nature genetics

Population genetic differentiation of height and body mass index across Europe

Matthew R Robinson[1], Gibran Hemani[1], Carolina Medina-Gomez[2], Massimo Mezzavilla[3,4], Tonu Esko[5–8], Konstantin Shakhbazov[1], Joseph E Powell[1,9], Anna Vinkhuyzen[1], Sonja I Berndt[10], Stefan Gustafsson[11], Anne E Justice[12], Bratati Kahali[13,14], Adam E Locke[15], Tune H Pers[6–8,16], Sailaja Vedantam[6,7], Andrew R Wood[17], Wouter van Rheenen[18], Ole A Andreassen[19], Paolo Gasparini[3,4], Andres Metspalu[5], Leonard H van den Berg[18], Jan H Veldink[18], Fernando Rivadeneira[2], Thomas M Werge[20–22], Goncalo R Abecasis[15], Dorret I Boomsma[23–25], Daniel I Chasman[8,26], Eco J C de Geus[23–25], Timothy M Frayling[17], Joel N Hirschhorn[5–8], Jouke Jan Hottenga[23–25], Erik Ingelsson[11,27], Ruth J F Loos[28–31], Patrik K E Magnusson[32], Nicholas G Martin[33], Grant W Montgomery[33], Kari E North[13,14,34], Nancy L Pedersen[32], Timothy D Spector[35], Elizabeth K Speliotes[15], Michael E Goddard[36,37], Jian Yang[1,9] & Peter M Visscher[1,9]

Across-nation differences in the mean values for complex traits are common[1–8], but the reasons for these differences are unknown. Here we find that many independent loci contribute to population genetic differences in height and body mass index (BMI) in 9,416 individuals across 14 European countries. Using discovery data on over 250,000 individuals and unbiased effect size estimates from 17,500 sibling pairs, we estimate that 24% (95% credible interval (CI) = 9%, 41%) and 8% (95% CI = 4%, 16%) of the captured additive genetic variance for height and BMI, respectively, reflect population genetic differences. Population genetic divergence differed significantly from that in a null model (height, $P < 3.94 \times 10^{-8}$; BMI, $P < 5.95 \times 10^{-4}$), and we find an among-population genetic correlation for tall and slender individuals ($r = -0.80$, 95% CI = −0.95, −0.60), consistent with correlated selection for both phenotypes. Observed differences in height among populations reflected the predicted genetic means ($r = 0.51$; $P < 0.001$), but environmental differences across Europe masked genetic differentiation for BMI ($P < 0.58$).

Many of the phenotypes that vary within human populations are complex, in that they are determined by alleles at multiple loci and many non-genetic factors[9–15]. Therefore, it is reasonable to assume that regional differences in such traits have a complex basis[16–18]. Understanding these regional differences requires knowledge of the relative roles of environmental and genetic effects, which can be gained through estimating the amount of population genetic variance in phenotype and by determining the amount of observed differences that are explained by population genetic effects[19]. Thus far, these estimates have yet to be made outside of laboratory study populations[20], and experimental designs for human populations have been lacking because of confounding from genetic and environmental effects.

At least 135 million European citizens are obese[21], resulting in major direct and indirect health and economic costs[16,22,23]. Regional differences across Europe in height and susceptibility to weight gain, as defined by BMI, are well documented[5,18,22–26], but the reasons for these differences are not well understood. Height and BMI are extensively studied, and there is strong evidence that a large number of genetic polymorphisms influence both traits, with a considerable proportion of genetic variance captured by common SNPs[11,26–30]. For height, there is empirical evidence for selection at height-associated SNP loci within Europe[26,31] and between European populations and the rest of the world[32]. However, the true extent of population genetic differentiation is unlikely to have been captured or well represented by the limited number of ascertained loci examined thus far. For BMI, it is not known whether genetic differentiation exists, and, for both traits, the extent to which common loci contribute to population genetic variance or the observed regional phenotypic differences remains unknown (although see ref. 26).

Here we estimate cumulative population genetic differentiation for height and BMI captured by multiple unlinked loci across 9,416 Europeans from 14 countries, using population genetic analyses (Online Methods and **Supplementary Figs. 1** and **2**). We performed genome-wide association study (GWAS) meta-analyses on data from recent studies[33,34], selecting independent loci (linkage disequilibrium (LD) r^2 <0.1 and >1 Mb apart[35]) associated with either trait in a European-ancestry sample (~250,000 individuals for height and ~350,000 individuals for BMI). We reestimated the effects of each SNP in a within-family study design, which is unbiased by population stratification, and used these effect sizes to create a genetic predictor for both phenotypes (also termed a 'profile' or 'polygenic score')[36].

A full list of affiliations appears at the end of the paper.

Received 26 June; accepted 19 August; published online 14 September 2015; doi:10.1038/ng.3401

