q. J. Econ.* **124**, 809 (2009). doi:10.1162/qjec.2009.124.2.809
3. D. J. Benjamin *et al*., The genetic architecture of economic and political preferences. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8026 (2012). doi:10.1073/pnas.1120666109 Medline
4. J. P. Beauchamp *et al*., Molecular genetics and economics. *J. Econ. Perspect.* **25**, 57 (2011). doi:10.1257/jep.25.4.57 Medline
5. Please see the supplementary materials on *Science* Online.
6. D. J. Benjamin *et al*., The promises and pitfalls of genoeconomics. *Annu. Rev. Econ.* **4**, 627 (2012). doi:10.1146/annurev-economics-080511-110939 Medline
7. R. P. Ebstein, S. Israel, S. H. Chew, S. Zhong, A. Knafo, Genetics of human social behavior. *Neuron* **65**, 831 (2010). doi:10.1016/j.neuron.2010.02.020 Medline

8. L. E. Duncan, M. C. Keller, A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *Am. J. Psychiatry* **168**, 1041 (2011). doi:10.1176/appi.ajp.2011.11020191 Medline
9. J. P. Ioannidis, Why most published research findings are false. *PLoS Med.* **2**, e124 (2005). doi:10.1371/journal.pmed.0020124 Medline
10. P. M. Visscher, M. A. Brown, M. I. McCarthy, J. Yang, Five years of GWAS discovery. *Am. J. Hum. Genet.* **90**, 7 (2012). doi:10.1016/j.ajhg.2011.11.029 Medline
11. J. P. Mackenbach *et al*.; European Union Working Group on Socioeconomic Inequalities in Health, Socioeconomic inequalities in health in 22 European countries. *N. Engl. J. Med.* **358**, 2468 (2008). doi:10.1056/NEJMsa0707519 Medline
12. I. J. Deary, S. Strand, P. Smith, C. Fernandes, Intelligence and educational achievement. *Intelligence* **35**, 13 (2007). doi:10.1016/j.intell.2006.02.001
13. J. J. Heckman, Y. Rubinstein, The importance of noncognitive skills: Lessons from the GED testing program. *Am. Econ. Rev.* **91**, 145 (2001). doi:10.1257/aer.91.2.145
14. UNESCO Institute for Statistics, International Standard Classification of Education (2006).
15. H. Lango Allen *et al*., Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832 (2010). doi:10.1038/nature09410 Medline
16. J. Yang *et al*.; GIANT Consortium, Genomic inflation factors under polygenic inheritance. *Eur. J. Hum. Genet.* **19**, 807 (2011). doi:10.1038/ejhg.2011.39 Medline
17. B. Carlstedt, *Cognitive Abilities: Aspects of Structure, Process and Measurement* (Acta Universitatis Gothoburgensis, Göteborg, Sweden, 2000).
18. E. K. Speliotes *et al*.; MAGIC; Procardis Consortium, Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **42**, 937 (2010). doi:10.1038/ng.686 Medline
19. M. H. de Moor *et al*., Meta-analysis of genome-wide association studies for personality. *Mol. Psychiatry* **17**, 337 (2012). doi:10.1038/mp.2010.128 Medline
20. B. Benyamin *et al*.; Wellcome Trust Case Control Consortium 2 (WTCCC2), Childhood intelligence is heritable, highly polygenic and associated with FNBP1L. *Mol. Psychiatry* 10.1038/mp.2012.184 (2013). doi:10.1038/mp.2012.184 Medline
21. C. Jencks, Heredity, environment, and public policy reconsidered. *Am. Sociol. Rev.* **45**, 723 (1980). doi:10.2307/2094892 Medline
22. A. S. Goldberger, Heritability. *Economica* **46**, 327 (1979). doi:10.2307/2553675
23. B. L. Browning, S. R. Browning, A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.* **84**, 210 (2009). doi:10.1016/j.ajhg.2009.01.005 Medline
24. B. Servin, M. Stephens, Imputation-based analysis of association studies: Candidate regions and quantitative traits. *PLoS Genet.* **3**, e114 (2007). doi:10.1371/journal.pgen.0030114 Medline
25. J. Marchini, B. Howie, S. Myers, G. McVean, P. Donnelly, A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39**, 906 (2007). doi:10.1038/ng2088 Medline
26. Y. Li, C. J. Willer, J. Ding, P. Scheet, G. R. Abecasis, MaCH: Using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet. Epidemiol.* **34**, 816 (2010). doi:10.1002/gepi.20533 Medline
27. S. Purcell *et al*., PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559 (2007). doi:10.1086/519795 Medline
28. B. Devlin, K. Roeder, Genomic control for association studies. *Biometrics* **55**, 997 (1999). doi:10.1111/j.0006-341X.1999.00997.x Medline
29. C. J. Willer, Y. Li, G. R. Abecasis, METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190 (2010). doi:10.1093/bioinformatics/btq340 Medline
30. SCAN, SNP and CNV Annotation Database (2012); www.scandb.org/
31. M. L. Freedman *et al*., Assessing the impact of population stratification on genetic association studies. *Nat. Genet.* **36**, 388 (2004). doi:10.1038/ng1333 Medline
32. P. I. W. de Bakker *et al*., Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum. Mol. Genet.* **17**, R122 (2008). doi:10.1093/hmg/ddn288 Medline
33. P. Taubman, Earnings, education, genetics, and environment. *J. Hum. Resour.* **11**, 447 (1976). doi:10.2307/145426 Medline
34. A. R. Branigan, K. J. McCallum, J. Freese, Variation in the heritability of educational attainment: An international meta-analysis. *Northwestern University Institute for Policy Research Working Paper.* **13-09** (2013).
35. D. Cesarini, *Essays on Genetic Variation and Economic Behavior* (Massachusetts Institute of Technology, 2010).
36. P. Lichtenstein *et al*., Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *N. Engl. J. Med.* **343**, 78 (2000). doi:10.1056/NEJM200007133430201 Medline
37. E. Turkheimer, Three laws of behavior genetics and what they mean. *Curr. Dir. Psychol. Sci.* **9**, 160 (2000). doi:10.1111/1467-8721.00084
38. J. Yang *et al*., Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565 (2010). doi:10.1038/ng.608 Medline
39. C. E. Ross, C. Wu, The links between education and health. *Am. Sociol. Rev.* **60**, 719 (1995). doi:10.2307/2096319
40. D. M. Cutler, A. Lleras-Muney, in *Making Americans Healthier: Social and Economic Policy as Health Policy*, J. House, R. Schoeni, G. Kaplan, H. Pollack, Eds. (Russell Sage Foundation, New York, 2008).
41. W. Johnson *et al*., Does education confer a culture of healthy behavior? Smoking and drinking patterns in Danish twins. *Am. J. Epidemiol.* **173**, 55 (2011). doi:10.1093/aje/kwq333 Medline
42. W. Johnson *et al*., Education reduces the effects of genetic susceptibilities to poor physical health. *Int. J. Epidemiol.* **39**, 406 (2010). doi:10.1093/ije/dyp314 Medline
43. A. P. Vermeiren *et al*., Do genetic factors contribute to the relation between education and metabolic risk factors in young adults? A twin study. *Eur. J. Public Health* 10.1093/eurpub/cks167 (2012). doi:10.1093/eurpub/cks167 Medline
44. A. Lleras-Muney, The relationship between education and adult mortality in the United States. *Rev. Econ. Stat.* **72**, 189 (2005). doi:10.1111/0034-6527.00329
45. A. C. J. Lager, J. Torssander, Causal effect of education on mortality in a quasi-experiment on 1.2 million Swedes. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8461 (2012). doi:10.1073/pnas.1105839109 Medline
46. J. N. Arendt, Does education cause better health? A panel data analysis using school reforms for identification. *Econ. Educ. Rev.* **24**, 149 (2005). doi:10.1016/j.econedurev.2004.04.008
47. T. Illig *et al*., A genome-wide perspective of genetic variation in human metabolism. *Nat. Genet.* **42**, 137 (2010). doi:10.1038/ng.507 Medline
48. J. Z. Liu *et al*.; AMFS Investigators, A versatile gene-based test for genome-wide association studies. *Am. J. Hum. Genet.* **87**, 139 (2010). doi:10.1016/j.ajhg.2010.06.009 Medline
49. S. H. Lee, J. Yang, M. E. Goddard, P. M. Visscher, N. R. Wray, Estimation of pleiotropy between complex diseases using using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* **28**, 2540 (2012). doi:10.1093/bioinformatics/bts474
50. G. Trynka *et al*., Chromatin marks identify critical cell types for fine mapping complex trait variants. *Nat. Genet.* **45**, 124 (2013). doi:10.1038/ng.2504 Medline
51. A. Cvejic *et al*., SMIM1 underlies the Vel blood group and influences red blood cell traits. *Nat. Genet.* **45**, 542 (2013). doi:10.1038/ng.2603 Medline
52. L. C. Andreae, A. Lumsden, J. D. Gilthorpe, Chick Lrrn2, a novel downstream effector of Hoxb1 and Shh, functions in the selective targeting of rhombomere 4 motor neurons. *Neural Dev.* **4**, 27 (2009). doi:10.1186/1749-8104-4-27 Medline
53. E. L. Heinzen *et al*., Tissue-specific genetic control of splicing: Implications for the study of complex traits. *PLoS Biol.* **6**, e1 (2008). doi:10.1371/journal.pbio.1000001 Medline
54. J. A. Webster *et al*.; NACC-Neuropathology Group, Genetic control of human brain transcript expression in Alzheimer disease. *Am. J. Hum. Genet.* **84**, 445 (2009). doi:10.1016/j.ajhg.2009.03.011 Medline
55. R. S. N. Fehrmann *et al*., Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. *PLoS Genet.* **7**, e1002197 (2011). doi:10.1371/journal.pgen.1002197 Medline
56. M. Nelis *et al*., Genetic structure of Europeans: A view from the North-East. *PLoS ONE* **4**, e5472 (2009). doi:10.1371/journal.pone.0005472 Medline

