Signals of polygenic adaptation on height have been overestimated due to uncorrected population structure in genome-wide association studies

Mashaal Sohail[1,2,3+], Robert M. Maier[3,4,5+], Andrea Ganna[3,4,5,6,7], Alexander Bloemendal[3,4,5], Alicia R. Martin[3,4,5], Michael C. Turchin[8,9], Charleston W. K. Chiang[10], Joel N. Hirschhorn[3,11,12], Mark Daly[3,4,5,7], Nick Patterson[3,13], Benjamin Neale[3,4,5*], Iain Mathieson[14*], David Reich[3,13,15*], Shamil Sunyaev[1,2,3*]

[1] Division of Genetics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA
[2] Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts, USA
[3] Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[4] Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[5] Analytical and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts, USA
[6] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[7] Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland
[8] Center for Computational Molecular Biology, Brown University, Providence, Rhode Island, USA
[9] Department of Ecology and Evolutionary Biology, Brown University, Providence, Rhode Island, USA
[10] Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, USA.
[11] Departments of Pediatrics and Genetics, Harvard Medical School, Boston, Massachusetts, USA
[12] Division of Endocrinology and Center for Basic and Translational Obesity Research, Boston Children's Hospital, Boston, Massachusetts, USA
[13] Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA
[14] Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19103, USA
[15] Howard Hughes Medical Institute, Harvard Medical School, Boston, MA 02115, USA

[+] Contributed Equally
* Co-supervised

Genetic predictions of height differ significantly among human populations and these differences are too large to be explained by random genetic drift. This observation has been interpreted as evidence of polygenic adaptation—natural selection acting on many positions in the genome simultaneously. Selected differences across populations were detected using single nucleotide polymorphisms (SNPs) that were genome-wide significantly associated with height, and many studies also found that the signals grew stronger when large numbers of sub-significant SNPs were analyzed. This has led to excitement about the prospect of analyzing large fractions of the genome to detect subtle signals of selection for diverse traits, the introduction of methods to do this, and claims of polygenic adaptation for multiple traits. All of the claims of polygenic adaptation for height to date have been based on SNP ascertainment or effect size measurement in the GIANT Consortium meta-analysis of studies in people of European ancestry. Here we repeat the height analyses in the UK Biobank, a much more homogeneously designed study. While we replicate most previous findings when restricting to genome-wide significant SNPs, when we extend the analyses to large fractions of SNPs in the genome, the differences across groups attenuate and some change ordering. Our results show that polygenic adaptation signals based on large numbers of SNPs below genome-wide significance are extremely sensitive to biases due to uncorrected population structure, a more severe problem in GIANT and possibly other meta-analyses than in the more homogeneous UK Biobank. Therefore, claims of polygenic adaptation for height and other traits—particularly those that rely on SNPs below genome-wide significance—should be viewed with caution.

Most human complex traits are highly polygenic.[1,2] For example, height has been estimated to be modulated by as much as 4% of human allelic variation.[2,3] Polygenic traits are expected to evolve differently from monogenic ones, through slight but coordinated shifts in the frequencies of a large number of alleles, each with mostly small effect. In recent years, multiple methods have sought to detect selection on polygenic traits by evaluating whether shifts in the frequency of trait-associated alleles are correlated with the signed effects of the alleles estimated by genome-wide association studies (GWAS).[4,5,6,7,8,9,10]

Here we focus on a series of recent studies—some involving co-authors of the present manuscript—that have reported evidence of polygenic adaptation at alleles associated with height in Europeans. One set of studies observed that height-increasing alleles are systematically elevated in frequency in northern compared to southern European populations, a result that has subsequently been extended to ancient DNA.[4,5,6,7,8,9,10,11] Another study using a very different methodology (singleton density scores, SDS) found that height-increasing alleles have systematically more recent coalescent times in the United Kingdom (UK) consistent with selection for increased height over the last few thousand years.[12]

All of these studies have been based on SNP associations, in most cases with effect sizes discovered by the GIANT Consortium, which most recently combined 79 individual GWAS through meta-analysis, encompassing a total of 253,288 individuals.[13,14] Here, we show that the selection effects described in these studies are severely attenuated and in some cases no longer significant when using

summary statistics derived from the UK Biobank, an independent and larger single study avoiding meta-analysis that includes 336,474 genetically unrelated individuals who derive their ancestry almost entirely from British Isles (identified as "white British ancestry" by the UK Biobank). The UK Biobank analysis is based on a single cohort drawn from a relatively homogeneous population enabling excellent control of potential population stratification. Our analysis of the UK Biobank data confirms that almost all genome-wide significant loci discovered by the GIANT consortium are real associations (and the genetic correlation between the two height studies is 0.94 [se=0.0078]). However, our analysis yields qualitatively different conclusions with respect to signals of polygenic adaptation.