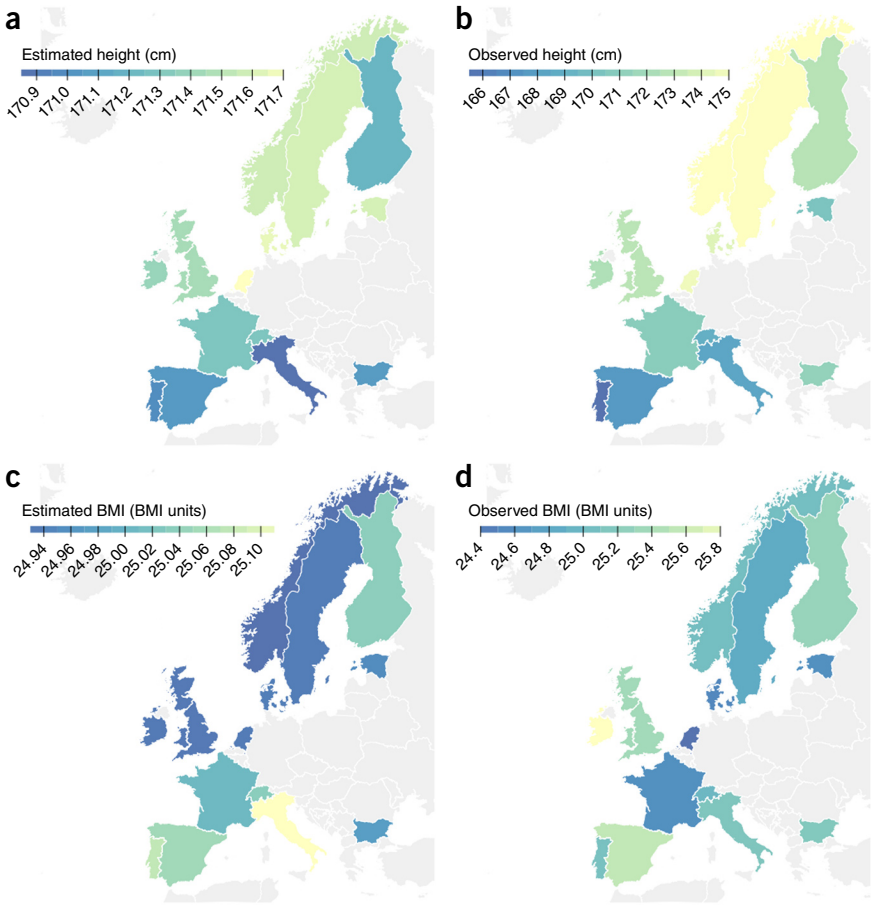

Figure 1 Observed divergence and predicted genetic divergence in height and BMI across 14 European nations. (**a**–**d**) The predicted genetic means (**a**,**c**) and observed means (**b**,**d**) for height and BMI for 14 European nations are shown across Europe. From recently published data, we estimated national differences in mean height and BMI for 14 European countries, accounting for trends over time, with a European average height of 171.1 cm (95% CI = 169.6, 172.8 cm) and an average BMI of 25.0 (95% CI = 24.7, 25.3) across nations for males between 2000 and 2010.

The proportion of variance in the genetic predictor attributable to population differences was estimated in a Bayesian mixed-effects model, alongside the co-differentiation of the phenotypes and the predicted means for each nation (Online Methods). Using theory and simulation study, we show that, whereas within-family effect sizes are unbiased, population genetic analyses conducted using loci ascertained from a standard GWAS can be biased if population stratification is not fully accounted for (Online Methods, **Supplementary Figs. 3–5** and **Supplementary Note**). There is no certainty that population stratification is completely controlled for in large-scale meta-analyses, and we thus repeated our analysis using (i) non-ascertained, unlinked (LD r^2 <0.1 and >1 Mb apart), common (minor allele frequency (MAF) >1%) HapMap 3 loci (~40,000 SNPs) and (ii) unlinked (LD r^2 <0.1 and >1 Mb apart), common (MAF >1%) HapMap 3 loci selected on the basis of their within-family association with each phenotype (~40,000 SNPs for both traits). This analysis provides an unbiased, genome-wide estimate, representing a lower limit of population genetic differentiation at common, unlinked loci.

The maximum proportion of variance in a polygenic predictor attributable to population genetic differences was 24% (95% CI = 9%, 41%) and 8% (95% CI = 4%, 16%) for height and BMI, respectively, using 2,660 SNPs for height and 11,919 SNPs for BMI. For height, the largest proportion of population-level variance was captured by SNPs of low *P* value in the meta-analysis (**Supplementary Fig. 6**). For BMI, the continual addition of SNPs increased the proportion of population-level variance captured (**Supplementary Fig. 6**). For both traits, the among-population variation was greater in predictors that explained greater phenotypic variance (**Supplementary Fig. 7**). Our results were confirmed using the non-ascertained independent, genome-wide loci (height: 8.6%, 95% CI = 3%, 15.7%; BMI: 2.8%, 95% CI = 1.1%, 5.3%) and the set of independent, genome-wide loci selected on the basis of their within-family association (height: 11.9%, 95% CI = 4.5%, 21.8%; BMI: 8%, 95% CI = 3.4%, 14.7%). The lower among-population variance captured using non-ascertained loci reflects reduced prediction accuracy, likely due to the addition of a large number of loci with no detectable association. Subsequent results are presented using the predictor for each trait that captured the greatest amount of population- and individual-level variance (comprising 2,660 SNPs for height and 11,919 SNPs for BMI); however, the results remained the same irrespective of the SNPs selected (**Supplementary Fig. 8**). The predicted population genetic means for the traits are shown in **Figure 1** alongside the observed values, estimated from an independent set of recently published data[25,37] accounting for trends over time. Among-population differences in allele frequency are expected to create genetic differences in height, such that people from the Netherlands are on average 1 cm taller than those from Italy, and genetic differences in BMI, such that, on average, the BMIs for people from Italy and Denmark differ by 0.2 units (**Figs. 1** and **2**).

Genetic differences among populations may occur by random, chance processes or through natural selection in the evolutionary past[19,38–46]. We thus compared our estimates to the values from a null quantitative genetic model of multivariate population differentiation[32,47]. We found strong evidence that the divergence of each trait was greater than expected under the neutral model (**Fig. 2**). The overall level of neutral genetic differentiation was small for both height (1.2%; 95% CI = 0.01%, 1.78%) and BMI (1.9%, 95% CI = 0.48%, 2.97), reflecting the average F_{ST} (a measure of population differentiation due to genetic structure; **Supplementary Note**) of the SNP sets between the populations of 1% for height and 1.2% for BMI. Our results were confirmed using non-ascertained independent, genome-wide loci (height, $P = 3.29 \times 10^{-6}$; BMI, $P = 0.018$) and independent, genome-wide loci selected on the basis of their within-family association (height, $P = 2.67 \times 10^{-6}$; BMI, $P = 8.35 \times 10^{-5}$). We therefore reject the null model, and our results suggest that population genetic differentiation across these 14 European countries for height and BMI has been driven by selection on standing genetic variation across geographical regions in the evolutionary past.

The significant departure from a neutral model occurs because, on average, the common loci comprising the genetic predictor are differentiated in a direction that is consistent with the direction of