57. H. J. Westra *et al*., MixupMapper: Correcting sample mix-ups in genome-wide datasets increases power to detect small genetic effects. *Bioinformatics* **27**, 2104 (2011). doi:10.1093/bioinformatics/btr323 Medline
58. P. H. Lee, C. O'Dushlaine, B. Thomas, S. M. Purcell, INRICH: Interval-based enrichment analysis for genome-wide association studies. *Bioinformatics* **28**, 1797 (2012). doi:10.1093/bioinformatics/bts191 Medline
59. M. Ashburner *et al*.; The Gene Ontology Consortium, Gene ontology: Tool for the unification of biology. *Nat. Genet.* **25**, 25 (2000). doi:10.1038/75556 Medline
60. C. M. Koch *et al*., The landscape of histone modifications across 1% of the human genome in five human cell lines. *Genome Res.* **17**, 691 (2007). doi:10.1101/gr.5704207 Medline
61. A. C. Need *et al*., A genome-wide study of common SNPs and CNVs in cognitive performance in the CANTAB. *Hum. Mol. Genet.* **18**, 4650 (2009). doi:10.1093/hmg/ddp413 Medline
62. M. W. Logue *et al*.; Multi-Institutional Research on Alzheimer Genetic Epidemiology (MIRAGE) Study Group, A comprehensive genetic association study of Alzheimer disease in African Americans. *Arch. Neurol.* **68**, 1569 (2011). doi:10.1001/archneurol.2011.646 Medline
63. J. Burroughs-Garcia, V. Sittaramane, A. Chandrasekhar, S. T. Waters, Evolutionarily conserved function of Gbx2 in anterior hindbrain development. *Dev. Dyn.* **240**, 828 (2011). doi:10.1002/dvdy.22589 Medline
64. K. M. Wassarman *et al*., Specification of the anterior hindbrain and establishment of a normal mid/hindbrain organizer is dependent on Gbx2 gene function. *Development* **124**, 2923 (1997). Medline
65. L. Chen, M. Chatterjee, J. Y. Li, The mouse homeobox gene Gbx2 is required for the development of cholinergic interneurons in the striatum. *J. Neurosci.* **30**, 14824 (2010). doi:10.1523/JNEUROSCI.3742-10.2010 Medline
66. M. Muers, Complex disease: Ups and downs at the MHC. *Nat. Rev. Genet.* **12**, 456 (2011). doi:10.1038/nrg3021 Medline
67. D. Migliorini *et al*., Mdm4 (Mdmx) regulates p53-induced growth arrest and neuronal cell death during early embryonic mouse development. *Mol. Cell. Biol.* **22**, 5527 (2002). doi:10.1128/MCB.22.15.5527-5538.2002 Medline
68. J. A. Grahn, J. A. Parkinson, A. M. Owen, The cognitive functions of the caudate nucleus. *Prog. Neurobiol.* **86**, 141 (2008). doi:10.1016/j.pneurobio.2008.09.004 Medline
69. W. D. Altrock *et al*., Functional inactivation of a fraction of excitatory synapses in mice deficient for the active zone protein bassoon. *Neuron* **37**, 787 (2003). doi:10.1016/S0896-6273(03)00088-6 Medline
70. P. R. Burton *et al*.; Wellcome Trust Case Control Consortium, Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661 (2007). doi:10.1038/nature05911 Medline
71. M. Parkes *et al*.; Wellcome Trust Case Control Consortium, Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat. Genet.* **39**, 830 (2007). doi:10.1038/ng2061 Medline
72. J. C. Barrett *et al*.; UK IBD Genetics Consortium; Wellcome Trust Case Control Consortium 2, Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. *Nat. Genet.* **41**, 1330 (2009). doi:10.1038/ng.483 Medline
73. A. Franke *et al*., Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat. Genet.* **42**, 1118 (2010). doi:10.1038/ng.717 Medline
74. J. C. Barrett *et al*.; NIDDK IBD Genetics Consortium; Belgian-French IBD Consortium; Wellcome Trust Case Control Consortium, Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat. Genet.* **40**, 955 (2008). doi:10.1038/ng.175 Medline
75. L. Jostins *et al*.; International IBD Genetics Consortium (IIBDGC), Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119 (2012). doi:10.1038/nature11582 Medline
76. D. P. McGovern *et al*.; NIDDK IBD Genetics Consortium, Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nat. Genet.* **42**, 332 (2010). doi:10.1038/ng.549 Medline
77. C. A. Anderson *et al*., Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat. Genet.* **43**, 246 (2011). doi:10.1038/ng.764 Medline
78. M. Imielinski *et al*.; Western Regional Alliance for Pediatric IBD; International IBD Genetics Consortium; NIDDK IBD Genetics Consortium; Belgian-French IBD Consortium; Wellcome Trust Case Control Consortium, Common variants at five new loci associated with early-onset inflammatory bowel disease. *Nat. Genet.* **41**, 1335 (2009). doi:10.1038/ng.489 Medline
79. E. A. Stahl *et al*.; BIRAC Consortium; YEAR Consortium, Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat. Genet.* **42**, 508 (2010). doi:10.1038/ng.582 Medline
80. A. Ferguson, D. M. Sedgwick, J. Drummond, Morbidity of juvenile onset inflammatory bowel disease: Effects on education and employment in early adult life. *Gut* **35**, 665 (1994). doi:10.1136/gut.35.5.665 Medline
81. L. M. Mackner, D. P. Sisson, W. V. Crandall, Review: Psychosocial issues in pediatric inflammatory bowel disease. *J. Pediatr. Psychol.* **29**, 243 (2004). doi:10.1093/jpepsy/jsh027 Medline
82. K. A. Frazer *et al*.; International HapMap Consortium, A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851 (2007). doi:10.1038/nature06258 Medline
83. J. Yang *et al*.; Genetic Investigation of ANthropometric Traits (GIANT) Consortium; DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369, S1 (2012). doi:10.1038/ng.2213 Medline
84. J. Yang, S. H. Lee, M. E. Goddard, P. M. Visscher, GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76 (2011). doi:10.1016/j.ajhg.2010.11.011 Medline
85. H. D. Daetwyler, B. Villanueva, J. A. Woolliams, Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLoS ONE* **3**, e3395 (2008). doi:10.1371/journal.pone.0003395 Medline
86. B. J. Hayes, P. M. Visscher, M. E. Goddard, Increased accuracy of artificial selection by using the realized relationship matrix. *Genet. Res.* **91**, 47 (2009). doi:10.1017/S0016672308009981 Medline
87. M. E. Goddard, N. R. Wray, K. Verbyla, P. M. Visscher, Estimating effects and making predictions from genome-wide marker data. *Stat. Sci.* **24**, 517 (2009). doi:10.1214/09-STS306
88. P. M. Visscher, J. Yang, M. E. Goddard, A commentary on 'common SNPs explain a large proportion of the heritability for human height' by Yang et al. (2010). *Twin Res. Hum. Genet.* **13**, 517 (2010). doi:10.1375/twin.13.6.517 Medline
89. R. G. Fryer, Financial incentives and student achievement: Evidence from randomized trials. *Q. J. Econ.* **126**, 1755 (2011). doi:10.1093/qje/qjr045
90. J. Heckman, S. H. Moon, R. Pinto, P. Savelyev, A. Yavitz, Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quant. Econ.* **1**, 1 (2010). doi:10.3982/QE8 Medline
91. J. Eckenrode *et al*., Long-term effects of prenatal and infancy nurse home visitation on the life course of youths: 19-year follow-up of a randomized trial. *Arch. Pediatr. Adolesc. Med.* **164**, 9 (2010). doi:10.1001/archpediatrics.2009.240 Medline
92. L. N. Masse, W. S. Barnett, A benefit-cost analysis of the Abecedarian early childhood intervention. *Cost-Effectiveness and Educational Policy, Larchmont, NY: Eye on Education, Inc.*, 157-173 (2002).
93. J. J. Heckman, S. H. Moon, R. Pinto, P. A. Savelyev, A. Yavitz, The rate of return to the HighScope Perry Preschool Program. *J. Public Econ.* **94**, 114 (2010). doi:10.1016/j.jpubeco.2009.11.001 Medline
94. T. B. Harris *et al*., Age, Gene/Environment Susceptibility–Reykjavik Study: Multidisciplinary applied phenomics. *Am. J. Epidemiol.* **165**, 1076 (2007). doi:10.1093/aje/kwk115 Medline
95. A. Fraser *et al*., Cohort profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int. J. Epidemiol.* **42**, 97 (2012). 10.1093/ije/dys066 Medline
96. R. Schmidt *et al*., Assessment of cerebrovascular risk profiles in healthy persons: Definition of research goals and the Austrian Stroke Prevention Study (ASPS). *Neuroepidemiology* **13**, 308 (1994). doi:10.1159/000110396 Medline
97. R. Schmidt, F. Fazekas, P. Kapeller, H. Schmidt, H. P. Hartung, MRI white matter hyperintensities: Three-year follow-up of the Austrian Stroke Prevention Study. *Neurology* **53**, 132 (1999). doi:10.1212/WNL.53.1.132 Medline
98. N. W. Shock *et al*., Normal human aging: The Baltimore Longitudinal Study of Aging. *NIH Publication* **84-2450** (1984).
99. K. Einarsdóttir *et al*., Linkage disequilibrium mapping of *CHEK2*: Common variation and breast cancer risk. *PLoS Med.* **3**, e168 (2006).

