Recovery of trait heritability from whole genome sequence data

Pierrick Wainschtein[1], Deepti P. Jain[2], Loic Yengo[1], Zhili Zheng[1], TOPMed Anthropometry Working Group, Trans-Omics for Precision Medicine Consortium, L. Adrienne Cupples[3], Aladdin H. Shadyab[4], Barbara McKnight[2], Benjamin M. Shoemaker[5], Braxton D. Mitchell[6,33], Bruce M. Psaty[7,8], Charles Kooperberg[9], Dan Roden[10], Dawood Darbar[11], Donna K. Arnett[12], Elizabeth A. Regan[13], Eric Boerwinkle[14], Jerome I. Rotter[15], Matthew A. Allison[16], Merry-Lynn N. McDonald[17], Mina K Chung[18], Nicholas L. Smith[7,8,34], Patrick T. Ellinor[19,20], Ramachandran S. Vasan[3], Rasika A. Mathias[21], Stephen S. Rich[22], Susan R. Heckbert[7], Susan Redline[23], Xiuqing Guo[15], Y.-D Ida Chen[15], Ching-Ti Liu[24], Mariza de Andrade[35], Lisa R. Yanek[37], Christine M. Albert[19,25,26], Ryan D. Hernandez[27,36], Stephen T. McGarvey[28], Kari E. North[29], Leslie A. Lange[30], Bruce S. Weir[2], Cathy C. Laurie[2], Jian Yang[1,31,32,*], Peter M. Visscher[1,32,*]

[1]Institute for Molecular Bioscience, The University of Queensland, Brisbane 4072, Australia; [2]Department of Biostatistics, University of Washington, Seattle, WA, USA; [3]Boston University and the Framingham Heart Study; [4]Department of Family Medicine and Public Health, University of California San Diego School of Medicine, La Jolla, CA; [5]Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA; [6]Department of Medicine, University of Maryland School of Medicine, Baltimore, USA; [7]Cardiovascular Health Research Unit and Department of Epidemiology, University of Washington, Seattle, WA, USA; [8]Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA; [9]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; [10]Departments of Medicine, Pharmacology and Bioinformatics, Vanderbilt University Medical Center, Nashville, TN, USA; [11]Department of Medicine. University of Illinois-Chicago, Chicago, IL; USA; [12]Dean's Office, College of Public Health, University of Kentucky, Lexington, KY, USA; [13]Department of Medicine, National Jewish Health, Denver, CO, USA; [14]University of Texas, Health Science Center, Houston, TX, USA; [15]The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA, USA; [16]Department of Family Medicine and Public Health, University of California San Diego, La Jolla, CA, USA; [17]Division of Pulmonary, Allergy and Critical Care Medicine, University of Alabama at Birmingham, Birmingham, AL, USA; [18]Department of Molecular Cardiology, Cleveland Clinic, Cleveland, OH, USA; [19]Harvard Medical School, Boston, MA, USA; [20]Cardiac Arrhythmia Service, Massachusetts General Hospital, Boston, MA, USA; [21]GeneSTAR Research Program, Divisions of Allergy and Clinical Immunology and General Internal Medicine, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA; [22]Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA; [23]Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, Boston, MA, USA; Division of Sleep Medicine, Harvard Medical School, Boston, MA, USA; Division of Pulmonary, Critical Care, and Sleep Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA; [24]Department of Biostatistics, Boston University School of Public Health Boston, MA, USA; [25]Division of Cardiovascular, Brigham and Women's Hospital, Boston, MA, USA; [26]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA, USA; [27]Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, CA, USA; [28]International Health Institute, Department of Epidemiology, Brown University School of Public Health, Providence, USA; [29]Department of Epidemiology and Carolina Center of

Genome Sciences, University of North Carolina , Chapel Hill , NC , USA; [30]Department of Medicine, University of Colorado Denver, Anschutz Medical Campus, Aurora, CO, USA; [31]Institute for Advanced Research, Wenzhou Medical University, Wenzhou, Zhejiang, China; [32]Queensland Brain Institute, The University of Queensland, Brisbane, Queensland, Australia; [33]Geriatrics Research and Education Clinical Center, Baltimore Veterans Administration Medical Center, Baltimore, MD, USA; [34]Seattle Epidemiologic Research and Information Center, Department of Veterans Affairs Office of Research and Development, Seattle, WA, USA; [35]Department of Health Sciences Research, Mayo Clinic, Rochester, MN, 55905, USA; [36]Department of Human Genetics, McGill University, Montreal, QC, Canada; [37]Division of General Internal Medicine, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA; * Equal contribution.