We began by estimating "polygenic height scores"—sums of allele frequencies at independent SNPs from GIANT weighted by their effect sizes—to study population level differences among ancient and present-day European samples. We used a set of different significance thresholds and strategies to correct for linkage disequilibrium as employed by previous studies, and replicated their signals for significant differences in genetic height across populations.[4,5,6,7,8,9,10,11] (**Figure 1a, Supplementary Figure S1**). We then repeated the analysis using summary statistics from a GWAS for height in the UK Biobank restricting to individuals of British Isles ancestry and correcting for population stratification based on the first ten principal components (UKB).[15] This analysis resulted in a dramatic attenuation of differences in polygenic height scores (**Figure 1a, Supplementary Figures S1-S3**). The differences between ancient European populations also greatly attenuated (**Figure 1a, Supplementary Figure S4**). Strikingly, the ordering of the scores for populations also changed depending on which GWAS was used to estimate genetic height both within Europe (**Figure 1a, Supplementary Figures S1-S4**) and globally (**Supplementary Figure S5**), consistent with reports from a recent simulation study.[16] The height scores were qualitatively similar only when we restricted to independent genome-wide significant SNPs in GIANT and the UK Biobank ($P < 5x10^{-8}$) (**Supplementary Figure S1b**). This replicates the originally reported significant north-south difference in the allele frequency of the height-increasing allele[4] or in genetic height[5] across Europe, as well as the finding of greater genetic height in ancient European steppe pastoralists than in ancient European farmers,[6] although the signals are attenuated even here. This suggests that tests of polygenic adaptation based on genome-wide significant SNPs may be relatively insensitive to confounding (**Supplementary Figure S1b**), and that confounding due to stratification is most likely to arise from sub-significant SNPs (**Figure 1a, Supplementary Figure S1a**).

Next, we looked at polygenic adaptation within the UK using the "singleton density score" (SDS)—an independent measure that uses the local density of alleles that occur only once in the sample as a proxy for coalescent branch lengths.[12,17] SDS can be combined with GWAS effect sizes estimates by aligning the SDS sign to the trait-increasing allele, after which it is referred to as tSDS. A tSDS score larger than zero implies that height-increasing alleles have been increasing in frequency over time due to natural selection. We replicate the finding that tSDS computed in the UK10K is positively rank-correlated with GIANT[12] height P values (Spearman's $\rho = 0.078$, $P = 1.55 \times 10^{-65}$, **Figure 1b**). However, this signal of polygenic adaptation in

the UK attenuated when we used UK Biobank height effect size estimates and P values and became formally non-significant ($\rho = 0.009$, $P = 0.077$, **Figure 1b**).

We propose that the qualitative difference between the polygenic adaptation signals in GIANT and the UK Biobank is the cumulative effect of subtle biases in each of the contributing SNPs in GIANT. This bias can arise due to incomplete control of the population structure in GWAS.[18] For example, if height were differentiated along a north-south axis because of differences in environment, any variant that is differentiated in frequency along the same axis would have an artifactually large effect size estimated in the GWAS. Population structure is substantially less well controlled for in the GIANT study than in the UK Biobank study, both because the GIANT study population is more heterogeneous than that in the UK Biobank, and because the population structure in GIANT may not have been well controlled in some cohorts due to the relatively small size of individual studies (i.e., the ability to detect and correct population structure is dependent on sample size[19,20]). The GIANT study also found that such stratification effects worsen as SNPs below genome-wide significance are used to estimate height scores,[14] consistent with our finding that the differences in genetic height increase when including these SNPs.

To obtain further insight into our observed discrepancy between polygenic adaptation signals in GIANT vs the UK Biobank, we repeated our analyses using estimates of height effect sizes computed using different methods, and then interrogated each of these for signs of population structure. Repeating our analysis with family-based effect size estimates from an independent study (NG2015 sibs),[7] we found evidence for significant differences in polygenic scores between northern and southern Europeans that were qualitatively similar to those obtained using GIANT effect size estimates (**Supplementary Figure S3-S4**). Inclusion of individuals from the UK Biobank who were not of British Isles ancestry without controlling for population structure (UKB all no PCs) in the measurements of effect sizes also produced this pattern (**Supplementary Figure S2-S4**). Thus, UK Biobank estimates that retain population structure show similar patterns to GIANT and previously published family-based estimates (NG 2015 sibs). In contrast, no significant signals of genetic stratification of height or a strong tSDS signal are present across populations from: 1) a genetically homogeneous sample of UK Biobank with entirely British Isles ancestry without controlling for population structure (UKB WB no PCs), or 2) effect size estimates based on UK Biobank families (UKB sibs, UKB sibs WB) (**Supplementary Figures S2-S4, S6-S7**). These analyses further suggest that the lack of signal in the UK biobank analysis is unlikely to be simply due to over-correction for structure in the original UKB estimates.

