



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

Cornelius A. Rietveld *et al.*
Science **340**, 1467 (2013);
DOI: 10.1126/science.1235488

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of June 23, 2013):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/340/6139/1467.full.html>

Supporting Online Material can be found at:

<http://www.sciencemag.org/content/suppl/2013/05/29/science.1235488.DC1.html>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/content/340/6139/1467.full.html#related>

This article **cites 162 articles**, 38 of which can be accessed free:

<http://www.sciencemag.org/content/340/6139/1467.full.html#ref-list-1>

This article has been **cited by** 1 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/340/6139/1467.full.html#related-urls>

This article appears in the following **subject collections**:

Genetics

<http://www.sciencemag.org/cgi/collection/genetics>

24. S. Chutinimitkul *et al.*, *J. Virol.* **84**, 6825 (2010).
 25. Y. Gao *et al.*, *PLoS Pathog.* **5**, e1000709 (2009).
 26. M. Hatta, P. Gao, P. Halfmann, Y. Kawaoka, *Science* **293**, 1840 (2001).

Acknowledgments: This work was supported by the National 973 Project (grant 2011CB504703), the National Natural Science Foundation of China (NSFC, grant 81290342), and the China National Grand S&T Special Project (grant 2013ZX10004-611). G.F.G. is a leading principal investigator

of the NSFC Innovative Research Group (grant 81021003). Assistance by the staff at the Shanghai Synchrotron Radiation Facility (SSRF-beamline 17U) and the KEK Synchrotron Radiation Facility (beamline 17A) is acknowledged. We are grateful to the Consortium for Functional Glycomics for providing the biotinylated SA analogs. Coordinates and structure factors have been deposited in the Protein Data Bank (PDB code 4K62 for InH5, 4K63 for InH5/LSTa, 4K64 for InH5/LSTc, 4K65 for InH5mut, 4K66 for InH5mut/LSTa, and 4K67 for InH5mut/LSTc).

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1236787/DC1
 Materials and Methods
 Figs. S1 to S3
 Tables S1 and S2
 References (27–36)

19 February 2013; accepted 24 April 2013
 Published online 2 May 2013;
 10.1126/science.1236787

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 (GWAS) of educational attainment was conducted in a discovery sample of 101,069 individuals and a replication sample of 25,490. Three independent single-nucleotide polymorphisms (SNPs) are genome-wide significant (rs9320913, rs11584700, rs4851266), and all three replicate. Estimated effects sizes are small (coefficient of determination $R^2 \approx 0.02\%$), approximately 1 month of schooling per allele. A linear polygenic score from all measured SNPs accounts for $\approx 2\%$ 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 may lead 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 probably because existing work [for a 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 participants (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 (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 participants were very likely to have completed their education [more than 95% of the sample was at least 30 (5)]. On average, participants 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 stage, 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 10 loci that reached at least suggestive significance, we brought forward for replication the single-nucleotide polymorphism (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 five out of seven 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 coefficient of determination R^2 corresponds to a difference of ≈ 1 month 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 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' untranslated region (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, expression quantitative trait locus (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 with the use of 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. 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 coexpression analysis revealed that several implicated genes (including *BSN*, *GBX2*, *LRRN2*, and *PIK3C2B*) are probably 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

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. \hat{r}^2 represents the percent heterogeneity of effect size between the discovery-stage studies. P_{het} is the heterogeneity P value. bp, base pair.