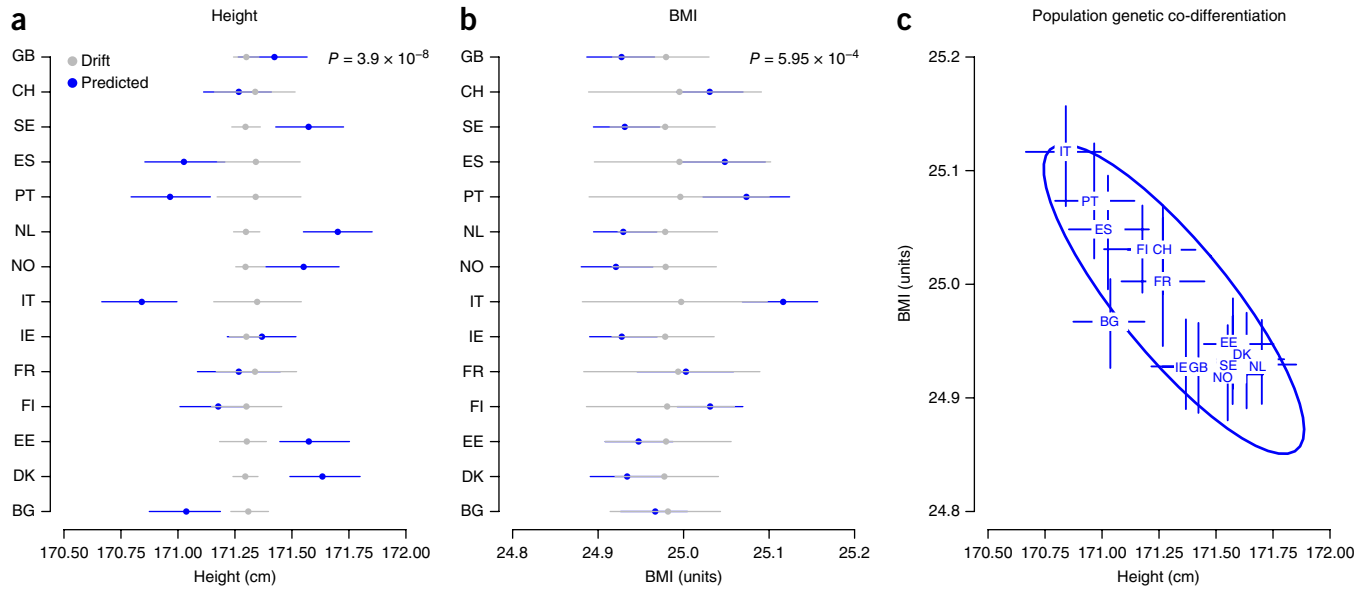

Figure 2 Predicted genetic differentiation compared to the expectation under genetic drift for height and BMI across 14 European nations. (**a**,**b**) The mean predicted genetic differentiation (blue) and differentiation under the null model representing genetic drift (gray) are shown for 14 European nations, with 95% credible intervals, for height (**a**) and BMI (**b**). ISO 2 country codes indicate each nation. The average *P* value of differentiation from the null expectation is shown in the figure. (**c**) The pattern of population-level co-differentiation for height and BMI across 14 European nations. The negative population genetic co-differentiation for these traits of –0.80 (95% CI = –0.95, –0.60) is represented by the blue ellipse.

their effects on each trait, which in turn creates differences among countries in a genetic predictor. Our interpretation does not depend on LD, as we used only independent loci, nor does it depend on the SNPs used, as we generated our null model estimates from the same set of SNPs that we used to create the genetic predictor. Loss of LD, genotype-environment interactions and genetic heterogeneity can change the relationships between SNPs and underlying causal variants across countries, but such changes would reduce the likelihood of detecting population genetic differentiation. In this study, we minimized these effects by using populations of European ancestry throughout.

We estimated the population genetic co-differentiation of height and BMI to ask whether selection has acted on both traits independently. We found a negative correlation between the population genetic means of –0.80 (95% CI = –0.95, –0.60; **Fig. 2**). This finding was consistent across predictors comprising non-ascertained genome-wide loci (–0.77, 95% CI = –0.94, –0.55) and independent genome-wide loci selected on the basis of their within-family association (–0.89, 95% CI = –0.97, –0.77). These results imply that selection has acted on common loci to increase height while reducing BMI and vice versa, and a genetic predisposition for tall stature at the population level was associated with a genetic predisposition for slenderness (low BMI). As height and BMI are nearly uncorrelated at the individual level (correlation among genetic profile scores within populations $r = -0.016$, 95% CI = –0.041, 0.001), selection for one trait should not elicit a response in the other. Our results suggest that selection has acted on both phenotypes, although, as some genes affect both phenotypes[48], we cannot rule out differentiation in one trait having been mediated by selection for the other. The population genetic co-divergence shown here is inconsistent with random genetic drift because the expectation with drift is that the among-population genetic correlation will equal the within-population correlation[47,49].

We tested whether the observed phenotypic differences across the 14 European countries reflect genetic differentiation at common loci or whether current environmental differences among countries (in diet, economy, climate, etc.) mask the population genetic

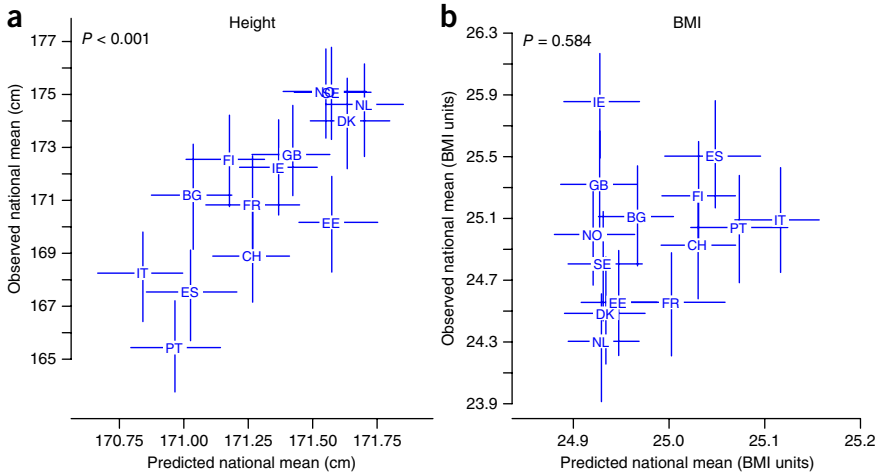

Figure 3 Association between observed population means and predicted genetic population means for height and BMI across 14 European nations. (**a**,**b**) Predicted population genetic means are plotted against observed population means for height (**a**) and BMI (**b**). *P* values give the significance of the multivariate Pearson product-moment correlation between the predicted population genetic means and the observed population means for both traits. For height, the correlation ($r = 0.51$, 95% CI = 0.39, 0.61) was greater than that expected under the null model ($r = 0.03$, 95% CI = –0.21, 0.17). For BMI, the correlation ($r = -0.10$, 95% CI = –0.19, 0.01) was not significantly different from the null expectation ($r = -0.08$, 95% CI = –0.24, 0.15).

differentiation that we detect. Our results show a strong association (r = 0.51, 95% CI = 0.39, 0.61; $P < 0.001$) between the population genetic values and the observed phenotypic pattern for height (**Fig. 3**), suggesting that the phenotypic differences that we observe across countries reflect differences in allele frequency at common height-associated loci. For BMI, the observed pattern did not reflect the pattern of population genetic differentiation (r = −0.10, 95% CI = −0.19, 0.01; P = 0.584; **Fig. 3**). This suggests that, although selection has created population genetic differentiation for BMI, environmental differences among countries mask this population genetic differentiation. We found no evidence of an association between the genetic differentiation expected under drift and the observed values (**Fig. 3**), implying that the observed national patterns do not reflect genetic drift.