Indeed, we confirmed that population structure is more correlated with effect size estimates in GIANT than to those in the UK Biobank. **Figure 2a** shows that the effect sizes estimated in GIANT are highly correlated with the SNP loadings of several principal components of population structure (PC loadings). Previously published family-based effect size estimates[7] (NG 2015 sibs) are similarly correlated with the PC loadings showing that they are also affected by population structure despite being computed within families; in other words, these empirical analyses show that these effect size estimates are not free from concerns about population structure either. The within-family design is not problematic on its own as our UK

Biobank family estimates (UKB sibs, UKB sibs WB) computed using the same method do not show any stratification effects (**Supplementary Figure S8-S9**). We do not see a strong correlation with PC loadings in our UK Biobank estimates computed from unrelated individuals (UKB) either (**Figure 2a**). However, the UK Biobank estimates including individuals not of British Isles ancestry and not correcting for population structure (UKB all no PCs) show the same stratification effects as GIANT and NG2015 sibs (**Supplementary Figure S8**). Similarly, we find that alleles that are more common in the Great Britain population (GBR) than in the Tuscan population from Italy (TSI) tend to be preferentially height-increasing according to the GIANT and NG2015 sibs estimates but not according to the UKB estimates (**Figure 2c, Supplementary Figures S9, S19**).

The tSDS analysis should be robust to the type of population structure discussed above.[12] However, there is also a north-south cline in singleton density in Europe, with singleton density being lower in northern than in southern regions.[21] This cline in singleton density coincidentally parallels the phenotypic cline in height and the major axis of genome-wide genetic variation. Therefore, when we perform the tSDS test using GIANT-estimated effect sizes and P values, we find fewer singletons around the inferred height-increasing alleles which tend, due to the uncontrolled population stratification in GIANT, to be at high frequency in northern Europe (**Figures 2c**). This effect thus does not appear when we use UK Biobank summary statistics because of the much lower level of population stratification and more modest variation in height. We confirmed that the tSDS statistic is correlated with principal components across all SNPs (**Figure 2b**), and that alleles that are more common in GBR than in TSI tend to produce higher tSDS scores (**Figure 2d**).

A striking pattern of the GIANT height tSDS signal is its almost linear increase over the whole range of P values (**Figure 1b**). This is not expected if the effect was driven by natural selection (because we expect that not all SNPs will be linked with SNPs that affect height) and is not observed with UK Biobank effect sizes and P values (**Figure 1b)**. We further find that the tSDS signal which is observed across the whole range of P values in some summary statistics can be mimicked by replacing SDS with GBR-TSI allele frequency differences (**Figures 3a, 3c, Supplementary Figures S6-S7, S10-S11**), suggesting that the tSDS signal at non-significant SNPs is driven by uncorrected stratification. However, as with the polygenic score analysis, a small but significant effect is observed when we restrict to genome-wide significant SNPs ($P < 5 \times 10^{-8}$). This effect persists when using UK Biobank family-based estimates for genome-wide significant SNPs (**Figure 3b**), and is not driven by allele frequency differences between GBR and TSI (**Figure 3d**), suggesting a true but attenuated signal of polygenic adaptation in the UK that is driven by a much smaller number of SNPs than previously thought.

Lastly, we asked whether any remaining differences in polygenic height scores among populations are driven by polygenic selection by using the Q_x framework to test against a null model of genetic drift.[5] We re-computed polygenic height scores in the POPRES dataset for this analysis as it has larger sample sizes of northern and southern Europeans than the 1000 Genomes project.[22] We computed height scores using independent SNPs that are 1) genome-wide significant in the UK Biobank ("gw-sig", $P < 5 \times 10^{-8}$) and 2) sub-significantly associated with height