SNP	Chr	Position (bp)	Nearest gene	Reference allele	Frequency	Discovery stage			Replication stage			Combined stage			Combined stage—sex-specific analysis		
						Beta/OR	P value	\hat{r}^2	P_{het}	Beta/OR	P value	P_{het}	Beta/OR (males)	P value (males)	Beta/OR (females)	P value (females)	
rs9320913	6	98691454	LOC100129158	A	0.483	4.19x10⁻⁹	18.3	0.097	0.077	0.012	0.101	3.50x10⁻¹⁰	0.350	0.095	1.87x10⁻⁴	0.100	1.43x10⁻⁶
rs3783006	13	97909210	STK24	C	0.454	2.29x10 ⁻⁷	0	0.982	0.056	0.055	0.088	8.45x10 ⁻⁸	0.959	0.064	1.44x10 ⁻²	0.108	3.35x10 ⁻⁷
rs8049439	16	28745016	ATXN2L	T	0.581	7.12x10 ⁻⁷	10.7	0.229	0.065	0.026	0.086	1.15x10 ⁻⁷	0.205	0.097	1.43x10 ⁻⁴	0.078	1.90x10 ⁻⁴
rs13188378	5	101958587	SLC06A1	A	0.878	7.49x10 ⁻⁷	0	0.791	0.091	0.914	-0.097	1.37x10 ⁻⁴	0.646	-0.134	8.21x10 ⁻³	-0.080	5.92x10 ⁻³
rs11584700	1	202843606	LRRN2	A	0.780	2.07x10⁻⁹	13.8	0.179	0.912	4.86x10⁻⁴	0.919	8.24x10⁻¹²	0.221	0.934	6.11x10⁻⁴	0.911	2.12x10⁻⁹
rs4851266	2	100184911	LOC150577	T	0.396	2.20x10⁻⁹	23.7	0.049	1.049	0.003	1.050	5.33x10⁻¹¹	0.072	1.054	1.55x10⁻⁵	1.052	6.74x10⁻⁸
rs2054125	2	199093966	PLCL1	T	0.064	5.55x10 ⁻⁸	7	0.325	1.098	0.225	1.376	2.12x10 ⁻⁷	0.268	1.264	1.74x10 ⁻²	1.503	1.95x10 ⁻⁷
rs3227	6	33770273	ITPR3	C	0.498	6.02x10 ⁻⁸	5	0.363	1.010	0.280	1.037	3.24x10 ⁻⁷	0.415	1.046	9.44x10 ⁻⁵	1.029	1.37x10 ⁻³
rs4073894	7	104254200	LHPL3	A	0.207	4.41x10 ⁻⁷	0	0.765	1.003	0.467	1.062	5.55x10 ⁻⁶	0.513	1.050	2.18x10 ⁻²	1.073	1.74x10 ⁻⁵
rs12640626	4	176863266	GPM6A	A	0.580	4.94x10 ⁻⁷	10.9	0.234	1.000	0.495	1.034	7.48x10 ⁻⁶	0.420	1.038	1.59x10 ⁻³	1.031	7.61x10 ⁻⁴

EduYears

College

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 [Queensland Institute of Medical Research (QIMR) and Swedish Twin Registry (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. Though it would have been preferable to explore a richer set of mediators, this variable was available in STR, a data set where we had access to the individual-level genotypic data. The Swedish Enlistment Battery (used for conscription) had

previously been administered to measure cognitive function in a subset of males (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 ge-

netic variance across individuals captured by current SNP microarrays. Using combined data from STR and QIMR, we estimate that this upper bound is 22.4% (SE = 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 more than an order of magnitude smaller than those observed for height and body mass index (BMI): 0.4% (15) and 0.3% (18), respectively. For comparison with the R^2 value of 2% from our linear polygenic score for education, 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 100 times larger than those we have found (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.

