

POPULATION GENETICS

Why structure matters

Great care is needed when interpreting claims about the genetic basis of human variation based on data from genome-wide association studies.

NICK BARTON, JOACHIM HERMISSON AND MAGNUS NORDBORG

Related research article Sohail M, Maier RM, Ganna A, Bloemendal A, Martin AR, Turchin MC, Chiang CWK, Hirschhorn J, Daly MJ, Patterson N, Neale B, Mathieson I, Reich D, Sunyaev SR. 2019. Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *eLife* **8**:e39702. DOI: [10.7554/eLife.39702](https://doi.org/10.7554/eLife.39702)

Related research article Berg JJ, Harpak A, Sinnott-Armstrong N, Joergensen AM, Mostafavi H, Field Y, Boyle EA, Zhang X, Racimo F, Pritchard JK, Coop G. 2019. Reduced signal for polygenic adaptation of height in UK Biobank. *eLife* **8**:e39725. DOI: [10.7554/eLife.39725](https://doi.org/10.7554/eLife.39725)

Human height is the classic example of a quantitative trait: its distribution is continuous, presumably because it is influenced by variation at a very large number of genes, most with a small effect (Fisher, 1918). Yet height is also strongly affected by the environment: average height in many countries increased during the last century and the children of immigrants are often taller than relatives in their country of origin – in both cases presumably due to changing diet and other environmental factors (Cavalli-Sforza and Bodmer, 1971; Grasgruber et al., 2016; NCD Risk Factor Collaboration, 2016). This makes it very difficult to determine the cause of geographic patterns for height, such as the ‘latitudinal cline’ seen in Europe (Figure 1).

Are such patterns caused by environmental or genetic differences – or by a complex combination of both? And to the extent that genetic differences are involved, do they reflect selection or simply random history? A number of recent papers have relied on so-called Genome-Wide Association Studies (GWAS) to address these questions, and reported strong evidence for both genetics and selection. Now, in *eLife*, two papers – one by Jeremy Berg, Arbel Harpak, Nasa Sinnott-Armstrong and colleagues (Berg et al., 2019); the other by Mashaal Sohail, Robert Maier and colleagues (Sohail et al., 2019) – independently reject these conclusions. Even more importantly, they identify problems with GWAS that have broader implications for human genetics.

As the name suggests, GWAS scan the genome for variants – typically single nucleotide polymorphisms (SNPs) – that are associated with a particular condition or trait (phenotype). The first GWAS for height found a small number of SNPs that jointly explained only a tiny fraction of the variation. Because this was in contrast with the high heritability seen in twin studies, it was dubbed ‘the missing heritability problem’ (reviewed in Yang et al., 2010). It was suggested that the problem was simply due to a lack of statistical power to detect polymorphisms of small effect. Subsequent studies with larger sample sizes have supported this explanation: more and more loci have been identified although most of the variation remains ‘unmappable’, presumably because sample sizes on the order of a million are still not large enough (Yengo et al., 2018).

One way in which the unmappable component of genetic variation can be included in a statistical measure is via so-called polygenic