("sub-sig", $P < 0.01$) in different GWAS datasets. For each of these, we tested if population differences were significant due to an overall overdispersion (P_{Qx}), and if they were significant along a north-south cline (P_{lat}) (**Figure 4**). Both gw-sig and sub-sig SNP-based scores computed using GIANT effect sizes showed significant overdispersion of height scores overall and along a latitude cline, consistent with previous results (**Figure 4**). However, the signal attenuated dramatically between sub-sig ($Q_x = 1100$, $P_{Qx} = 1 \times 10^{-220}$) and gw-sig ($Q_x = 48$, $P_{Qx} = 2 \times 10^{-4}$) height scores. In comparison, scores that were computed using the UK Biobank (UKB) effect sizes showed substantially attenuated differences using both sub-sig ($Q_x = 64$, $P_{Qx} = 5 \times 10^{-7}$) and gw-sig ($Q_x = 33$, $P_{Qx} = 0.02$) SNPs, and a smaller difference between the two scores. This suggests that the attenuation of the signal in GIANT is not only driven by a loss of power when using fewer gw-sig SNPs, but also reflects a decrease in stratification effects. The overdispersion signal disappeared entirely when the UK Biobank family based effect sizes were used (**Figure 4**). Moreover, there is likely residual population structure even within the UK Biobank,[23] as randomly ascertained Q_x P values based on the UK Biobank summary statistics are not uniformly distributed as would be expected if population structure was absent (**Supplementary Figure S15**). Therefore, we remain cautious about interpreting any residual signals as "real" signals of polygenic adaptation.

In sum, estimates of population differences in polygenic height scores are strikingly attenuated with the UK Biobank GWAS data relative to previous analyses. We find some evidence for population-level differences in genetic height, but it can only be robustly seen at highly significant SNPs, because any signal at less significant P values is dominated by the effect of residual population structure. Even genome-wide significant SNPs in these analyses may be subtly affected by population structure, leading to continued overestimation of the effect. Thus, it is difficult to arrive at any quantitative conclusion regarding the proportion of the population differences that are due to statistical biases vs. population stratification of genetic height. It is equally challenging to test whether differences in genetic height are due to adaptation in response to environmental differences, migration and admixture (e.g. fraction of Steppe pastoralist ancestry), or relaxation of negative selection. Further, estimates of the number of independent genetic loci contributing to height variation are sensitive to and likely confounded by residual population stratification.

We conclude that while GIANT SNP effect estimates are highly concordant with the UK Biobank individually (**Supplementary Tables S4-S6, Supplementary Figure S18**), they are also influenced by residual population stratification that can mislead inferences about polygenic selection across populations in aggregate. Although these biases are subtle, in the context of tests for polygenic adaptation, which are driven by small systematic shifts in allele frequency, they can create highly significant artificial signals especially when SNPs that are not genome-wide significant are used to estimate genetic height. In no way do our results question the reliability of the genome-wide significant associations discovered in the GIANT cohort or the validity of the statistical methodology used in previously reported polygenic tests for adaptation. However, we urge caution in the interpretation of genome-wide signals of polygenic adaption that are based on large number of sub-

significant SNPs–particularly when using effect sizes derived from meta-analysis of heterogeneous cohorts which may be unable to fully control for population structure.

Acknowledgements
We thank Alkes Price, Jeremy Berg, Graham Coop, Jonathan Pritchard, Matthew Robinson, Jian Yang, Peter Visscher and Hilary Finucane for useful discussions and comments that significantly improved the manuscript. The study was supported by National Institute of Health grants HG009088, MH101244 (M.S., R.M., B.N. and S.S.) and GM127131 (S.S.). D.R. was supported by National Institutes of Health grant GM100233 and HG006399, an Allen Discovery Center of the Paul Allen Foundation, and the Howard Hughes Medical Institute.

References

1. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42,** 565–9 (2010).
2. Boyle, E. A., Li, Y. I. & Pritchard, J. K. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* **169,** 1177–1186 (2017).
3. Zeng, J. *et al.* Signatures of negative selection in the genetic architecture of human complex traits. *Nat. Genet.* **50,** 746–753 (2018).
4. Turchin, M. C. *et al.* Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nat. Genet.* **44,** 1015–1019 (2012).
5. Berg, J. J. & Coop, G. A Population Genetic Signal of Polygenic Adaptation. *PLoS Genet.* **10,** e1004412 (2014).
6. Mathieson, I. et al. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528,** 499–503 (2016).
7. Robinson, M. R. *et al.* Population genetic differentiation of height and body mass index across Europe. *Nat. Genet.* **47,** 1357–1361 (2015).
8. Berg, J. J., Zhang, X. & Coop, G. Polygenic Adaptation has Impacted Multiple Anthropometric Traits. *bioRxiv* (2017). doi:10.1101/167551
9. Racimo, F., Berg, J. J. & Pickrell, J. K. Detecting polygenic adaptation in admixture graphs. *Genetics* **208,** 1565–1584 (2018).
10. Guo, J. *et al.* Global genetic differentiation of complex traits shaped by natural selection in humans. *Nat. Commun.* **9,** 1–9 (2018).
11. Simonti, C., Stein, J., Thompson, P., Fisher, S. E. & Dan, J. Polygenic selection underlies evolution of human brain structure and behavioral traits. *BioRxiv* 1–51 (2017).
12. Field, Y. *et al.* Detection of human adaptation during the past 2,000 years. *Science.* **354,** 760–764 (2016).
13. Allen, H. L. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467,** 832–838 (2010).
14. 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).