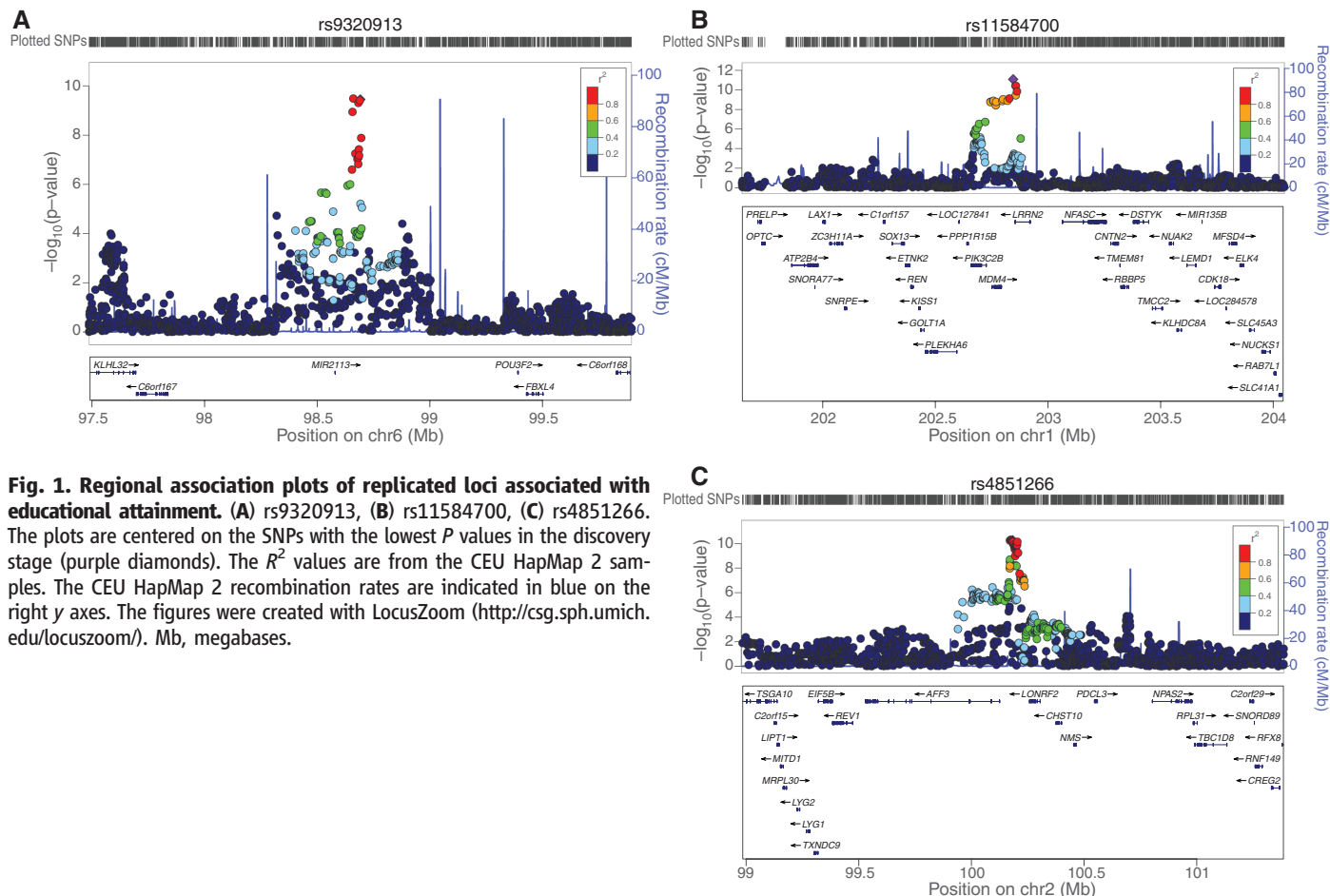


Fig. 1. Regional association plots of replicated loci associated with educational attainment. (A) rs9320913, (B) rs11584700, (C) rs4851266. The plots are centered on the SNPs with the lowest P values in the discovery stage (purple diamonds). The R^2 values are from the CEU HapMap 2 samples. The CEU HapMap 2 recombination rates are indicated in blue on the right y axes. The figures were created with LocusZoom (<http://csg.sph.umich.edu/locuszoom/>). Mb, megabases.