doi:10.1371/journal.pmed.0030168 Medline
100. E. T. Chang, M. Hedelin, H. O. Adami, H. Grönberg, K. A. Bälter, Alcohol drinking and risk of localized versus advanced and sporadic versus familial prostate cancer in Sweden. *Cancer Causes Control* **16**, 275 (2005). doi:10.1007/s10552-004-3364-2 Medline
101. M. Hedelin *et al*., Dietary phytoestrogen, serum enterolactone and risk of prostate cancer: The cancer prostate Sweden study (Sweden). *Cancer Causes Control* **17**, 169 (2006). doi:10.1007/s10552-005-0342-2 Medline
102. F. Lindmark *et al*., H6D polymorphism in macrophage-inhibitory cytokine-1 gene associated with prostate cancer. *J. Natl. Cancer Inst.* **96**, 1248 (2004). doi:10.1093/jnci/djh227 Medline
103. M. Firmann *et al*., The CoLaus study: A population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. *BMC Cardiovasc. Disord.* **8**, 6 (2008). doi:10.1186/1471-2261-8-6 Medline
104. I. Rudan *et al*., "10001 Dalmatians:" Croatia launches its national biobank. *Croat. Med. J.* **50**, 4 (2009). doi:10.3325/cmj.2009.50.4 Medline
105. G. B. Ehret *et al*.; International Consortium for Blood Pressure Genome-Wide Association Studies; CARDIoGRAM consortium; CKDGen Consortium; KidneyGen Consortium; EchoGen consortium; CHARGE-HF consortium, Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* **478**, 103 (2011). doi:10.1038/nature10405 Medline
106. K. Sleegers *et al*., Cerebrovascular risk factors do not contribute to genetic variance of cognitive function: The ERF study. *Neurobiol. Aging* **28**, 735 (2007). doi:10.1016/j.neurobiolaging.2006.03.012 Medline
107. F. A. Sayed-Tabatabaei *et al*., Heritability of the function and structure of the arterial wall: Findings of the Erasmus Rucphen Family (ERF) study. *Stroke* **36**, 2351 (2005). doi:10.1161/01.STR.0000185719.66735.dd Medline
108. E. Vartiainen *et al*., Thirty-five-year trends in cardiovascular risk factors in Finland. *Int. J. Epidemiol.* **39**, 504 (2010). doi:10.1093/ije/dyp330 Medline
109. J. Kaprio, L. Pulkkinen, R. J. Rose, Genetic and environmental factors in health-related behaviors: Studies on Finnish twins and twin families. *Twin Res.* **5**, 366 (2002). Medline
110. S. M. Purcell *et al*.; International Schizophrenia Consortium, Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748 (2009). Medline
111. FBPP Investigators, Multi-center genetic study of hypertension: The Family Blood Pressure Program (FBPP). *Hypertension* **39**, 3 (2002). doi:10.1161/hy1201.100415 Medline
112. T. B. Harris *et al*., Waist circumference and sagittal diameter reflect total body fat better than visceral fat in older men and women. The Health, Aging and Body Composition Study. *Ann. N. Y. Acad. Sci.* **904**, 462 (2000). doi:10.1111/j.1749-6632.2000.tb06501.x Medline
113. D. J. P. Barker, C. Osmond, T. J. Forsén, E. Kajantie, J. G. Eriksson, Trajectories of growth among children who have coronary events as adults. *N. Engl. J. Med.* **353**, 1802 (2005). doi:10.1056/NEJMoa044160 Medline
114. L. Ferrucci *et al*., Subsystems contributing to the decline in ability to walk: bridging the gap between epidemiology and geriatric practice in the InCHIANTI study. *J. Am. Geriatr. Soc.* **48**, 1618 (2000). Medline
115. H.-E. Wichmann, C. Gieger, R. Illig; MONICA/KORA Study Group, KORA-gen - Resource for population genetics, controls and a broad specturm of disease phenotypes. *Gesundheitswesen* **67**, 26 (2005). doi:10.1055/s-2005-858226 Medline
116. R. P. Stolk *et al*., Universal risk factors for multifactorial diseases. LifeLines: A three-generation population-based study. *Eur. J. Epidemiol.* **23**, 67 (2008). doi:10.1007/s10654-007-9204-4 Medline
117. I. J. Deary, M. C. Whiteman, J. M. Starr, L. J. Whalley, H. C. Fox, The impact of childhood intelligence on later life: following up the Scottish mental surveys of 1932 and 1947. *J. Pers. Soc. Psychol.* **86**, 130 (2004). doi:10.1037/0022-3514.86.1.130 Medline
118. I. J. Deary *et al*., The Lothian Birth Cohort 1936: A study to examine influences on cognitive ageing from age 11 to age 70 and beyond. *BMC Geriatr.* **7**, 28 (2007). doi:10.1186/1471-2318-7-28 Medline
119. P. Magnus *et al*.; MoBa Study Group, Cohort profile: The Norwegian Mother and Child Cohort Study (MoBa). *Int. J. Epidemiol.* **35**, 1146 (2006). doi:10.1093/ije/dyl170 Medline
120. L. M. Irgens, The Medical Birth Registry of Norway. Epidemiological research and surveillance throughout 30 years. *Acta Obstet. Gynecol. Scand.* **79**, 435 (2000). doi:10.1080/j.1600-0412.2000.079006435.x Medline
121. B. W. J. H. Penninx *et al*.; NESDA Research Consortium, The Netherlands Study of Depression and Anxiety (NESDA): Rationale, objectives and methods. *Int. J. Methods Psychiatr. Res.* **17**, 121 (2008). doi:10.1002/mpr.256 Medline
122. P. Rantakallio, Groups at risk in low birth weight infants and perinatal mortality. *Acta Paediatr. Scand.* **193** (suppl.), 193, 1 (1969). Medline
123. C. Sabatti *et al*., Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat. Genet.* **41**, 35 (2009). doi:10.1038/ng.271 Medline
124. N. W. Martin *et al*., Educational attainment: A genome wide association study in 9538 Australians. *PLoS ONE* **6**, e20128 (2011). doi:10.1371/journal.pone.0020128 Medline
125. K. Estrada *et al*., GRIMP: A web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. *Bioinformatics* **25**, 2750 (2009). doi:10.1093/bioinformatics/btp497 Medline
126. A. Hofman *et al*., The Rotterdam Study: 2012 objectives and design update. *Eur. J. Epidemiol.* **26**, 657 (2011). doi:10.1007/s10654-011-9610-5 Medline
127. D. A. Bennett *et al*., Overview and findings from the rush Memory and Aging Project. *Curr. Alzheimer Res.* **9**, 646 (2012). Medline
128. L. J. Bierut *et al*.; Gene, Environment Association Studies Consortium, A genome-wide association study of alcohol dependence. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 5082 (2010). doi:10.1073/pnas.0911109107 Medline
129. G. Pilia *et al*., Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet.* **2**, e132 (2006). doi:10.1371/journal.pgen.0020132 Medline
130. H. Völzke *et al*., Cohort profile: The study of health in Pomerania. *Int. J. Epidemiol.* **40**, 294 (2011). doi:10.1093/ije/dyp394 Medline
131. P. K. E. Magnusson *et al*., The Swedish Twin Registry: Establishment of a biobank and other recent developments. *Twin Res. Hum. Genet.* **16**, 317 (2013). doi:10.1017/thg.2012.104 Medline
132. A. Moayyeri, C. J. Hammond, A. M. Valdes, T. D. Spector, Cohort profile: TwinsUK and Healthy Ageing Twin Study. *Int. J. Epidemiol.* **42**, 76 (2013). Medline
133. O. T. Raitakari *et al*., Cohort profile: The cardiovascular risk in Young Finns Study. *Int. J. Epidemiol.* **37**, 1220 (2008). doi:10.1093/ije/dym225 Medline
134. V. Pfaffenrath *et al*., Regional variations in the prevalence of migraine and tension-type headache applying the new IHS criteria: The German DMKG Headache Study. *Cephalalgia* **29**, 48 (2009). doi:10.1111/j.1468-2982.2008.01699.x Medline
135. M. M. Vennemann, T. Hummel, K. Berger, The association between smoking and smell and taste impairment in the general population. *J. Neurol.* **255**, 1121 (2008). doi:10.1007/s00415-008-0807-9 Medline
136. A. Aromaa, Health and functional capacity in Finland: Baseline results of the Health 2000 health examination survey. *Kansanterveyslaitos Folkhälsoinstitutet National Public Health Institute Kansanterveyslaitoksen Julkaisuja B12.* (2004).
137. M. McEvoy *et al*., Cohort profile: The Hunter Community Study. *Int. J. Epidemiol.* **39**, 1452 (2010). doi:10.1093/ije/dyp343 Medline
138. D. Weir, in *Biosocial Surveys*, Committee on Advances in Collecting and Utilizing Biological Indicators and Genetic Information in Social Science Surveys, M. Weinstein, J. W. Vaupel, K. W. Wachter, Eds. (2007), pp. 78, chap. 4.
139. M. B. Miller *et al*., The Minnesota Center for Twin and Family Research genome-wide association study. *Twin Res. Hum. Genet.* **15**, 767 (2012). doi:10.1017/thg.2012.62 Medline
140. J. H. Lee, R. Cheng, N. Graff-Radford, T. Foroud, R. Mayeux; National Institute on Aging Late-Onset Alzheimer's Disease Family Study Group, Analyses of the National Institute on Aging late-onset Alzheimer's disease family study: Implication of additional loci. *Arch. Neurol.* **65**, 1518 (2008). doi:10.1001/archneur.65.11.1518 Medline
141. D. I. Boomsma *et al*., Netherlands Twin Register: From twins to twin families. *Twin Res. Hum. Genet.* **9**, 849 (2006). doi:10.1375/twin.9.6.849 Medline
142. R. McQuillan *et al*., Runs of homozygosity in European populations. *Am. J. Hum. Genet.* **83**, 359 (2008). doi:10.1016/j.ajhg.2008.08.007 Medline
143. E. V. Theodoraki *et al*., Fibrinogen beta variants confer protection against coronary artery disease in a Greek case-control study. *BMC Med. Genet.* **11**, 28 (2010). doi:10.1186/1471-2350-11-28 Medline