15. http://www.nealelab.is/blog/2017/7/19/rapid-gwas-of-thousands-of-phenotypes-for-337000-samples-in-the-uk-biobank.
16. Martin, A. R. *et al.* Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am. J. Hum. Genet.* **100,** 635–649 (2017).
17. Field, Y. *et al.* Data from: Detection of human adaptation during the past 2000 years. Dyrad Digital Repository. doi:https://doi.org/10.5061/dryad.kd58f
18. Novembre, J. & Barton, N. H. Tread Lightly Interpreting Polygenic Tests of Selection. *Genetics* **208,** 1351–1355 (2018).
19. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2,** 2074–2093 (2006).
20. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38,** 904–909 (2006).
21. Sohail, M. *et al.* Negative selection in humans and fruit flies involves synergistic epistasis. *Science.* **356,** 539–542 (2017).
22. Nelson, M. R. *et al.* The Population Reference Sample, POPRES: A Resource for Population, Disease, and Pharmacological Genetics Research. *Am. J. Hum. Genet.* **83,** 347–358 (2008).
23. Haworth, S. *et al.* Common genetic variants and health outcomes appear geographically structured in the UK Biobank sample: Old concerns returning and their implications. *bioRxiv* 294876 (2018). doi:10.1101/294876

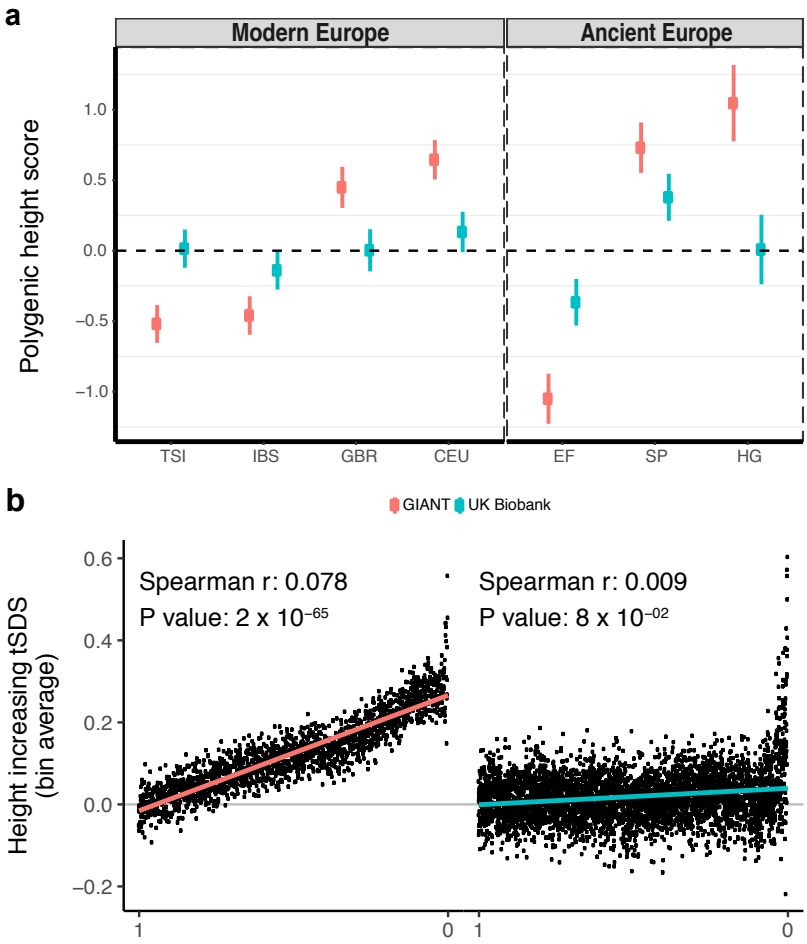

Figure 1. Polygenic height scores and tSDS scores based on GIANT and UK Biobank GWAS.