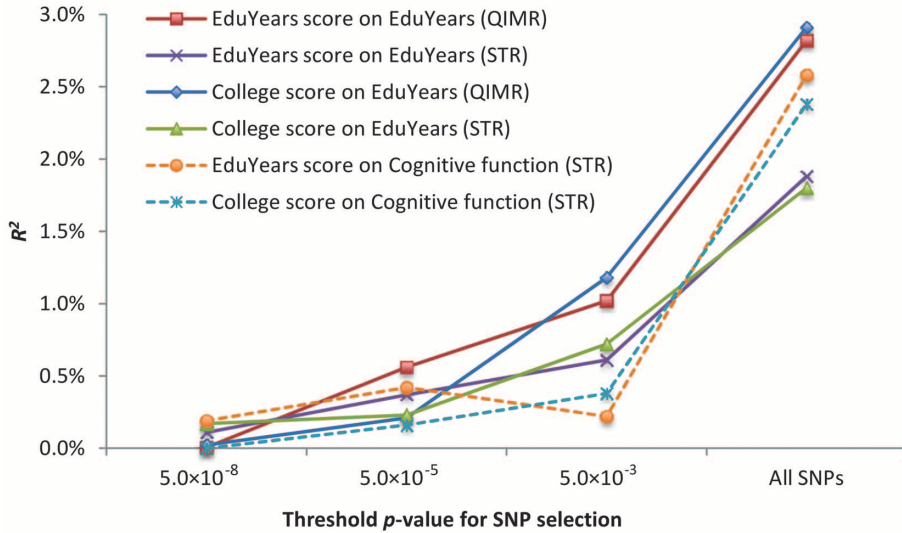


Fig. 2. Explanatory power of the linear polygenic scores estimated for EduYears or College. Solid lines show results from regressions of EduYears on linear polygenic scores in a set of unrelated individuals from the QIMR ($n = 3526$) and STR ($n = 6770$) cohorts. Dashed lines show results from regressions 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 cohort (either QIMR or STR) subsequently used as the prediction sample.

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).
2. D. Cesarini, C. T. Dawes, M. Johannesson, P. Lichtenstein, B. Wallace, *Q. J. Econ.* **124**, 809 (2009).

3. D. J. Benjamin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8026 (2012).
4. J. P. Beauchamp *et al.*, *J. Econ. Perspect.* **25**, 57 (2011).
5. See the supplementary materials on Science Online.
6. D. J. Benjamin *et al.*, *Annu. Rev. Econ.* **4**, 627 (2012).
7. R. P. Eibstein, S. Israel, S. H. Chew, S. Zhong, A. Knafo, *Neuron* **65**, 831 (2010).
8. L. E. Duncan, M. C. Keller, *Am. J. Psychiatry* **168**, 1041 (2011).
9. J. P. Ioannidis, *PLoS Med.* **2**, e124 (2005).
10. P. M. Visscher, M. A. Brown, M. I. McCarthy, J. Yang, *Am. J. Hum. Genet.* **90**, 7 (2012).
11. J. P. Mackenbach *et al.*; European Union Working Group on Socioeconomic Inequalities in Health, *N. Engl. J. Med.* **358**, 2468 (2008).
12. I. J. Deary, S. Strand, P. Smith, C. Fernandes, *Intelligence* **35**, 13 (2007).
13. J. J. Heckman, Y. Rubinstein, *Am. Econ. Rev.* **91**, 145 (2001).
14. UNESCO Institute for Statistics, *International Standard Classification of Education* (UNESCO Institute for Statistics, Montreal, 2006).
15. H. Lango Allen *et al.*, *Nature* **467**, 832 (2010).
16. J. Yang *et al.*; GIANT Consortium, *Eur. J. Hum. Genet.* **19**, 807 (2011).
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, *Nat. Genet.* **42**, 937 (2010).
19. M. H. de Moor *et al.*, *Mol. Psychiatry* **17**, 337 (2012).
20. B. Benyamin *et al.*; Wellcome Trust Case Control Consortium 2 (WTCCC2), *Mol. Psychiatry* **10.1038/mp.2012.184** (2013).
21. C. Jencks, *Am. Sociol. Rev.* **45**, 723 (1980).
22. A. S. Goldberger, *Economica* **46**, 327 (1979).