144. S. Mukherjee *et al*., Cohort profile: The Western Australian Sleep Health Study. *Sleep Breath.* **16**, 205 (2012). doi:10.1007/s11325-011-0491-3 Medline
145. L. A. Baker, S. A. Treloar, C. A. Reynolds, A. C. Heath, N. G. Martin, Genetics of educational attainment in Australian twins: Sex differences and secular changes. *Behav. Genet.* **26**, 89 (1996). doi:10.1007/BF02359887 Medline
146. P. Miller, C. Mulvey, N. Martin, The return to schooling: Estimates from a sample of young Australian twins. *Labour Econ.* **13**, 571 (2006). doi:10.1016/j.labeco.2004.10.008
147. K. Silventoinen, R. F. Krueger, T. J. Bouchard Jr., J. Kaprio, M. McGue, Heritability of body height and educational attainment in an international context: Comparison of adult twins in Minnesota and Finland. *Am. J. Hum. Biol.* **16**, 544 (2004). doi:10.1002/ajhb.20060 Medline
148. A. C. Heath *et al*., Education policy and the heritability of educational attainment. *Nature* **314**, 734 (1985). doi:10.1038/314734a0 Medline
149. G. Isacsson, Estimating the economic return to educational levels using data on twins. *J. Appl. Econ.* **19**, 99 (2004). doi:10.1002/jae.724
150. P. Taubman, The determinants of earnings: Genetics, family, and other environments: A study of white male twins. *Am. Econ. Rev.* **66**, 858 (1976).
151. D. T. Lykken, T. J. Bouchard Jr., M. McGue, A. Tellegen, The Minnesota Twin Family Registry: Some initial findings. *Acta Genet. Med. Gemellol. (Roma)* **39**, 35 (1990). Medline
152. J. R. Behrman, P. Taubman, T. Wales, in Kinometrics: Determinants of Socioeconomic Success Within and Between Families (North-Holland Publishing Company, New York, 1977), pp. 35.
153. M. Soler Artigas *et al*.; GIANT consortium, Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. *Nat. Genet.* **43**, 1082 (2011). doi:10.1038/ng.941 Medline
154. T. Thye *et al*.; African TB Genetics Consortium; Wellcome Trust Case Control Consortium, Genome-wide association analyses identifies a susceptibility locus for tuberculosis on chromosome 18q11.2. *Nat. Genet.* **42**, 739 (2010). doi:10.1038/ng.639 Medline
155. J. R. Shaffer *et al*., GWAS of dental caries patterns in the permanent dentition. *J. Dent. Res.* **92**, 38 (2013). doi:10.1177/0022034512463579 Medline
156. R. A. Eeles *et al*.; UK Genetic Prostate Cancer Study Collaborators/British Association of Urological Surgeons' Section of Oncology; UK ProtecT Study Collaborators; PRACTICAL Consortium, Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. *Nat. Genet.* **41**, 1116 (2009). doi:10.1038/ng.450 Medline
157. N. M. Pajewski *et al*., A genome-wide association study of host genetic determinants of the antibody response to Anthrax Vaccine Adsorbed. *Vaccine* **30**, 4778 (2012). doi:10.1016/j.vaccine.2012.05.032 Medline
158. N. Sandholm *et al*.; DCCT/EDIC Research Group, New susceptibility loci associated with kidney disease in type 1 diabetes. *PLoS Genet.* **8**, e1002921 (2012). doi:10.1371/journal.pgen.1002921 Medline
159. B. Benyamin *et al*., Variants in *TF* and *HFE* explain ~40% of genetic variation in serum-transferrin levels. *Am. J. Hum. Genet.* **84**, 60 (2009). doi:10.1016/j.ajhg.2008.11.011 Medline
160. R. Qayyum *et al*., A meta-analysis and genome-wide association study of platelet count and mean platelet volume in african americans. *PLoS Genet.* **8**, e1002491 (2012). doi:10.1371/journal.pgen.1002491 Medline
161. C. Gieger *et al*., New gene functions in megakaryopoiesis and platelet formation. *Nature* **480**, 201 (2011). doi:10.1038/nature10659 Medline
162. C. S. Fox *et al*.; GIANT Consortium; MAGIC Consortium; GLGC Consortium, Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. *PLoS Genet.* **8**, e1002695 (2012). doi:10.1371/journal.pgen.1002695 Medline
163. M. Kolz *et al*.; EUROSPAN Consortium; ENGAGE Consortium; PROCARDIS Consortium; KORA Study; WTCCC, Meta-analysis of 28,141 individuals identifies common variants within five new loci that influence uric acid concentrations. *PLoS Genet.* **5**, e1000504 (2009). doi:10.1371/journal.pgen.1000504 Medline
164. M. Man *et al*., Beyond single-marker analyses: Mining whole genome scans for insights into treatment responses in severe sepsis. *Pharmacogenomics J.* **13**, 218 (2012). Medline
165. J. E. Landers *et al*., Reduced expression of the Kinesin-Associated Protein 3 (KIFAP3) gene increases survival in sporadic amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9004 (2009). doi:10.1073/pnas.0812937106 Medline
166. C. Cotsapas *et al*.; GIANT Consortium, Common body mass index-associated variants confer risk of extreme obesity. *Hum. Mol. Genet.* **18**, 3502 (2009). doi:10.1093/hmg/ddp292 Medline
167. K. Nakabayashi *et al*., Identification of independent risk loci for Graves' disease within the MHC in the Japanese population. *J. Hum. Genet.* **56**, 772 (2011). doi:10.1038/jhg.2011.99 Medline
168. B. Kestenbaum *et al*., Common genetic variants associate with serum phosphorus concentration. *J. Am. Soc. Nephrol.* **21**, 1223 (2010). doi:10.1681/ASN.2009111104 Medline
169. M. J. Barber *et al*., Genome-wide association of lipid-lowering response to statins in combined study populations. *PLoS ONE* **5**, e9763 (2010). doi:10.1371/journal.pone.0009763 Medline
170. A. I. Yashin, D. Wu, K. G. Arbeev, S. V. Ukraintseva, Joint influence of small-effect genetic variants on human longevity. *Aging* **2**, 612 (2010). Medline
171. G. Thorleifsson *et al*., Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nat. Genet.* **41**, 18 (2009). doi:10.1038/ng.274 Medline
172. C. J. Willer *et al*.; Wellcome Trust Case Control Consortium; Genetic Investigation of ANthropometric Traits Consortium, Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat. Genet.* **41**, 25 (2009). doi:10.1038/ng.287 Medline
173. E. Melum *et al*., Genome-wide association analysis in primary sclerosing cholangitis identifies two non-HLA susceptibility loci. *Nat. Genet.* **43**, 17 (2011). doi:10.1038/ng.728 Medline
174. J. M. Robins, S. Greenland, Identifiability and exchangeability for direct and indirect effects. *Epidemiology* **3**, 143 (1992). doi:10.1097/00001648-199203000-00013 Medline
175. M. J. H. M. van der Loos *et al*., The molecular genetic architecture of self-employment. *PLoS ONE* **8**, e60542 (2013). doi:10.1371/journal.pone.0060542 Medline