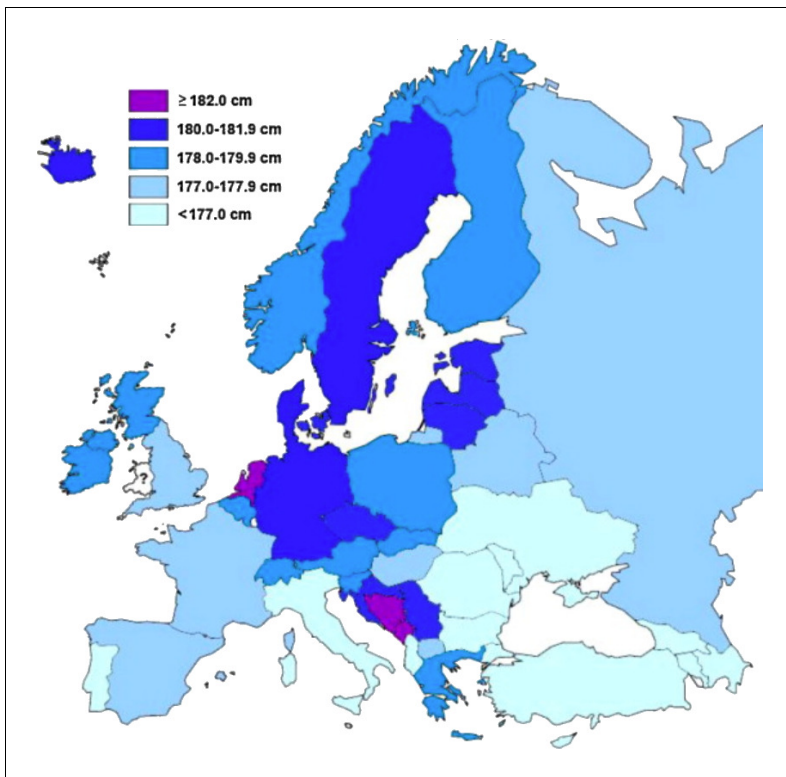


Figure 1. Distribution of average male height in Europe, calculated from studies performed between 1999–2013. In general, southern Europeans tend to be shorter than northern Europeans. Image reproduced from [Grasgruber et al., 2014](#) (CC BY 3.0).

scores. These scores sum the estimated contributions to the trait across many SNPs, including those whose effects, on their own, are not statistically significant. Polygenic scores thus represent a shift from the goal of identifying major genes to predicting phenotype from genotype. Originally designed for plant and animal breeding purposes, polygenic scores can, in principle, also be used to study the genetic basis of differences between individuals and groups.

This, however, requires accurate and unbiased estimation of the effects of all SNPs included in the score, which is difficult in a structured (non-homogeneous) population when environmental differences cannot be controlled. To see why this is a problem, consider the classic example of chopstick-eating skills ([Lander and Schork, 1994](#)). While there surely are genetic variants affecting our ability to handle chopsticks, most of the variation for this trait across the globe is due to environmental differences (cultural background), and a GWAS would mostly identify variants that had nothing to do with chopstick skills, but simply happened to differ in frequency between East Asia and the rest of the world.

Several methods for dealing with this problem have been proposed. When a GWAS is carried out to identify major genes, it is relatively simple to avoid false positives by eliminating associations outside major loci regardless of whether they are due to population structure confounding or an unmappable polygenic background ([Vilhjálmsson and Nordborg, 2013](#)). However, if the goal is to make predictions, or to understand differences among populations (such as the latitudinal cline in height), we need accurate and unbiased estimates for all SNPs. Accomplishing this is extremely challenging, and it is also difficult to know whether one has succeeded.

One possibility is to compare the population estimates with estimates taken from sibling data, which should be relatively unbiased by environmental differences. In one of many examples of this, Robinson et al. used data from the GIANT Consortium ([Wood et al., 2014](#)) together with sibling data to estimate that genetic variation contributes significantly to height variation across Europe ([Robinson et al., 2015](#)). They also argued that selection must have occurred, because the differences were too large to have arisen by chance. Using estimated effect sizes provided by Robinson et al., a more sophisticated analysis by Field et al. found extremely strong evidence for selection for height across Europe ($p=10^{-74}$; [Field et al., 2016](#)). Several other studies reached the same conclusion based on the GIANT data (reviewed in [Berg et al., 2019](#); [Sohail et al., 2019](#)).

Berg et al. (who are based at Columbia University, Stanford University, UC Davis and the University of Copenhagen) and Sohail et al. (who are based at Harvard Medical School, the Broad Institute, and other institutes in the US, Finland and Sweden) now re-examine these conclusions using the recently released data from the UK Biobank ([Sudlow et al., 2015](#)). Estimating effect sizes from these data allows possible biases due to population structure confounding to be investigated, because the UK Biobank data comes from a (supposedly) more homogenous population than the GIANT data.

Using these new estimates, Berg et al. and Sohail et al. independently found that evidence for selection vanishes – along with evidence for a genetic cline in height across Europe. Instead, they show that the previously published results were due to the cumulative effects of slight biases in the effect-size estimates in the GIANT data. Surprisingly, they also found evidence for confounding in the sibling data used as a control

by Robinson et al. and Field et al. This turned out to be due to a technical error in the data distributed by Robinson et al. after they published their paper.