Polygenic scores in present-day and ancient European populations are shown, centered by the average score across populations and standardized by the square root of the additive variance. Independent SNPs for the polygenic score from both GIANT (*red*) and the UK Biobank (*blue*) were selected by picking the SNP with the lowest P value in each of 1700 independent LD blocks similarly to refs [8,9](see methods). Present-day populations are shown from Northern Europe (CEU, GBR) and Southern Europe (IBS, TSI) from the 1000 genomes project. Ancient populations are shown in three meta-populations (HG = Hunter-Gatherer (n=162 individuals), EF = Early Farmer (n=485 individuals), and SP = Steppe Ancestry (n=465 individuals)). Error bars are drawn at 95% credible intervals. See Supplementary figures S1-S5 for polygenic height scores computed using other linkage disequilibrium pruning procedures, significance thresholds, summary statistics and populations. (b) tSDS for height-increasing allele in GIANT (left) and UK Biobank (right). The tSDS method was applied using pre-computed Singleton Density Scores for 4,451,435 autosomal SNPs obtained from 3,195 individuals from the UK10K project[12,17] for SNPs associated with height in GIANT and the UK biobank. SNPs were ordered by GWAS P value and grouped into bins of 1000 SNPs each. The mean tSDS score within each P value bin is shown on the y-axis. The Spearman correlation coefficient between the tSDS scores and GWAS P values, as well as the correlation standard errors and P values, were computed on the un-binned data. The gray line indicates the null-expectation, and the colored lines are the linear regression fit. The correlation is significant for GIANT (Spearman r = 0.078, P = 1.55×10^{-65}) but not for UK Biobank (Spearman r = -0.009, P = 0.077).