Acknowledgments: This research was carried out under the auspices of the Social Science Genetic Association Consortium (SSGAC), a cooperative enterprise among medical researchers and social scientists that coordinates genetic association studies for social science variables. Data for our analyses come from many studies and organizations, some of which are subject to an MTA (*5*). Results from the meta-analysis are available at the website of the consortium, www.ssgac.org. The formation of the SSGAC was made possible by an EAGER grant from the NSF and a supplemental grant from the NIH/OBSSR (SES-1064089). This research was also funded in part by the Söderbergh Foundation (E9/11), the NIA/NIH through grants P01-AG005842, P01-AG005842-20S2, P30-AG012810, and T32-AG000186-23 and the Intramural Research Program of the NIA/NIH. For a full list of acknowledgments, see (*5*).

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1235488/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S22
Tables S1 to S27
References (*23–175*)

22 January 2013; accepted 16 May 2013
Published online 30 May 2013
10.1126/science.1235488

Cornelius A. Rietveld,[1,2] Sarah E. Medland,[3] Jaime Derringer,[4] Jian Yang,[5] Tõnu Esko,[6] Nicolas W. Martin,[3,7] Harm-Jan Westra,[8] Konstantin Shakhbazov,[5,9] Abdel Abdellaoui,[10] Arpana Agrawal,[11] Eva Albrecht,[12] Behrooz Z. Alizadeh,[13] Najaf Amin,[14] John Barnard,[15] Sebastian E. Baumeister,[16] Kelly S. Benke,[17] Lawrence F. Bielak,[18] Jeffrey A. Boatman,[19] Patricia A. Boyle,[20] Gail Davies,[21] Christiaan de Leeuw,[22] Niina Eklund,[24,25] Daniel S. Evans,[26] Rudolf Ferhmann,[8] Krista Fischer,[6] Christian Gieger,[12] Håkon K. Gjessing,[27] Sara Hägg,[28,29,30] Jennifer R. Harris,[27] Caroline Hayward,[31] Christina Holzapfel,[32,33] Carla A. Ibrahim-Verbaas,[14,34] Erik Ingelsson,[28,29,30] Bo Jacobsson,[27,35] Peter K. Joshi,[36] Astanand Jugessur,[27] Marika Kaakinen,[37,38] Stavroula Kanoni,[39] Juha Karjalainen,[8] Ivana Kolcic,[40] Kati Kristiansson,[24,25] Zoltán Kutalik,[41,42] Jari Lahti,[43] Sang H. Lee,[3] Peng Lin,[11] Penelope