This means we still do not know whether genetics and selection are responsible for the pattern of height differences seen across Europe. That genetics plays a major role in height differences between individuals is not in doubt, and it is also clear that the signal from GWAS is mostly real. The issue is that there is no perfect way to control for complex population structure and environmental heterogeneity. Biases at individual loci may be tiny, but they become highly significant when summed across thousands of loci – as is done in polygenic scores. Standard methods to control for these biases, such as principal component analysis, may work well in simulations but are often insufficient when confronted with real data. Importantly, no natural population is unstructured: indeed, even the data in the UK Biobank seems to contain significant structure (*Haworth et al., 2019*).

Berg et al. and Sohail et al. demonstrate the potential for population structure to create spurious results, especially when using methods that rely on large numbers of small effects, such as polygenic scores. Caution is clearly needed when interpreting and using the results of such studies. For clinical predictions, risks must be weighed against benefits (*Rosenberg et al., 2019*). In some cases, such as recommendations for more frequent medical checkups for patients found at higher ‘genetic’ risk of a condition, it may not matter greatly whether predictors are confounded as long as they work. By contrast, the results of behavioral studies of traits such as IQ and educational attainment (*Plomin and von Stumm, 2018*) must be presented carefully, because while the benefits are far from obvious, the risks of such results being misinterpreted and misused are quite clear. The problem is worsened by the tendency of popular media to ignore caveats and uncertainties of estimates.

Finally, although quantitative genetics has proved highly successful in plant and animal breeding, it should be remembered that this success has been based on large pedigrees, well-controlled environments, and short-term prediction. When these methods have been applied to natural populations, even the most basic predictions fail, in large part due to poorly understood environmental factors (*Charmantier et al., 2014*). Natural populations are never homogeneous, and it is therefore

misleading to imply there is a qualitative difference between ‘within-population’ and ‘between-population’ comparisons – as was recently done in connection with James Watson’s statements about race and IQ (*Harmon, 2019*). With respect to confounding by population structure, the key qualitative difference is between controlling the environment experimentally, and not doing so. Once we leave an experimental setting, we are effectively skating on thin ice, and whether the ice will hold depends on how far out we skate.

Acknowledgments

We thank Jeremy Berg and Peter Visscher for answering our questions, and Molly Przeworski for helpful discussions.

Nick Barton is at IST Austria, Klosterneuburg, Austria
nick.barton@ist-austria.ac.at

 <https://orcid.org/0000-0002-8548-5240>

Joachim Hermisson is at the Department of Mathematics and at the Max F. Perutz Laboratories, University of Vienna, Vienna, Austria
joachim.hermisson@univie.ac.at

 <http://orcid.org/0000-0001-7476-9283>

Magnus Nordborg is at the Gregor Mendel Institute, Austrian Academy of Sciences, Vienna BioCenter, Vienna, Austria

magnus.nordborg@gmi.oeaw.ac.at

 <https://orcid.org/0000-0001-7178-9748>

Competing interests: The authors declare that no competing interests exist.

Published 21 March 2019

References

- Berg JJ**, Harpak A, Sinnott-Armstrong N, Joergensen AM, Mostafavi H, Field Y, Boyle EA, Zhang X, Racimo F, Pritchard JK, Coop G. 2019. Reduced signal for polygenic adaptation of height in UK Biobank. *eLife* **8**: e39725. DOI: <https://doi.org/10.7554/eLife.39725>, PMID: 30895923
- Cavalli-Sforza LL**, Bodmer WF. 1971. *The Genetics of Human Populations*. San Francisco: WH Freeman.
- Charmantier A**, Garant D, Kruuk LEB. 2014. *Quantitative Genetics in the Wild*. Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199674237.001.0001>
- Field Y**, Boyle EA, Telis N, Gao Z, Gaulton KJ, Golan D, Yengo L, Rocheleau G, Froguel P, McCarthy MI, Pritchard JK. 2016. Detection of human adaptation during the past 2000 years. *Science* **354**:760–764. DOI: <https://doi.org/10.1126/science.aag0776>, PMID: 27738015
- Fisher RA**. 1918. XV.—The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh* **52**: 399–433. DOI: <https://doi.org/10.1017/S0080456800012163>