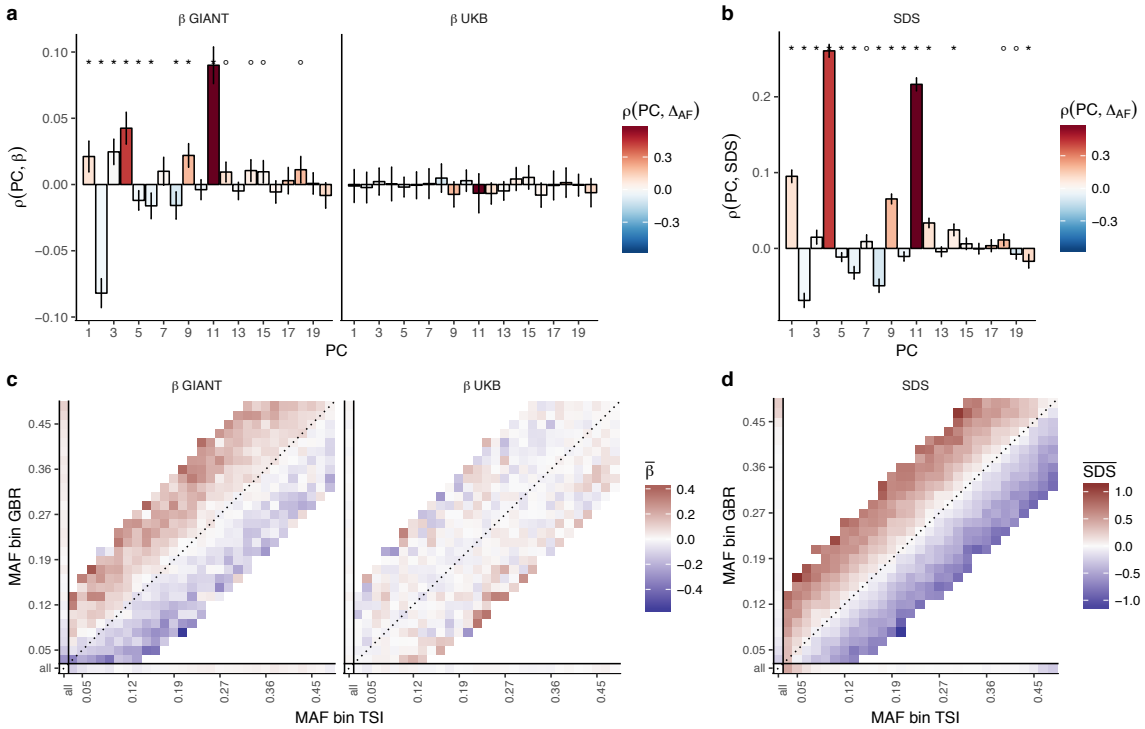

Figure 2. Evidence of stratification in height summary statistics.
Top row: Pearson Correlation coefficients of (a) PC loadings and height beta coefficients from GIANT and UKB, and (b) PC loadings and SDS (pre-computed in the UK10K) across all SNPs. PCs were computed in all 1000 genomes phase 1 samples. Colors indicate the correlation of each PC loading with the allele frequency difference between GBR and TSI, a proxy for the European North-South genetic differentiation. PC 4 and 11 are most highly correlated with the GBR - TSI allele frequency difference. Confidence intervals and P values are based on Jackknife standard errors (1000 blocks). Open circles indicate correlations significant at alpha = 0.05, stars indicate correlations significant after Bonferroni correction in 20 PCs ($P < 0.0025$).
Bottom row: Heat map after binning all SNPs by GBR and TSI minor allele frequency of (c) mean beta coefficients from GIANT and UKB, and (d) SDS scores for all SNPs. Only bins with at least 300 SNPs are shown. While the stratification effect in SDS is not unexpected, it can lead to false conclusions when applied to summary statistics that exhibit similar stratification effects. UKB height betas exhibit stratification effects that are weaker, and in the opposite direction of the stratification effects in GIANT (see Figure S14 for a possible explanation).

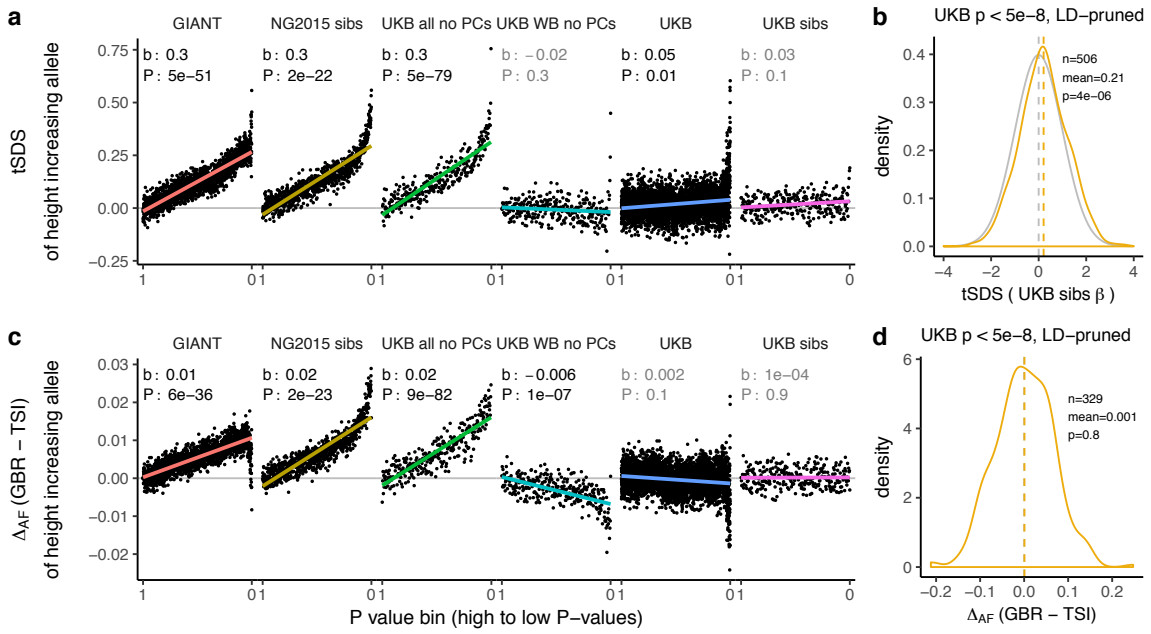

Figure 3. Height tSDS results for different summary statistics.
(a) Mean tSDS of the height increasing allele in each P value bin for six different summary statistics. The first two panels are computed analogously to Figure 4A and Figure S22 in Field et al. In contrast to those Figures and to Figure 1b, the displayed betas and P values correspond to the slope and P value of the linear regression across all un-binned SNPs (rather than the Spearman correlation coefficient and Jackknife P values). The y-axis has been truncated at 0.75, and does not show the top bin for UKB all no PCs, which has a mean tSDS of 1.5. See Figure S6 for other GWAS summary statistics. (b) tSDS distribution of the height increasing allele in 506 LD-independent SNPs which are genome-wide significant in a UKB height GWAS, where the beta coefficient is taken from a within sibling analysis in the UKB. The gray curve represents the standard normal null distribution, and we observe a significant shift providing confirmation of a real SDS signal of polygenic adaptation for height. (c) Allele frequency difference between GBR and TSI of the height increasing allele in each P value bin for six different summary statistics. Betas and P values correspond to the slope and P value of the linear regression across all un-binned SNPs. The lowest P value bin in UKB all no PCs with a y-axis value of 0.06 has been omitted. See Figure S10 for other GWAS summary statistics. (d) Allele frequency difference between GBR and TSI of the height increasing allele in 329 LD-independent SNPs which are genome-wide significant in a UKB height GWAS and were intersecting with our set of 1000 genomes SNPs. There is no significant difference in frequency in these two populations, suggesting that tSDS shift at the gw-significant SNPs is not driven by population stratification.
The patterns shown here suggest that the positive tSDS values across the whole range of P values is a consequence of residual stratification. At the same time, the increase in tSDS at genome-wide significant, LD-independent SNPs in (b) cannot be explained by GBR - TSI allele frequency differences as shown in (d). Binning SNPs by P value without LD-pruning can lead to unpredictable patterns at the low P value end, as the SNPs at the low P value end are less independent of each other than higher P value SNPs (Figure S12). Figures S7 and S11 therefore show the same data for a set of LD-pruned SNPs.