We identified the loci that contributed the most to the population genetic differentiation for both traits. We found that the SNPs making the largest contributions to the phenotypic variance were enriched for association with the genome-wide pattern of population genetic differentiation (**Supplementary Figs. 9** and **10**), a pattern expected on the basis of our theory (**Supplementary Figs. 11** and **12**) and supported by the fact that the proportion of population-level variance was greater in a predictor explaining greater phenotypic variance. We found no evidence for significant among-population differentiation at any SNP across the genome, implying that many loci of small effect combine across the genome to create the detected genome-wide population genetic differentiation. Annotation of the 500 SNPs contributing the most to each trait with genes suggests that population genetic variation across the 14 countries for height and BMI is underlain by the combined effects of multiple pathways, with an overlap in the genes involved (**Supplementary Fig. 13**).

Finally, we examined population genetic differentiation in height and BMI on both a local and worldwide scale. We found no evidence for population genetic differentiation across six northern Italian villages[50] (**Supplementary Fig. 14**). We then examined population differentiation for height and BMI in the Human Genetic Diversity Panel (HGDP), as used in a previous study[32] (**Supplementary Fig. 15**). For both phenotypes, we found evidence to reject the null hypothesis that population genetic differentiation reflects neutrality (**Supplementary Fig. 15**), extending previous work that reported no significant differentiation for BMI using a limited number of loci[32]. Additionally, we found no evidence for population genetic co-differentiation of height and BMI, implying a European-specific pattern of selection.

The conclusions of our study are fourfold: (i) many common loci combine in a consistent manner to create population genetic differences for height and BMI; (ii) population genetic differentiation for height and BMI does not reflect a pattern expected under neutrality, and selection has thus driven the differences observed; (iii) population genetic divergences for height and BMI are correlated within Europe, with this correlation reaching a greater level than expected under neutrality, implying that the height-associated loci under selection are enriched for loci with effects that reduce BMI; and (iv) selection-driven population genetic differences for height reflect the phenotypic patterns we see across Europe, whereas, for BMI, environmental factors are masking the population genetic differences. Although genotype-environment effects and rare variants will have a role in shaping population genetic differentiation, the focus of our approach is on estimating the amount of population-level variance that is 'tagged' by a specific set of common SNP markers. As additional genetic variation is captured for both traits, it is likely that power will increase to fully capture the among-population genetic effects. The theoretical and analysis framework builds upon previous approaches[26,32], is entirely general and can be applied to estimate the role of commonly varying loci in shaping population differences for any set of phenotypes.

URLs. Online interactive version of **Figure 1**, http://www.kn3in.com/eu_traits/; imputation procedure, https://github.com/CNSGenomics/impute-pipe; R library biomaRt from Bioconductor, http://bioconductor.org/packages/release/bioc/html/biomaRt.html.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank the reviewers for their very helpful and insightful comments that greatly improved the manuscript. We also thank B. Hill and O. Ovaskainen for useful discussions. The University of Queensland group is supported by the Australian National Health and Medical Research Council (NHMRC; grants 1078037, 1048853 and 1050218). J.E.P. is supported by Australian Research Council grant DE130100691. J.Y. is supported by a Charles and Sylvia Viertel Senior Medical Research Fellowship and by NHMRC grant 1052684. We thank our colleagues at the Centre for Neurogenetics and Statistical Genomics for comments and suggestions. We are grateful to the twins and their families for their generous participation in the full-sibling family data set, which includes data from many cohorts and received support from many funding bodies. TWINGENE was supported by the Swedish Research Council (M-2005-1112), GenomEUtwin (EU/QLRT-2001-01254 and QLG2-CT-2002-01254), US National Institutes of Health (NIH) grant DK U01-066134, the Swedish Foundation for Strategic Research (SSF), and the Heart and Lung Foundation (20070481). For the Netherlands Twin Register (NTR), funding was obtained from the Netherlands Organization for Scientific Research (NWO; MagW/ZonMW grants 904-61-090, 985-10-002, 904-61-193, 480-04-004, 400-05-717, Addiction-31160008, Middelgroot-11-09-032 and Spinozapremie 56-464-14192), the Center for Medical Systems Biology (CSMB; NWO Genomics), NBIC/BioAssist/RK(2008.024), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL; 184.021.007), the VU University's Institute for Health and Care Research (EMGO+) and Neuroscience Campus Amsterdam (NCA), the European Science Foundation (ESF; EU/QLRT-2001-01254), the European Community's Seventh Framework Programme (FP7/2007-2013) under the ENGAGE project grant agreement (HEALTH-F4-2007-201413), the European Research Council (ERC Advanced; 230374), the Rutgers University Cell and DNA Repository (National Institute for Mental Health (NIMH), U24-MH068457-06), the Avera Institute (Sioux Falls, South Dakota, USA) and the US NIH (R01-D0042157-01A, Grand Opportunity grants 1RC2-MH089951-01 and 1RC2-MH089995-01). Part of the genotyping and analyses were funded by the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health. The TwinsUK study was funded by the Wellcome Trust and the European Community's Seventh Framework Programme (FP7/2007-2013) under the ENGAGE project grant agreement (HEALTH-F4-2007-201413). TwinsUK also receives support from the UK Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's and St Thomas' National Health Service (NHS) Foundation Trust in partnership with King's College London. T.D.S. is the holder of an ERC Advanced Principal Investigator award. Genotyping for the TwinsUK study was performed by the Wellcome Trust Sanger Institute, with the support of the National Eye Institute via a US NIH/Center for Inherited Disease Research (CIDR) genotyping project. The Framingham Heart Study is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with Boston University (contract N01-HC-25195). This manuscript was not prepared in collaboration with investigators of the Framingham Heart Study and does not necessarily reflect the opinions or views of the Framingham Heart Study, Boston University or NHLBI. Funding for SHARe Affymetrix genotyping was provided by NHLBI contract N02-HL-64278. Funding for SHARe Illumina genotyping was provided under an agreement between Illumina and Boston University. The QIMR researchers acknowledge funding from the Australian NHMRC (grants 241944, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 496688 and 552485) and the US NIH (grants AA07535, AA10248, AA014041, AA13320, AA13321, AA13326 and

DA12854). We are grateful to M. Gill (Trinity College Dublin) and K. Nicodemus (University of Edinburgh) for access to the ISC–Trinity College Dublin cohort, which was supported by the Wellcome Trust and the Health Research Board, Ireland. Access to the Bulgarian cohort data was kindly facilitated by G. Kirov and V. Excott-Price. For the Danish cohort, the Danish Scientific Committees and the Danish Data Protection Agency approved the study and all the patients gave written informed consent before inclusion in the project. The National Institute on Aging (NIA) provided funding for the Health and Retirement Study (HRS; U01-AG09740). The HRS is performed at the Institute for Social Research at the University of Michigan. This manuscript was not prepared in collaboration with investigators of the HRS and does not necessarily reflect the opinions or views of the HRS, University of Michigan or NIA. The Netherlands genotype samples were part of Project MinE, which was supported by the ALS Foundation Netherlands. Research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013).