- Grasgruber P**, Cacek J, Kalina T, Sebera M. 2014. The role of nutrition and genetics as key determinants of the positive height trend. *Economics & Human Biology* **15**:81–100. DOI: <https://doi.org/10.1016/j.ehb.2014.07.002>, PMID: 25190282
- Grasgruber P**, Sebera M, Hrazdíra E, Cacek J, Kalina T. 2016. Major correlates of male height: A study of 105 countries. *Economics & Human Biology* **21**:172–195. DOI: <https://doi.org/10.1016/j.ehb.2016.01.005>, PMID: 26948573
- Harmon A**. 2019. James Watson had a chance to salvage his reputation on race. He made things worse. *The New York Times*. <https://www.nytimes.com/2019/01/01/science/watson-dna-genetics-race.html> January 1, 2019.
- Haworth S**, Mitchell R, Corbin L, Wade KH, Dudding T, Budu-Aggrey A, Carslake D, Hemani G, Paternoster L, Smith GD, Davies N, Lawson DJ, J Timpson N. 2019. Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. *Nature Communications* **10**:333. DOI: <https://doi.org/10.1038/s41467-018-08219-1>, PMID: 30659178
- Lander ES**, Schork NJ. 1994. Genetic dissection of complex traits. *Science* **265**:2037–2048. DOI: <https://doi.org/10.1126/science.8091226>, PMID: 8091226
- NCD Risk Factor Collaboration**. 2016. A century of trends in adult human height. *eLife* **5**:e13410. DOI: <https://doi.org/10.7554/eLife.13410>, PMID: 27458798
- Plomin R**, von Stumm S. 2018. The new genetics of intelligence. *Nature Reviews Genetics* **19**:148–159. DOI: <https://doi.org/10.1038/nrg.2017.104>, PMID: 29335645
- Robinson MR**, Hemani G, Medina-Gomez C, Mezzavilla M, Esko T, Shakhbazov K, Powell JE, Vinkhuyzen A, Berndt SI, Gustafsson S, Justice AE, Kahali B, Locke AE, Pers TH, Vedantam S, Wood AR, van Rheenen W, Andreassen OA, Gasparini P, Metspalu A, et al. 2015. Population genetic differentiation of height and body mass index across Europe. *Nature Genetics* **47**:1357–1362. DOI: <https://doi.org/10.1038/ng.3401>, PMID: 26366552
- Rosenberg NA**, Edge MD, Pritchard JK, Feldman MW. 2019. Interpreting polygenic scores, polygenic adaptation, and human phenotypic differences. *Evolution, Medicine, and Public Health* **2019**:26–34. DOI: <https://doi.org/10.1093/emph/eoy036>, PMID: 30838127
- Sohail M**, Maier RM, Ganna A, Bloemendal A, Martin AR, Turchin MC, Chiang CW, Hirschhorn J, Daly MJ, Patterson N, Neale B, Mathieson I, Reich D, Sunyaev SR. 2019. Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *eLife* **8**:e39702. DOI: <https://doi.org/10.7554/eLife.39702>, PMID: 30895926
- Sudlow C**, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T, Collins R. 2015. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLOS Medicine* **12**:e1001779. DOI: <https://doi.org/10.1371/journal.pmed.1001779>, PMID: 25826379
- Vilhjálmsón BJ**, Nordborg M. 2013. The nature of confounding in genome-wide association studies. *Nature Reviews Genetics* **14**:1–2. DOI: <https://doi.org/10.1038/nrg3382>, PMID: 23165185
- Wood AR**, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, Chu AY, Estrada K, Luan J, Kutalik Z, Amin N, Buchkovich ML, Croteau-Chonka DC, Day FR, Duan Y, Fall T, Fehrmann R, Ferreira T, Jackson AU, Karjalainen J, et al. 2014. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics* **46**:1173–1186. DOI: <https://doi.org/10.1038/ng.3097>, PMID: 25282103
- Yang J**, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, Goddard ME, Visscher PM. 2010. Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* **42**:565–569. DOI: <https://doi.org/10.1038/ng.608>, PMID: 20562875
- Yengo L**, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, Frayling TM. 2018. Meta-analysis of genome-wide association studies for height and body mass index in ~700,000 individuals of European ancestry. *bioRxiv*. DOI: <https://doi.org/10.1101/274654>