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 a materials transfer agreement (5). Results from the meta-analysis are available at the Web site 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 Office of Behavioral and Social Sciences Research (SES-1064089). This research was also funded in part by the Ragnar Söderberg Foundation (E9/11); the National Institute on Aging (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).

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 Lottfelder,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 Polasek,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,⁵ Stefania Bandinelli,⁶⁰ François Bastardot,⁶¹ Jonathan Beauchamp,⁶² David A. Bennett,²⁰ Klaus Berger,⁵⁰ Laura J. Bierut,¹¹ Dorret I. Boomsma,¹⁰ Ute Bültmann,⁶³ Harry Campbell,⁶³ 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 Guöny Eiriksdóttir,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 Vilumundur Guðnason,^{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} Milka Kähönen,⁸⁸ Jaakko Kaprio,^{24,89,90} Robert M. Kirkpatrick,⁸² Matthew Kowiger,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 W. J. H. Penninx,⁵² Markus Perola,^{24,25,6} Patricia A. Peyser,¹⁸ Martin Preisig,⁶¹ Katri Raitakinen,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,⁷³ Henning Tiemeier,^{2,110} Frank J. A. van 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 Willemssen,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. Kardina,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. Timponi,⁷⁸ André G. Uitterlinden,^{2,105} Cornelia M. van Duin,^{14,125} Peter M. Visscher,^{3,5,9*} Daniel J. Benjamin,^{126*}† David Cesarini,^{127,128,129*}† Philipp D. Koellinger,^{1,2,*}†