A. Lind,[3] Yongmei Liu,[44] Kurt Lohman,[45] Marisa Loitfelder,[46] George McMahon,[47] Pedro Marques Vidal,[48] Osorio Meirelles,[49] Lili Milani,[6] Ronny Myhre,[27] Marja-Liisa Nuotio,[24,25] Christopher J. Oldmeadow,[50] Katja E. Petrovic,[51] Wouter J. Peyrot,[52] Ozren Polašek,[40] Lydia Quaye,[53] Eva Reinmaa,[6] John P. Rice,[11] Thais S. Rizzi,[22] Helena Schmidt,[54] Reinhold Schmidt,[46] Albert V. Smith,[55,56] Jennifer A. Smith,[18] Toshiko Tanaka,[49] Antonio Terracciano,[49,57] Matthijs J.H.M. van der Loos,[1,2] Veronique Vitart,[31] Henry Völzke,[16] Jürgen Wellmann,[58] Lei Yu,[20] Wei Zhao,[18] Jüri Allik,[59] John R. Attia,[50] Stefania Bandinelli,[60] François Bastardot,[61] Jonathan Beauchamp,[62] David A. Bennett,[20] Klaus Berger,[58] Laura J. Bierut,[11] Dorret I. Boomsma,[10] Ute Bültmann,[63] Harry Campbell,[36] Christopher F. Chabris,[64] Lynn Cherkas,[53] Mina K. Chung,[15] Francesco Cucca,[65,66] Mariza de Andrade,[67] Philip L. De Jager,[68] Jan-Emmanuel De Neve,[69,70] Ian J. Deary,[21,71] George V. Dedoussis,[72] Panos Deloukas,[39] Maria Dimitriou,[72] Gudny Eiriksdottir,[55] Martin F. Elderson,[73] Johan G. Eriksson,[74,75,76,77] David M. Evans,[78] Jessica D. Faul,[79] Luigi Ferrucci,[49] Melissa E. Garcia,[49] Henrik Grönberg,[30] Vilmundur Gudnason,[55,56] Per Hall,[30] Juliette M. Harris,[53] Tamara B. Harris,[49] Nicholas D. Hastie,[31] Andrew C. Heath,[80] Dena G. Hernandez,[49] Wolfgang Hoffmann,[16] Adriaan Hofman,[81] Rolf Holle,[83] Elizabeth G. Holliday,[50] Jouke-Jan Hottenga,[10] William G. Iacono,[82] Thomas Illig,[33,84] Marjo-Riitta Järvelin,[37,38,85,86,87] Mika Kähönen,[88] Jaakko Kaprio,[24,89,90] Robert M. Kirkpatrick,[82] Matthew Kowgier,[91] Antti Latvala,[89,90] Lenore J. Launer,[49] Debbie A. Lawlor,[78] Terho Lehtimäki,[92] Jingmei Li,[93] Paul Lichtenstein,[30] Peter Lichtner,[94] David C. Liewald,[21] Pamela A. Madden,[11] Patrik K. E. Magnusson,[30] Tomi E. Mäkinen,[95] Marco Masala,[65] Matt McGue,[82] Andres Metspalu,[6] Andreas Mielck,[83] Michael B. Miller,[82] Grant W. Montgomery,[3] Sutapa Mukherjee,[96,97,98] Dale R. Nyholt,[3] Ben A. Oostra,[14] Lyle J. Palmer,[91] Aarno Palotie,[24,39,99] Brenda Penninx,[52] Markus Perola,[24,25,6] Patricia A. Peyser,[18] Martin Preisig,[61] Katri Räikkönen,[43] Olli T. Raitakari,[100,101] Anu Realo,[59] Susan M. Ring,[47] Samuli Ripatti,[24,25,39] Fernando Rivadeneira,[2,102] Igor Rudan,[36] Aldo Rustichini,[103] Veikko Salomaa,[104] Antti-Pekka Sarin,[24] David Schlessinger,[49] Rodney J. Scott,[50] Harold Snieder,[13] Beate St Pourcain,[78,105] John M. Starr,[21,106] Jae Hoon Sul,[107] Ida Surakka,[24,25] Rauli Svento,[108] Alexander Teumer,[109] The LifeLines Cohort Study,[Error! Reference source not found.] Henning Tiemeier,[2,110] Frank J.A. an Rooij,[2] David R. Van Wagoner,[15] Erkki Vartiainen,[111] Jorma Viikari,[112] Peter Vollenweider,[61] Judith M. Vonk,[13] Gérard Waeber,[61] David R. Weir,[79] H.-Erich Wichmann,[113,114,115] Elisabeth Widen,[24] Gonneke Willemsen,[10] James F. Wilson,[36] Alan F. Wright,[31] Dalton Conley,[116] George Davey-Smith,[78] Lude Franke,[8] Patrick J. F. Groenen,[121] Albert Hofman,[2] Magnus Johannesson,[122] Sharon L.R. Kardia,[18] Robert F. Krueger,[82] David Laibson,[117] Nicholas G. Martin,[3] Michelle N. Meyer,[118,119] Danielle Posthuma,[22,110,120] A. Roy Thurik,[1,123,124] Nicholas J. Timpson,[78] André G. Uitterlinden,[2,102] Cornelia M. van Duijn,[14,125] Peter M. Visscher,[3,5,9]*† Daniel J. Benjamin,[126]*† David Cesarini,[127,128,129]*† Philipp D. Koellinger[1,2]*†

[1]Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, 3000 DR Rotterdam, The Netherlands. [2]Department of Epidemiology, Erasmus Medical Center, Rotterdam 3000 CA, The Netherlands. [3]Queensland Institute of Medical Research, 300 Herston Road, Brisbane, Queensland 4006, Australia. [4]Institute for Behavioral Genetics, University of Colorado Boulder, Boulder, CO 80309–0447, USA. [5]University of Queensland Diamantina Institute, The University of Queensland, Princess Alexandra Hospital, Brisbane, Queensland 4102, Australia. [6]Estonian Genome Center, University of Tartu, Tartu 51010, Estonia. [7]School of Psychology, University of Queensland, Brisbane, Queensland 4072, Australia. [8]Department of Genetics, University Medical Center Groningen, University of Groningen, 9713 GZ Groningen, The Netherlands. [9]Queensland Brain Institute, The University of Queensland, Brisbane, Queensland 4072, Australia. [10]Department of Biological Psychology, VU University Amsterdam, 1081 BT Amsterdam, The Netherlands. [11]Department of Psychiatry, Washington University School of Medicine, St. Louis, MO 63110, USA. [12]Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany. [13]Department of Epidemiology, University Medical Center Groningen, University of Groningen, 9700 RB Groningen, The Netherlands. [14]Genetic Epidemiology Unit, Department of Epidemiology, Erasmus Medical Center, Rotterdam 3000 CA, the Netherlands. [15]Heart and Vascular and Lerner Research Institutes, Cleveland Clinic, Cleveland, OH 44195, USA. [16]Institute for Community Medicine, University Medicine Greifswald, Greifswald 17489, Germany. [17]Samuel Lunenfeld Research Institute, Mount Sinai Hospital, University of Toronto, Toronto, Ontario M5G 1X5, Canada. [18]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI 48109–2029, USA. [19]Division of Biostatistics, University of Minnesota, Minneapolis, MN 55455, USA. [20]Rush University Medical Center, Rush Alzheimer's Disease Center, Chicago, IL 60612, USA. [21]Centre for Cognitive Aging and Cognitive Epidemiology, The University of Edinburgh, Edinburgh EH8 9JZ, Scotland, UK. [22]Department of Functional Genomics, VU University Amsterdam and VU Medical Center, 1081 HV Amsterdam, The Netherlands. [23]Machine Learning Group, Intelligent Systems, Institute for Computing and Information Sciences, Faculty of Science, Radboud University Nijmegen, 6500 GL Nijmegen, The Netherlands. [24]Institute for Molecular Medicine Finland, University of Helsinki, Helsinki 00014, Finland. [25]Public Health Genomics Unit, Department of Chronic Disease Prevention, The National Institute for Health and Welfare, Helsinki 00014, Finland. [26]California Pacific Medical Center Research Institute, San Francisco, CA 94107–1728, USA. [27]Department of Genes and Environment, Division of Epidemiology, Norwegian Institute of Public Health, Nydalen, N-0403 Oslo, Norway. [28]Molecular Epidemiology, Department of Medical Sciences, Uppsala University, 751 85 Uppsala, Sweden. [29]Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, 751 23 Uppsala, Sweden. [30]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 171 77 Stockholm, Sweden. [31]Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK. [32]Else Kroener-Fresenius-Centre for Nutritional Medicine, Technische Universität München, 81675 Munich, Germany. [33]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany. [34]Department of Neurology, Erasmus Medical Center, Rotterdam 3000 CA, The Netherlands. [35]Department of Obstetrics and Gynecology, Institute of Public Health, Sahlgrenska Academy, Sahgrenska University Hospital, Gothenburg, 413 45, Sweden. [36]Centre for Population Health Sciences, The University of Edinburgh, Edinburgh EH8 9AG, UK. [37]Institute of Health Sciences, University of Oulu, Oulu 90014, Finland. [38]Biocenter Oulu, University of Oulu, Oulu 90014, Finland. [39]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK. [40]Faculty of Medicine, University of Split, 21000 Split, Croatia. [41]Department of Medical Genetics, University of Lausanne, 1005 Lausanne, Switzerland. [42]Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland. [43]Institute of Behavioral Sciences, University of Helsinki, Helsinki 00014, Finland. [44]Department of Epidemiology & Prevention, Division of Public Health Sciences, Wake Forest University Health Sciences, Winston-Salem, NC 27157–1063, USA. [45]Department of Biostatistical Sciences, Division of Public Health Sciences, Wake Forest University Health Sciences, Winston-Salem, NC 27157–1063, USA. [46]Division for Neurogeriatrics, Department of Neurology, Medical University of Graz, Graz 8036, Austria. [47]School of Social and Community Medicine, University of Bristol, Bristol BS8 2PR, UK. [48]Institute of Social and Preventive Medicine, Lausanne University Hospital, 1005 Lausanne, Switzerland. [49]National Institute on Aging, National Institutes of Health, Baltimore, MD 20892, USA. [50]Hunter Medical Research Institute and Faculty of Health, University of Newcastle, Newcastle, NSW 2308, Australia. [51]Division of General Neurology, Department of Neurology, General Hospital and Medical University of Graz, Graz 8036, Austria. [52]Department of Psychiatry, VU University Medical Center, 1081 HL Amsterdam, The Netherlands. [53]Department of Twin Research and Genetic Epidemiology, King's College London, London SE1 7EH, UK. [54]Institute of Molecular Biology and Biochemistry, Medical University of Graz, Graz 8036, Austria. [55]Icelandic Heart Association, Kopavogur 201, Iceland. [56]Department of Medicine, University of Iceland, Reykjavik 101, Iceland. [57]College of Medicine, Florida State University, Tallahassee, FL 32306–4300, USA. [58]Institute of Epidemiology and Social Medicine, University of Muenster, 48129 Muenster, Germany. [59]Department of Psychology, University of Tartu, Tartu 50410, Estonia. [60]Geriatric Unit, Azienda Sanitaria Firenze, 50125 Florence, Italy. [61]Department of Internal Medicine, University Hospital, 1011 Lausanne, Switzerland. [62]Department of Economics, Harvard University, Cambridge, MA 02138, USA. [63]Department of Health Sciences, Community & Occupational Medicine, University Medical Center Groningen, 9700 AD Groningen, The Netherlands. [64]Department of Psychology, Union College, Schenectady, NY 12308, USA. [65]Istituto di Ricerca Genetica e Biomedica, CNR, Monserrato, 09042, Cagliari,