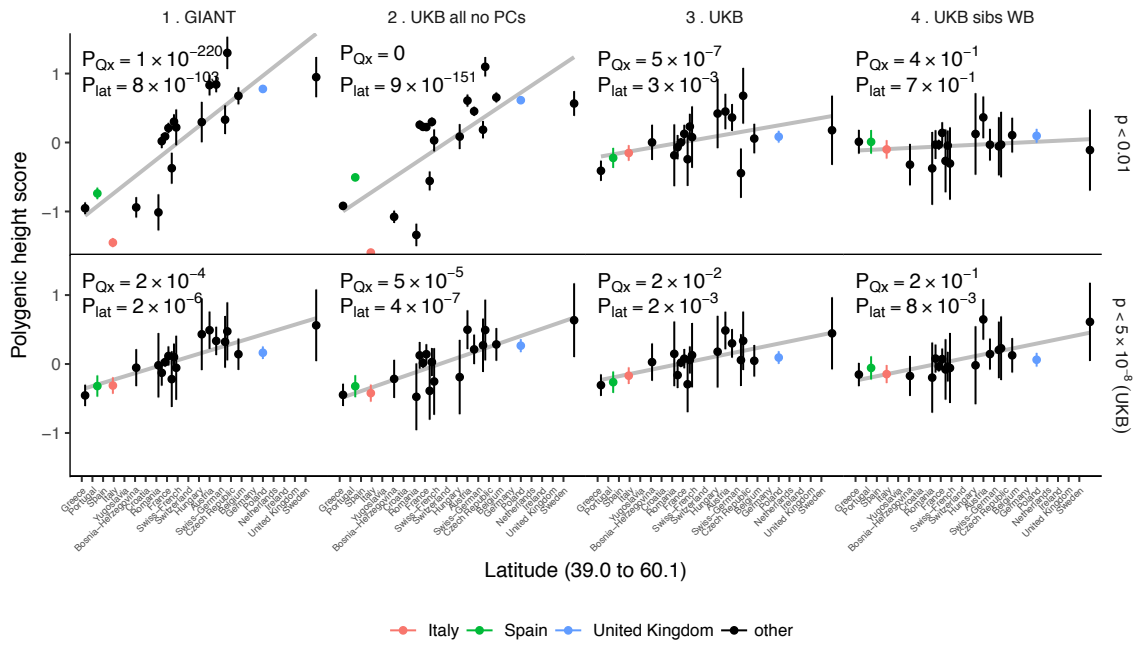

Figure 4. Polygenic height scores in POPRES populations show a residual albeit attenuated signal of polygenic adaptation for height.
Standardized polygenic height scores from six summary statistics for 19 POPRES populations with at least 10 samples per population, ordered by latitude (see Table S3). The grey line is the linear regression fit to the mean polygenic scores per population. Error bars represent 95% confidence intervals and are calculated in the same way as in Figure 1. SNPs which were overlapping between each set of the summary statistics and the POPRES SNPs were clumped using PLINK 1.9 with parameters r^2 < 0.1, 1 Mb distance, P < 1.
(Top) A number of independent SNPs was chosen for each summary statistic to match the number of SNPs which remained when clumping UKB at P < 0.01.
(Bottom) A set of independent SNPs with $P < 5 \times 10^{-8}$ in the UK Biobank was selected and used to compute polygenic scores along with effect size estimates from each of the different summary statistics.
SNP numbers for each panel ($P < 0.01$ / $P < 5 \times 10^{-8}$): GIANT: 9463 / 869; UKB all no PCs: 9341 / 359; UKB: 9341 / 1034; UKB sibs WB: 9391 / 380. The numbers on each plot show the Q_x P value and the latitude covariance P value respectively for each summary statistic. See Supplementary Figure S13, S16 and S17 for other clumping strategies and GWAS summary statistics.