Downloaded from www.sciencemag.org on June 24, 2013

- ¹Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, 3000 DR Rotterdam, Netherlands. ²Department of Epidemiology, Erasmus Medical Center, Rotterdam 3000 CA, Netherlands. ³Queensland Institute of Medical Research, 300 Herston Road, Brisbane, Queensland 4006, Australia. ⁴Institute for Behavioral Genetics, University of Colorado Boulder, Boulder, CO 80309–0447, USA. ⁵University of Queensland Diamantina Institute, The University of Queensland, Princess Alexandra Hospital, Brisbane, Queensland 4102, Australia. ⁶Estonian Genome Center, University of Tartu, Tartu 51010, Estonia. ⁷School of Psychology, University of Queensland, Brisbane, Queensland 4072, Australia. ⁸Department of Genetics, University Medical Center Groningen, University of Groningen, 9713 GZ Groningen, Netherlands. ⁹Queensland Brain Institute, The University of Queensland, Brisbane, Queensland 4072, Australia. ¹⁰Department of Biological Psychology, VU University Amsterdam, 1081 BT Amsterdam, Netherlands. ¹¹Department of Psychiatry, Washington University School of Medicine, St. Louis, MO 63110, USA. ¹²Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany. ¹³Department of Epidemiology, University Medical Center Groningen, University of Groningen, 9700 RB Groningen, Netherlands. ¹⁴Genetic Epidemiology Unit, Department of Epidemiology, Erasmus Medical Center, Rotterdam 3000 CA, Netherlands. ¹⁵Heart and Vascular and Lerner Research Institutes, Cleveland Clinic, Cleveland, OH 44195, USA. ¹⁶Institute for Community Medicine, University Medicine Greifswald, Greifswald 17489, Germany. ¹⁷Samuel Lunenfeld Research Institute, Mount Sinai Hospital, University of Toronto, Toronto, Ontario M5G 1X5, Canada. ¹⁸Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI 48109–2029, USA. ¹⁹Division of Biostatistics, University of Minnesota, Minneapolis, MN 55455, USA. ²⁰Rush University Medical Center, Rush Alzheimer's Disease Center, Chicago, IL 60612, USA. ²¹Centre for Cognitive Aging and Cognitive Epidemiology, The University of Edinburgh, Edinburgh EH8 9JZ, UK. ²²Department of Functional Genomics, VU University Amsterdam and VU Medical Center, 1081 HV Amsterdam, Netherlands. ²³Machine Learning Group, Intelligent Systems, Institute for Computing and Information Sciences, Faculty of Science, Radboud University Nijmegen, 6500 GL Nijmegen, Netherlands. ²⁴Institute for Molecular Medicine Finland, University of Helsinki, Helsinki 00014, Finland. ²⁵Public Health Genomics Unit, Department of Chronic Disease Prevention, The National Institute for Health and Welfare, Helsinki 00014, Finland. ²⁶California Pacific Medical Center Research Institute, San Francisco, CA 94107–1728, USA. ²⁷Department of Genes and Environment, Division of Epidemiology, Norwegian Institute of Public Health, Nydalen, N-0403 Oslo, Norway. ²⁸Department of Medical Sciences, Molecular Epidemiology, Uppsala University, 751 85 Uppsala, Sweden. ²⁹Science for Life Laboratory, Uppsala University, 751 23 Uppsala, Sweden. ³⁰Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 171 77 Stockholm, Sweden. ³¹Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK. ³²Else Kroener-Fresenius-Centre for Nutritional Medicine, Technische Universität München, University Hospital "Klinikum rechts der Isar," 81675 Munich, Germany. ³³Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany. ³⁴Department of Neurology, Erasmus Medical Center, Rotterdam 3000 CA, Netherlands. ³⁵Department of Obstetrics and Gynecology, Institute of Public Health, Sahlgrenska Academy, Sahlgrenska University Hospital, Gothenburg, 413 45, Sweden. ³⁶Centre for Population Health Sciences, The University of Edinburgh, Edinburgh EH8 9AG, UK. ³⁷Institute of Health Sciences, University of Oulu, Oulu 90014, Finland. ³⁸Biocentre Oulu, University of Oulu, Oulu 90014, Finland. ³⁹Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK. ⁴⁰Faculty of Medicine, University of Split, 21000 Split, Croatia. ⁴¹Department of Medical Genetics, University of Lausanne, 1005 Lausanne, Switzerland. ⁴²Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland. ⁴³Institute of Behavioral Sciences, University of