Italy. [66]Dipartimento di Scienze Biomediche, Università di Sassari, 07100 SS, Italy. [67]Department of Health Sciences Research, Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN 55905, USA. [68]Program in Translational Neuropsychiatric Genomics, Department of Neurology, Brigham and Women's Hospital, Boston, MA 02115, USA. [69]School of Public Policy, University College London, London WC1H 9QU, UK. [70]Centre for Economic Performance, London School of Economics, London WC2A 2AE, UK. [71]Department of Psychology, The University of Edinburgh, Edinburgh EH8 9JZ, Scotland, UK. [72]Department of Nutrition and Dietetics, Harokopio University of Athens, Athens 17671, Greece. [73]LifeLines Cohort Study, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands. [74]Department of General Practice and Primary Health Care, University of Helsinki, Helsinki 00014, Finland. [75]Unit of General Practice, Helsinki University Central Hospital, Helsinki 00280, Finland. [76]Folkhälsan Research Center, Helsinki 00250, Finland. [77]Vaasa Central Hospital, Vaasa 65130, Finland. [78]MRC Centre for Causal Analyses in Translational Epidemiology, School of Social and Community Medicine, University of Bristol, Bristol BS8 2PR, UK. [79]Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI 48106, USA. [80]Division of Biology and Biomedical Sciences, Washington University, St. Louis, MO 63110–1093, USA. [81]Faculty of Behavioral and Social Sciences, University of Groningen, 9747 AD Groningen, The Netherlands. [82]Department of Psychology, University of Minnesota, Minneapolis, MN 55455–0344, USA. [83]Institute of Health Economics and Health Care Management, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany. [84]Hannover Unified Biobank, Hannover Medical School, 30625 Hannover, Germany. [85]Department of Epidemiology and Biostatistics, MRC-HPA Centre for Environment and Health, Imperial College London, London W2 1PG, UK. [86]Unit of Primary Care, Oulu University Hospital, Oulu 90220, Finland. [87]Department of Children and Young People and Families, National Institute for Health and Welfare, Oulu 90101, Finland. [88]Department of Clinical Physiology, Tampere University Hospital and University of Tampere School of Medicine, Tampere 33520, Finland. [89]Department of Public Health, University of Helsinki, 00014 Helsinki, Finland. [90]Department of Mental Health and Substance Abuse Services, National Institute for Health and Welfare, 00300 Helsinki, Finland. [91]Ontario Institute for Cancer Research, Toronto, Ontario M5G 0A3, Canada. [92]Department of Clinical Chemistry, Fimlab Laboratories, Tampere University Hospital, Tampere 33520, Finland. [93]Human Genetics, Genome Institute of Singapore, Singapore 138672, Singapore. [94]Institute of Human Genetics, Helmholtz Centre Munich, German Research Center for Environmental Health, 85764 Neuherberg, Germany. [95]Department of Health, Functional Capacity and Welfare, National Institute for Health and Welfare, Helsinki 00271, Finland. [96]Western Australia Sleep Disorders Research Institute, Sir Charles Gairdner Hospital, Perth, Western Australia 6009, Australia. [97]Department of Medicine, University of Toronto, Toronto, Ontario M5S 1A8, Canada. [98]Women's College Research Institute, University of Toronto, Toronto, Ontario M5G 1N8, Canada. [99]Department of Medical Genetics, University of Helsinki, 00014 Helsinki, Finland. [100]Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku 20520, Finland. [101]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku 20520, Finland. [102]Department of Internal Medicine, Erasmus Medical Center, Rotterdam 3000 CA, The Netherlands. [103]Department of Economics, University of Minnesota, Minneapolis, MN 55455–0462, USA. [104]Chronic Disease Epidemiology Unit, Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki 00271, Finland. [105]School of Oral and Dental Sciences, University of Bristol, Bristol BS1 2LY, UK. [106]Alzheimer Scotland Dementia Research Centre, The University of Edinburgh, Edinburgh EH8 9JZ, Scotland, UK. [107]Department of Computer Science, University of California, Los Angeles, CA 90095, USA. [108]Department of Economics, Oulu Business School, University of Oulu, Oulu 90014, Finland. [109]Interfaculty Institute for Genetics and Functional Genomics, Department of Functional Genomics, University Medicine Greifswald, Greifswald 17487, Germany. [110]Department of Child and Adolescent Psychiatry, Erasmus Medical Center, 3000 CB Rotterdam, The Netherlands. [111]Division of Welfare and Health Promotion, National Institute for Health and Welfare, Helsinki 00271, Finland. [112]Department of Medicine, Turku University Hospital, Turku 20520, Finland. [113]Institute of Medical Informatics, Biometry and Epidemiology, Chair of Epidemiology, Ludwig-Maximilians-Universität, 81377 Munich, Germany. [114]Klinikum Grosshadern, 81377 Munich, Germany. [115]Institute of Epidemiology I, Helmholtz Zentrum München, German Research Centre for Environmental Health, 85764 Neuherberg, Germany. [116]Department of Sociology, New York University, New York, NY 10012, USA. [117]Department of Economics, Harvard University, Cambridge, MA 02138, USA. [118]Petrie-Flom Center for Health Law Policy, Biotechnology, & Bioethics, Harvard Law School, Cambridge, MA 02138, USA. [119]Nelson A. Rockefeller Institute of Government, State University of New York, Albany, NY 12203–1003, USA. [120]Department of Clinical Genetics, VU University Medical Centrer, 1081 BT Amsterdam, The Netherlands. [121]Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam 3000 DR, The Netherlands. [122]Department of Economics, Stockholm School of Economics, Stockholm 113 83, Sweden. [123]Panteia, Zoetermeer 2701 AA, Netherlands. [124]GSCM-Montpellier Business School, Montpellier 34185, France. [125]Centre for Medical Systems Biology, Leiden University Medical Center, 2300 RC Leiden, The Netherlands. [126]Department of Economics, Cornell University, Ithaca, NY 14853, USA. [127]Center for Experimental Social Science, Department of Economics, New York University, New York, NY 10012, USA. [128]Division of Social Science, New York University Abu Dhabi, PO Box 129188, Abu Dhabi, UAE. [129]Research Institute of Industrial Economics, Stockholm 102 15, Sweden.