AUTHOR CONTRIBUTIONS

Conception and design of the study: M.R.R., M.E.G., J.Y. and P.M.V. Data analysis: M.R.R., with additional contributions from G.H., C.M.-G., M.M., K.S., T.E., J.E.P., A.V., S.I.B., S.G., A.E.J., B.K., A.E.L., T.H.P., S.V., A.R.W. and W.v.R. Study oversight, sample collection and management: J.H.V., L.H.v.d.B., O.A.A., P.G., A.M., F.R., T.M.W., G.R.A., D.I.B., D.I.C., E.J.C.d.G., T.M.F., J.N.H., J.J.H., E.I., R.J.F.L., P.K.E.M., N.G.M., G.W.M., K.E.N., N.L.P., T.D.S. and E.K.S. Manuscript writing: M.R.R. and P.M.V., with contributions from all authors on the final version.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Moussavi, S. *et al.* Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *Lancet* **370**, 851–858 (2007).
2. Wild, S., Roglic, G., Green, A., Sicree, R. & King, H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care* **27**, 1047–1053 (2004).
3. Dye, C. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. *J. Am. Med. Assoc.* **282**, 677–686 (1999).
4. Lopez, A.D., Mathers, C.D., Ezzati, M., Jamison, D.T. & Murray, C.J.L. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* **367**, 1747–1757 (2006).
5. Wang, H. *et al.* Age-specific and sex-specific mortality in 187 countries, 1970–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2071–2094 (2012).
6. Jemal, A., Center, M.M., DeSantis, C. & Ward, E.M. Global patterns of cancer incidence and mortality rates and trends. *Cancer Epidemiol. Biomarkers Prev.* **19**, 1893–1907 (2010).
7. Kim, A.S. & Johnston, S.C. Global variation in the relative burden of stroke and ischemic heart disease. *Circulation* **124**, 314–323 (2011).
8. Johnston, S.C., Mendis, S. & Mathers, C.D. Global variation in stroke burden and mortality: estimates from monitoring, surveillance, and modelling. *Lancet Neurol.* **8**, 345–354 (2009).
9. Yang, J., Visscher, P.M. & Wray, N.R. Sporadic cases are the norm for complex disease. *Eur. J. Hum. Genet.* **18**, 1039–1043 (2010).
10. Hill, W.G., Goddard, M.E. & Visscher, P.M. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet.* **4**, e1000008 (2008).
11. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
12. Morris, A.P. *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* **44**, 981–990 (2012).
13. Lee, S.H. *et al.* Estimation and partitioning of polygenic variation captured by common SNPs for Alzheimer's disease, multiple sclerosis and endometriosis. *Hum. Mol. Genet.* **22**, 832–841 (2013).
14. Yang, J. *et al.* Ubiquitous polygenicity of human complex traits: genome-wide analysis of 49 traits in Koreans. *PLoS Genet.* **9**, e1003355 (2013).
15. Robinson, M.R., Wray, N.R. & Visscher, P.M. Explaining additional genetic variation in complex traits. *Trends Genet.* **30**, 124–132 (2014).
16. Abegunde, D.O., Mathers, C.D., Adam, T., Ortegon, M. & Strong, K. The burden and costs of chronic diseases in low-income and middle-income countries. *Lancet* **370**, 1929–1938 (2007).
17. Kim, A.S. & Johnston, S.C. Temporal and geographic trends in the global stroke epidemic. *Stroke* **44**, S123–S125 (2013).
18. Ezzati, M. & Riboli, E. Can noncommunicable diseases be prevented? Lessons from studies of populations and individuals. *Science* **337**, 1482–1487 (2012).
19. Hartl, D.L. & Clark, A.G. *Principles of Population Genetics* (Sinauer Associates, 1997).
20. Leinonen, T., McCairns, R.J.S., O'Hara, R.B. & Merilä, J. Q_{ST}–F_{ST} comparisons: evolutionary and ecological insights from genomic heterogeneity. *Nat. Rev. Genet.* **14**, 179–190 (2013).
21. James, P.T., Rigby, N. & Leach, R. The obesity epidemic, metabolic syndrome and future prevention strategies. *Eur. J. Cardiovasc. Prev. Rehabil.* **11**, 3–8 (2004).
22. Popkin, B.M. Global nutrition dynamics: the world is shifting rapidly toward a diet linked with noncommunicable diseases. *Am. J. Clin. Nutr.* **84**, 289–298 (2006).
23. Wang, Y.C., McPherson, K., Marsh, T., Gortmaker, S.L. & Brown, M. Health and economic burden of the projected obesity trends in the USA and the UK. *Lancet* **378**, 815–825 (2011).
24. Ng, M. *et al.* Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **384**, 766–781 (2014).
25. Finucane, M.M. *et al.* National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9·1 million participants. *Lancet* **377**, 557–567 (2011).
26. Turchin, M.C. *et al.* Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nat. Genet.* **44**, 1015–1019 (2012).
27. Speliotes, E.K. *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **42**, 937–948 (2010).
28. Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832–838 (2010).
29. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).
30. Yang, J. *et al. FTO* genotype is associated with phenotypic variability of body mass index. *Nature* **490**, 267–272 (2012).
31. Amato, R., Miele, G., Monticelli, A. & Cocozza, S. Signs of selective pressure on genetic variants affecting human height. *PLoS ONE* **6**, e27588 (2011).
32. Berg, J.J. & Coop, G. A population genetic signal of polygenic adaptation. *PLoS Genet.* **10**, e1004412 (2014).
33. Wood, A.R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46**, 1173–1186 (2014).
34. Locke, A.E. *et al.* Genetic studies of body mass index yeild new insights for obesity biology. *Nature* **518**, 197–206 (2015).
35. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
36. Dudbridge, F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet.* **9**, e1003348 (2013).
37. Baten, J. & Blum, M. Growing tall but unequal: new findings and new background evidence on anthropometric welfare in 156 countries, 1810–1989. *Econ. Hist. Dev. Reg.* **27**, S66–S85 (2012).
38. Sabeti, P.C. *et al.* Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913–918 (2007).
39. Nielsen, R., Hellmann, I., Hubisz, M., Bustamante, C. & Clark, A.G. Recent and ongoing selection in the human genome. *Nat. Rev. Genet.* **8**, 857–868 (2007).
40. Bustamante, C.D. *et al.* Natural selection on protein-coding genes in the human genome. *Nature* **437**, 1153–1157 (2005).
41. Blekhman, R. *et al.* Natural selection on genes that underlie human disease susceptibility. *Curr. Biol.* **18**, 883–889 (2008).
42. Barreiro, L.B., Laval, G., Quach, H., Patin, E. & Quintana-Murci, L. Natural selection has driven population differentiation in modern humans. *Nat. Genet.* **40**, 340–345 (2008).
43. Akey, J.M. *et al.* Population history and natural selection shape patterns of genetic variation in 132 genes. *PLoS Biol.* **2**, e286 (2004).
44. Barreiro, L.B. & Quintana-Murci, L. From evolutionary genetics to human immunology: how selection shapes host defence genes. *Nat. Rev. Genet.* **11**, 17–30 (2010).
45. Vasseur, E. & Quintana-Murci, L. The impact of natural selection on health and disease: uses of the population genetics approach in humans. *Evol. Appl.* **6**, 596–607 (2013).
46. Chiaroni, J., Underhill, P.A. & Cavalli-Sforza, L.L. Y chromosome diversity, human expansion, drift, and cultural evolution. *Proc. Natl. Acad. Sci. USA* **106**, 20174–20179 (2009).
47. Ovaskainen, O., Karhunen, M., Zheng, C., Arias, J.M.C. & Merilä, J. A new method to uncover signatures of divergent and stabilizing selection in quantitative traits. *Genetics* **189**, 621–632 (2011).
48. Diverse Populations Collaborative Group. Weight-height relationships and body mass index: some observations from the Diverse Populations Collaboration. *Am. J. Phys. Anthropol.* **128**, 220–229 (2005).
49. Lande, R. Genetic variation and phenotypic evolution during allopatric speciation. *Am. Nat.* **116**, 463–479 (1980).
50. Esko, T. *et al.* Genetic characterization of northeastern Italian population isolates in the context of broader European genetic diversity. *Eur. J. Hum. Genet.* **21**, 659–665 (2013).