Helsinki, Helsinki 00014, Finland. ⁴⁴Department of Epidemiology and Prevention, Division of Public Health Sciences, Wake Forest University Health Sciences, Winston-Salem, NC 27157–1063, USA. ⁴⁵Department of Biostatistical Sciences, Division of Public Health Sciences, Wake Forest University Health Sciences, Winston-Salem, NC 27157–1063, USA. ⁴⁶Division for Neurogeriatrics, Department of Neurology, Medical University of Graz, Graz 8036, Austria. ⁴⁷School of Social and Community Medicine, University of Bristol, Bristol BS8 2PR, UK. ⁴⁸Institute of Social and Preventive Medicine, Lausanne University Hospital, 1005 Lausanne, Switzerland. ⁴⁹National Institute on Aging, National Institutes of Health, Baltimore, MD 20892, USA. ⁵⁰Hunter Medical Research Institute and Faculty of Health, University of Newcastle, Newcastle, NSW 2308, Australia. ⁵¹Division of General Neurology, Department of Neurology, General Hospital and Medical University of Graz, Graz 8036, Austria. ⁵²Department of Psychiatry, VU University Medical Center, 1081 HL Amsterdam, Netherlands. ⁵³Department of Twin Research and Genetic Epidemiology, King's College London, London SE1 7EH, UK. ⁵⁴Institute of Molecular Biology and Biochemistry, Medical University of Graz, Graz 8036, Austria. ⁵⁵Icelandic Heart Association, Kopavogur 201, Iceland. ⁵⁶Department of Medicine, University of Iceland, Reykjavik 101, Iceland. ⁵⁷College of Medicine, Florida State University, Tallahassee, FL 32306–4300, USA. ⁵⁸Institute of Epidemiology and Social Medicine, University of Muenster, 48129 Muenster, Germany. ⁵⁹Department of Psychology, University of Tartu, Tartu 50410, Estonia. ⁶⁰Geriatric Unit, Azienda Sanitaria Firenze, 50125 Florence, Italy. ⁶¹Department of Internal Medicine, University Hospital, 1011 Lausanne, Switzerland. ⁶²Department of Economics, Harvard University, Cambridge, MA 02138, USA. ⁶³Department of Health Sciences, Community and Occupational Medicine, University Medical Center Groningen, University of Groningen, 9700 AD Groningen, The Netherlands. ⁶⁴Department of Psychology, Union College, Schenectady, NY 12308, USA. ⁶⁵Istituto di Ricerca Genetica e Biomedica, Consejo de Investigación Nacional Italiano (CNR), Monserrato, 09042, Cagliari, Italy. ⁶⁶Dipartimento di Scienze Biomediche, Università di Sassari, 07100 SS, Italy. ⁶⁷Department of Health Sciences Research, Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN 55905, USA. ⁶⁸Program in Translational Neuropsychiatric Genomics, Department of Neurology, Brigham and Women's Hospital, Boston, MA 02115, USA. ⁶⁹School of Public Policy, University College London, London WC1H 9QU, UK. ⁷⁰Centre for Economic Performance, London School of Economics, London WC2A 2AE, UK. ⁷¹Department of Psychology, The University of Edinburgh, Edinburgh EH8 9JZ, UK. ⁷²Department of Nutrition and Dietetics, Harokopio University of Athens, Athens 17671, Greece. ⁷³LifeLines Cohort Study, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, Netherlands. ⁷⁴Department of General Practice and Primary Health Care, University of Helsinki, Helsinki 00014, Finland. ⁷⁵Unit of General Practice, Helsinki University Central Hospital, Helsinki 00280, Finland. ⁷⁶Folkhälsan Research Center, Helsinki 00250, Finland. ⁷⁷Vaasa Central Hospital, Vaasa 65130, Finland. ⁷⁸MRC Centre for Causal Analyses in Translational Epidemiology, School of Social and Community Medicine, University of Bristol, Bristol BS8 2PR, UK. ⁷⁹Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI 48106, USA. ⁸⁰Division of Biology and Biomedical Sciences, Washington University, St. Louis, MO 63110–1093, USA. ⁸¹Faculty of Behavioral and Social Sciences, University of Groningen, 9747 AD Groningen, Netherlands. ⁸²Department of Psychology, University of Minnesota, Minneapolis, MN 55455–0344, USA. ⁸³Institute of Health Economics and Health Care Management, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany. ⁸⁴Hannover Unified Biobank, Hannover Medical School, 30625 Hannover, Germany. ⁸⁵Department of Epidemiology and Biostatistics, MRC-HPA Centre for Environment and Health, Imperial College London, London W2 1PG, UK. ⁸⁶Unit of Primary Care, Oulu University Hospital, Oulu 