*These authors contributed equally to this work.
†Corresponding author. E-mail: db468@cornell.edu (D.J.B.); dac12@nyu.edu (D.C.); koellinger@ese.eur.nl (P.D.K.); peter.visscher@uq.edu.au (P.M.V.)

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**	**3.50×10^{-8}**	**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	8.45×10^{-8}	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	1.15×10^{-7}	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	1.37×10^{-6}	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**	**8.24×10^{-8}**	**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**	**5.33×10^{-11}**	**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	2.12×10^{-7}	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	3.24×10^{-7}	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	5.55×10^{-7}	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	7.48×10^{-7}	0.420	1.038	1.59×10^{-3}	1.031	7.61×10^{-4}

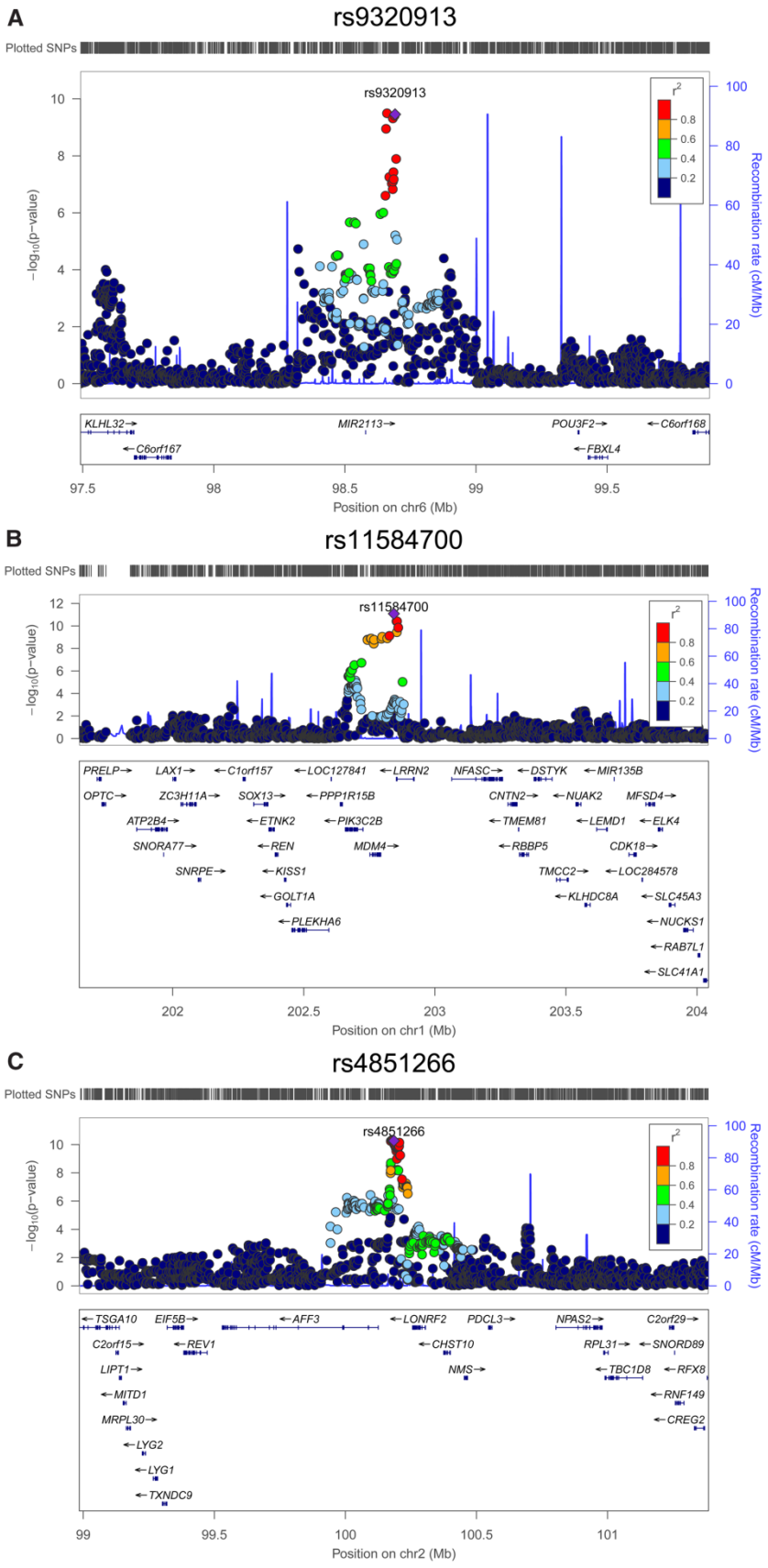

Fig. 1. Regional association plots of replicated loci associated with educational attainment [(**A**): rs9320913, (**B**): rs11584700, (**C**): rs4851266]. The plots are centered on the SNPs with the lowest *p*-values in the discovery stage (purple diamond). The R^2 values are from the CEU HapMap 2 samples. The CEU HapMap 2 recombination rates are indicated with a blue line on the right-hand *y*-axis. The figures were created with LocusZoom (http://csg.sph.umich.edu/locuszoom/).

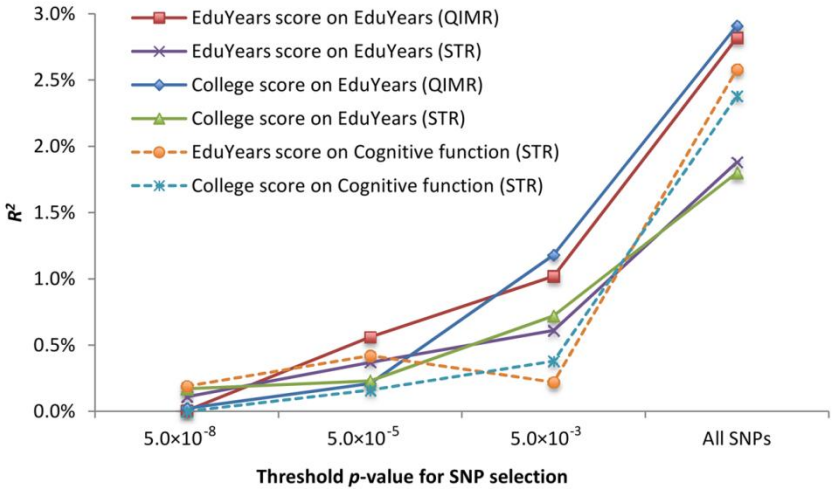

Fig. 2. Solid lines show results from regressions of *EduYears* on linear polygenic scores in a set of unrelated individuals from the QIMR (*N* = 3526) and STR (*N* = 6770) cohorts. Dashed lines show results from regressions of *Cognitive function* on linear polygenic scores in a sample from STR (*N*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