[1]Queensland Brain Institute, University of Queensland, Brisbane, Queensland, Australia. [2]Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, the Netherlands. [3]Institute for Maternal and Child Health, Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS) 'Burlo Garofolo', Trieste, Italy. [4]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK. [5]Estonian Genome Center, University of Tartu, Tartu, Estonia. [6]Center for Basic and Translational Obesity Research, Boston Children's Hospital, Boston, Massachusetts, USA. [7]Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA. [8]Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA. [9]University of Queensland Diamantina Institute, University of Queensland, Translational Research Institute, Brisbane, Queensland, Australia. [10]Division of Cancer Epidemiology and Genetics, National Cancer Institute, US National Institutes of Health, Bethesda, Maryland, USA. [11]Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden. [12]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [13]Division of Gastroenterology, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA. [14]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA. [15]Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA. [16]Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark. [17]Genetics of Complex Traits, University of Exeter Medical School, University of Exeter, Exeter, UK. [18]Department of Neurology and Neurosurgery, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, the Netherlands. [19]Norwegian Centre for Mental Disorders Research (NORMENT), KG Jebsen Centre for Psychosis Research, Division of Mental Health and Addiction, Oslo University Hospital and Institute of Clinical Medicine, University of Oslo, Oslo, Norway. [20]Institute of Biological Psychiatry, MHC Sct. Hans, Mental Health Devices Copenhagen, Roskilde, Denmark. [21]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [22]Lundbeck Foundation Initiative for Integrative Psychiatric Research, (iPSYCH), Aarhus, Denmark. [23]Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam, the Netherlands. [24]EMGO+ Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands. [25]Department of Biological Psychology, VU University Amsterdam, Amsterdam, the Netherlands. [26]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA. [27]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. [28]Medical Research Council (MRC) Epidemiology Unit, University of Cambridge, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK. [29]Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, USA. [30]Genetics of Obesity and Related Metabolic Traits Program, Icahn School of Medicine at Mount Sinai, New York, New York, USA. [31]Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA. [32]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [33]QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [34]Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [35]Department of Twin Research and Genetic Epidemiology, King's College London, St. Thomas' Hospital, London, UK. [36]Biosciences Research Division, Department of Primary Industries, Melbourne, Victoria, Australia. [37]Department of Food and Agricultural Systems, University of Melbourne, Melbourne, Victoria, Australia. Correspondence should be addressed to M.R.R. (m.robinson11@uq.edu.au) or P.M.V. (peter.visscher@uq.edu.au).

ONLINE METHODS

Genetic differentiation of height and BMI across Europe. The theoretical basis of our analysis framework builds upon those from refs. 32,47,51–53 and is outlined in the **Supplementary Note**. We apply this framework to examine population genetic differentiation in height and BMI across Europe. An overview of the analysis steps is provided in **Supplementary Figure 1**.

Imputation. All of the cohorts used in this study were independently imputed to the 1000 Genomes Project reference panel, using identical quality control procedures on the initial data sets of per-SNP missing data rate <0.02, MAF >0.01, per-individual missing data rate <0.05 and Hardy-Weinberg disequilibrium P <0.0001. Imputation for the majority of the cohorts was performed in two stages. First, the target data were assigned to haplotypes using HAPI-UR[54]. Second, IMPUTE2 (ref. 55) was used to impute the haplotypes to the 1000 Genomes Project reference panel[56] (release 1, version 3). We then selected SNPs that were present across all data sets at an imputation information score of >0.8. The imputation for the Netherlands cohort was identical, except that SHAPEIT2 (ref. 57) was used for haplotyping. We performed these same quality control steps again after combining the data from the different cohorts, including comparisons of allele frequencies across populations.

Selection of SNPs for genomic profiling across Europe. We performed GWAS meta-analyses on data from recent studies[33,34] to select independent loci (r^2 <0.1 and >1 Mb apart using the PLINK clumping procedure[35]) that were associated with the traits in a large sample of individuals (~250,000 for height and ~350,000 for BMI) of European ancestry. We excluded cohorts overlapping with our within-family and prediction samples. For height, we excluded the TWINGENE study, the TwinsUK study, the QIMR sample, the Framingham Heart Study sample and the Netherlands Twin Register. For BMI, cohort-level summary statistics were not available for all samples, and we could only exclude the QIMR sample.