90220, Finland. ⁸⁷Department of Children and Young People and Families, National Institute for Health and Welfare, Oulu 90101, Finland. ⁸⁸Department of Clinical Physiology, Tampere University Hospital and University of Tampere School of Medicine, Tampere 33520, Finland. ⁸⁹Department of Public Health, University of Helsinki, 00014 Helsinki, Finland. ⁹⁰Department of Mental Health and Substance Abuse Services, National Institute for Health and Welfare, 00300 Helsinki, Finland. ⁹¹Ontario Institute for Cancer Research, Toronto, Ontario M5G 0A3, Canada. ⁹²Department of Clinical Chemistry, Fimlab Laboratories, Tampere University Hospital, Tampere 33520, Finland. ⁹³Human Genetics, Genome Institute of Singapore, Singapore 138672, Singapore. ⁹⁴Institute of Human Genetics, Helmholtz Centre Munich, German Research Center for Environmental Health, 85764 Neuherberg, Germany. ⁹⁵Department of Health, Functional Capacity and Welfare, National Institute for Health and Welfare, Helsinki 00271, Finland. ⁹⁶Western Australia Sleep Disorders Research Institute, Sir Charles Gairdner Hospital, Perth, Western Australia 6009, Australia. ⁹⁷Department of Medicine, University of Toronto, Toronto, Ontario M5S 1A8, Canada. ⁹⁸Women's College Research Institute, University of Toronto, Toronto, Ontario M5G 1N8, Canada. ⁹⁹Department of Medical Genetics, University of Helsinki, 00014 Helsinki, Finland. ¹⁰⁰Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku 20520, Finland. ¹⁰¹Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku 20520, Finland. ¹⁰²Department of Internal Medicine, Erasmus Medical Center, Rotterdam 3000 CA, Netherlands. ¹⁰³Department of Economics, University of Minnesota, Minneapolis, MN 55455–0462, USA. ¹⁰⁴Chronic Disease Epidemiology Unit, Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki 00271, Finland. ¹⁰⁵School of Oral and Dental Sciences, University of Bristol, Bristol BS1 2LY, UK. ¹⁰⁶Alzheimer Scotland Dementia Research Centre, The University of Edinburgh, Edinburgh EH8 9JZ, UK. ¹⁰⁷Department of Computer Science, University of California, Los Angeles, CA 90095, USA. ¹⁰⁸Department of Economics, Oulu Business School, University of Oulu, Oulu 90014, Finland. ¹⁰⁹Interfaculty Institute for Genetics and Functional Genomics, Department of Functional Genomics, University Medicine Greifswald, Greifswald 17487, Germany. ¹¹⁰Department of Child and Adolescent Psychiatry, Erasmus Medical Center, 3000 CB Rotterdam, Netherlands. ¹¹¹Division of Welfare and Health Promotion, National Institute for Health and Welfare, Helsinki 00271, Finland. ¹¹²Department of Medicine, Turku University Hospital, Turku 20520, Finland. ¹¹³Institute of Medical Informatics, Biometry and Epidemiology, Chair of Epidemiology, Ludwig-Maximilians-Universität, 81377 Munich, Germany. ¹¹⁴Klinikum Grosshadern, 81377 Munich, Germany. ¹¹⁵Institute of Epidemiology I, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany. ¹¹⁶Department of Sociology, New York University, New York, NY 10012, USA. ¹¹⁷Department of Economics, Harvard University, Cambridge, MA 02138, USA. ¹¹⁸Petrie-Floem Center for Health Law Policy, Biotechnology, and Bioethics, Harvard Law School, Cambridge, MA 02138, USA. ¹¹⁹Nelson A. Rockefeller Institute of Government, State University of New York, Albany, NY 12203–1003, USA. ¹²⁰Department of Clinical Genetics, VU University Medical Center, 1081 BT Amsterdam, Netherlands. ¹²¹Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam 3000 DR, Netherlands. ¹²²Department of Economics, Stockholm School of Economics, Stockholm 113 83, Sweden. ¹²³Panteia, Zoetermeer 2701 AA, Netherlands. ¹²⁴GSCM-Montpellier Business School, Montpellier 34185, France. ¹²⁵Centre for Medical Systems Biology, Leiden University Medical Center, 2300 RC Leiden, Netherlands. ¹²⁶Department of Economics, Cornell University, Ithaca, NY 14853, USA. ¹²⁷Center for Experimental Social Science, Department of Economics, New York University, New York, NY 10012, USA. ¹²⁸Division of Social Science, New York University Abu Dhabi, PO Box 129188, Abu Dhabi, UAE. ¹²⁹Research 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.)

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