Within-family estimation of SNP effects. We reestimated the SNP effects at these loci in a within-family sibling pair data set (**Supplementary Table 1**) using the QFAM procedure in PLINK described in equation (3.1) in the **Supplementary Note**.

Population genetic analyses conducted using ascertained loci from a standard GWAS can be biased if population stratification is not fully accounted for (**Supplementary Note**). Thus, to confirm our results, we also selected a non-ascertained genome-wide set of unlinked (LD r^2 <0.1 and >1 Mb apart), common (MAF >1%) HapMap 3 loci for height and BMI (~40,000 SNP loci) that passed quality control in both the within-family and prediction samples. We estimated the effects of these SNPs again using our within-family sibling pair data set with the QFAM procedure in PLINK described in equation (3.1) in the **Supplementary Note**. Additionally, we used the clumping procedure in PLINK to select genome-wide unlinked (LD r^2 <0.1 and >1 Mb apart), common (MAF >1%) HapMap 3 loci on the basis of their within-family association with height and BMI (~40,000 SNP loci for both traits).

Genomic profile scoring in a collection of European genomic data. We used the within-family effect sizes to create genetic predictors for individuals across a collection of European genomics data. All data were imputed as described above, and details of the cohort are provided in **Supplementary Table 1**.

From the POPRES study, we selected individuals from France, Portugal, Spain, Italy and Switzerland whose grandparents were born in the same country as the sampled individuals. From the Estonian and Finnish cohorts, we selected 1,000 individuals at random who were included in all analyses. For the Netherlands cohort, we selected 1,000 control individuals from the MinE ALS study. These individuals were healthy controls born in the Netherlands, whose grandparents and parents were also born in the Netherlands. From the Psychiatric Genomic Consortium and the Wellcome Trust Case Control Consortium 2, we used control individuals from Bulgaria, Ireland, Norway, Denmark, Sweden and the UK (**Supplementary Table 1**). The POPRES data set used in this manuscript was obtained from the database of Genotypes and Phenotypes (dbGaP) through accession phs000145.v4.p2.

We used the genetic predictors for height and BMI as response variables in a bivariate Bayesian mixed-effects model, which is outlined in equation (2.2) of the **Supplementary Note**. This model was estimated using the R package MCMCglmm[58], with uninformative inverse Wishart priors, a burn-in period of 7,000 iterations, a sampling interval of 10 iterations and a total of 17,000 iterations, providing 1,000 posterior estimates. This provided estimates of the population genetic (co)variance of height and BMI, the residual (individual-level) (co)variance and the best linear unbiased predictors of the means of each nation along with the 95% credible intervals. For each set of SNPs, the genetic predictor was standardized to a z score to allow comparison across SNP sets and comparison with the null model.

Comparison with a null model. At each stage of the analysis, the estimates obtained were compared to the expectations under random genetic drift, using a quantitative genetic framework for studying population differentiation that is outlined in the **Supplementary Note**.

In brief, the within-family regression coefficients were randomized across SNPs 1,000 times, and 1,000 genetic predictors were created in the European prediction sample. By keeping the effect sizes consistent but attributing these effects across SNPs at random, the genetic predictors generated reflect the action of genetic drift. Second, each set of genetic predictors was standardized to a z score and used as a response variable in the Bayesian mixed-effects model outlined in equation (2.2) of the **Supplementary Note**. This provided 1,000 estimates of the population genetic variance and population means under drift; these values are displayed in the figures as the estimates from the neutral model. Third, the sample covariance matrix of these 1,000 estimates was calculated, which provided an estimate of the expected population-level covariance in phenotype under drift. We then used a Mahalanobis distance statistic to provide a measure of the relative deviation of our predicted population-level means from their multivariate theoretical expectations under drift (**Supplementary Note**). This calculation provided the χ^2 test statistic used to compare our predicted estimates to the expected values under drift. As both the drift profile and trait profile scores were transformed to a z score, this comparison was on the same standard deviation scale.

European phenotypic data. Height and BMI measures for males for each of the 14 European countries were taken from recently published estimates[25,37]. For height, measures were available from 1860 to the present day, and, for BMI, measures were available from 1980 to the present day. Values for both phenotypes were adjusted for trends over time before estimating the population means within a mixed-effects model. The model-generated estimates of these means were then compared to the predicted genetic means as described in equation (5.1) to equation (5.4) in the **Supplementary Note**.

Transforming population means onto the observed scale. For graphical presentation of the results, each genetic predictor $\hat{g}$, created for each individual i for each trait m, was approximately transformed back to the observed scale as outlined in equation (6.1) in the **Supplementary Note**.

We used an independent sample of population data to determine the amount of phenotypic variance explained by the genetic predictors created from different sets of SNPs, as measured by $\mathrm{cov}(z_y, z_g)$, where z_y and z_g are the z scores of the phenotype and genetic predictor, respectively, in equation (6.1) of the **Supplementary Note**. Individuals within the Health and Retirement Study (HRS) (**Supplementary Table 1**) were unrelated, and phenotypic values were adjusted by the first 20 principal components of the SNPs used in the predictor, to account for any population stratification before estimating the within-population covariance. The standard deviation for height and BMI measured in the sample of 17,500 quasi-independent sibling pairs was estimated accounting for sex differences. The HRS data set was obtained from dbGaP through accession phs000428.v1.p1.

Testing the contribution of each SNP to the pattern of population differentiation. As described in equation (4.4) in the **Supplementary Note**, we estimated a χ^2 value for each SNP representing its contribution to the pattern of population genetic differentiation. We tested for the association between a SNP's contribution to differentiation and its MAF and expected contribution to phenotypic variation estimated as $2pq\hat{\beta}^2$, where p is the frequency of the minor allele, q is the frequency of the major allele and β is the within-family effect size.

Positional and functional annotation of SNPs. Our aim was to describe the positional genic annotation and gene ontology (GO) categories for the height- and BMI-associated SNPs that contributed the most to the genome-wide pattern of population genetic variance. We selected the 500 height- and BMI-associated SNPs that contributed the most to differentiation. First, on the basis of the genomic position of these SNPs, we assigned them to genes and simply estimated the number of overlapping genes involved in the genetic differentiation of height and BMI. Second, SNPs were assigned to the following positional genic annotation categories: 3′ UTR variant, 5′ UTR variant, intronic variant, noncoding transcription variant, variant 1–1,000 bp downstream of a gene, variant 1–1,000 bp upstream of a gene, missense transcription variant, synonymous transcription variant and noncoding exonic variant. Finally, SNPs were assigned Ensembl gene identities and then GO terms. The top GO annotations are listed in **Supplementary Table 2**.

We then conducted statistical testing. As a baseline, we used the top 10,000 SNPs for height and BMI. We first compared the number of SNPs among the 500 SNPs contributing the most to differentiation in each genic category, to count data from the top 10,000 SNPs. This provided a list of genic categories potentially enriched for highly differentiated SNPs. We used Fisher's exact tests (hypergeometric) with Bonferroni *P*-value correction. Second, we compared the number of SNPs among the 500 SNPs contributing the most to differentiation in each GO term, to count data from the top 10,000 SNPs. Again, we used Fisher's exact tests with Bonferroni *P*-value correction. From this analysis, we selected a list of the top 20 functional categories potentially enriched for each trait. We assigned *P* values to our comparisons of count data but used these only as a guide to select the top categories, rather than as a definitive test of enrichment. All annotation was conducted using the R library biomaRt from Bioconductor.

Genomic profile scoring worldwide. We repeated our analyses for height and BMI using data from the HGDP as analyzed in ref. 32. We imputed the data following our protocol outlined above.

Simulation study using real genotype data. In addition to the simulations described in the **Supplementary Note**, we used the common, independent HapMap 3 SNPs from the 17,500 sibling pairs in the main empirical analyses (**Supplementary Table 1**) as the basis for a series of simulations. We randomly assigned 5,000 of the independent loci across the genome to be causal variants, with their effects sampled from a normal distribution with a mean of 0 and a variance of 1. The heritability of the trait was simulated to be 90%, and a phenotype was created as $y = \sum x_k b_k + e$, where x_k is the indicator value of locus k (0, 1, 2), b_k is the effect size of the locus and e is the residual variance, with $e = N(0, 1- h^2)$, where h^2 is the heritability. Fifty simulation replicates were conducted.

For each simulation replicate, we randomly selected 16,000 sibling pairs to create an estimation set, leaving 1,500 pairs as a prediction set. We then tested the effects of each SNP on the phenotype in the estimation set in three ways. First, we used a within-family sibling pair analysis implemented using the QFAM procedure in PLINK as described in equation (3.1) in the **Supplementary Note**. Second, we selected one member of a sibling pair at random to create an unrelated set of individuals and then estimated SNP effects using ordinary least-squares regression (the standard GWAS approach) without any control for population stratification. Finally, we repeated the GWAS estimation and controlled for the first 20 principal components estimated from the sibling pair sample.

We then used these three sets of estimates to create three different profile scores in the prediction set. We followed a recent approach to partition variance in a predictor in sibling pairs into genetic, environmental, common genetic and common environmental terms[33]. Variance attributable to common genetic or common environmental terms indicates population stratification bias in the effect size estimates, enabling us to demonstrate that our within-family estimates are unbiased by population stratification.

Additionally, we used the 3 sets of effect size estimates to create 3 different profile scores in the European prediction data of 9,416 individuals used in the main empirical analyses. Our European prediction data were projected onto the first principal component estimated in the within-family sample, and two groups of individuals were then created based on the upper and lower quartiles of the distribution of the projected principal component. The first principal component has been shown to reflect population stratification within genotype samples, and we thus stratified the independent prediction sample by the predominant axis of potential bias in the discovery sample. If there is no population stratification bias in estimates of SNP effects, then there should be no significant differentiation between the groups corresponding to the upper and lower quartiles with respect to the mean for the genetic predictor. We therefore compared the estimates obtained from the three predictors to those generated when a predictor was created using the true simulated effects and to those generated when a predictor was created using randomly allocated effect sizes under our null model. To test for ascertainment biases, we selected the top 100 and top 500 SNPs identified by a GWAS that did not control for population stratification, created 2 predictors containing these SNPs—one from the GWAS effect size estimates and one from the within-family estimates—and then compared the mean differences between the groups corresponding to the upper and lower quartiles. This procedure identifies SNPs from a GWAS and then tests for population differentiation as in previous approaches.

Finally, we repeated the 50 simulations described above, but we created a genotype-environment correlation. We used the same effect sizes and phenotype as in the simulation described above, but we added a z score of the standardized individual-level eigenvalue of the first principal component estimated in the within-family data. Thus, our phenotype was described as $y = \sum x_k b_k + 0.2\mathrm{PC}_1 + e$, where PC_1 is the z score–standardized value for each individual at the first principal component. This creates a correlation between the phenotype and the first principal component of ~0.2 across simulations, representing a phenotypic difference that aligns to the major axis of genetic differentiation. Note that, in this scenario, there is no selection because causal variants are allocated at random across the genome such that, on average, their frequencies across simulations are not expected to differ in a manner that will create a consistent directional difference in profile score along any axis of population stratification. We repeated the estimation and prediction in our independent European prediction sample to test for a directional deviation in mean profile scores between the groups corresponding to the upper and lower quartiles. Under this scenario, ascertainment biases are expected to be large because the SNPs identified by a GWAS that does not control for population stratification should correspond to the ones with the strongest genotype-environment correlation, creating a predictor with mean differences across the leading principal component. However, across the genome, the expectation is that the direction of SNP differentiation should not align with the direction of the effect size and the direction of the phenotypic differentiation, provided that the predictors are created from SNP effect sizes that are unbiased by population stratification.

Code availability. Full computer code for the analysis and the values used to produce the figures are available from the lead author upon request.

51. Weir, B.S. & Hill, W.G. Estimating F-statistics. *Annu. Rev. Genet.* **36**, 721–750 (2002).
52. Weir, B.S. & Cockerham, C.C. Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370 (1984).
53. Cockerham, C.C. & Weir, B.S. Correlations, descent measures: drift with migration and mutation. *Proc. Natl. Acad. Sci. USA* **84**, 8512–8514 (1987).
54. Williams, A.L., Patterson, N., Glessner, J., Hakonarson, H. & Reich, D. Phasing of many thousands of genotyped samples. *Am. J. Hum. Genet.* **91**, 238–251 (2012).
55. Howie, B., Marchini, J. & Stephens, M. Genotype imputation with thousands of genomes. *G3* (*Bethesda*) **1**, 457–470 (2011).
56. 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
57. Delaneau, O., Marchini, J. & Zagury, J.-F. A linear complexity phasing method for thousands of genomes. *Nat. Methods* **9**, 179–181 (2012).
58. Hadfield, J.D. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* **33** (2), 1–22 (2010).

doi:10.1038/ng.3401