Abstract

Heritability, the proportion of phenotypic variance explained by genetic factors, can be estimated from pedigree data [1], but such estimates are uninformative with respect to the underlying genetic architecture. Analyses of data from genome-wide association studies (GWAS) on unrelated individuals have shown that for human traits and disease, approximately one-third to two-thirds of heritability is captured by common SNPs [2-5]. It is not known whether the remaining heritability is due to the imperfect tagging of causal variants by common SNPs, in particular if the causal variants are rare, or other reasons such as over-estimation of heritability from pedigree data. Here we show that pedigree heritability for height and body mass index (BMI) appears to be fully recovered from whole-genome sequence (WGS) data on 21,620 unrelated individuals of European ancestry. We assigned 47.1 million genetic variants to groups based upon their minor allele frequencies (MAF) and linkage disequilibrium (LD) with variants nearby, and estimated and partitioned variation accordingly. The estimated heritability was 0.79 (SE 0.09) for height and 0.40 (SE 0.09) for BMI, consistent with pedigree estimates. Low-MAF variants in low LD with neighbouring variants were enriched for heritability, to a greater extent for protein altering variants, consistent with negative selection thereon. Cumulatively variants in the MAF range of 0.0001 to 0.1 explained 0.54 (SE 0.05) and 0.51 (SE 0.11) of heritability for height and BMI, respectively. Our results imply that the still missing heritability of complex traits and disease is accounted for by rare variants, in particular those in regions of low LD.

Introduction

Natural selection shapes the joint distribution of effect size and allele frequency of genetic variants for complex traits in populations, including that of common disease in humans, and determines the amount of additive genetic variation in segregating outbred populations [6]. Traditionally, additive genetic variation, and its ratio to total phenotypic variation (narrow-sense heritability) is estimated using the resemblance between relatives, by equating the expected proportion of genotypes shared identical-by-descent with the observed correlation between relatives [1,6]. Such methods are powerful but blind with respect to genetic architecture. In the last decade, experimental designs that use observed genotypes at many loci in the genome have facilitated the mapping of genetic variants associated with complex traits. In particular, genome-wide association studies (GWAS) in humans have discovered thousands of variants associated with complex traits and diseases [7]. GWAS to date have mainly relied on arrays of common SNPs that are in LD with underlying causal variants. Despite their success in mapping trait-associated variants and detecting evidence for negative selection [8,9], the proportion of phenotypic variance captured by all common SNPs, the SNP-heritability (h^2_{SNP}), is significantly less than the estimates of pedigree heritability (h^2_{ped}) [2,3]. Using SNP genotypes imputed to a fully sequenced reference panel recovers additional additive variance [5,10] but there is still a gap between SNP and pedigree heritability estimates. The most plausible hypotheses for this discrepancy are that causal variants are not well tagged (or imputed) by common SNPs because they are rare and/or that pedigree heritability is over-estimated due to confounding with common environmental effects or non-additive genetic variation [3,11,12].

Understanding the source of still missing heritability and achieving a better quantification of the genetic architecture of complex traits is important for experimental designs to map additional trait loci, for precision medicine and to understand the association between specific traits and fitness. Here we address the hypothesis that the still missing heritability is due to rare variants not sufficiently tagged by common SNPs, by estimating additive genetic variance for height and body mass index (BMI) from whole genome sequence (WGS) data on a large sample of 21,620 unrelated individuals from the Trans-Omics for Precision Medicine (TOPMed) program.

Results

Narrow-sense heritability estimates of height and BMI using WGS data

We used a dataset of 38,466 genomes (Supplementary Table 1) of which we selected a sample of 21,620 genomes with European ancestry (Online Methods; Supplementary Figure 1). After stringent quality control (QC), we retained 47.1M variants, including SNPs and insertion-deletions (indels). With a MAF threshold of 0.0001 during the QC process, each variant was observed at least 3 times in our dataset. The available phenotypes, height and BMI, were adjusted for age and standardized to follow a N(0,1) distribution in each gender group (Online methods, Supplementary Figure 2). We also analysed BMI with a rank based inverse normal transformation (BMI_{RINT}) after adjustment for age and sex.

To verify that we could replicate prior estimates of h^2_{SNP} based on common SNPs, we selected ~1.33 M HapMap 3 (HM3) SNPs from the sequence variants and estimated h^2_{SNP} for height and BMI using the GREML approach implemented in GCTA [13]. We estimated a SNP-based heritability (h^2_{SNP}) of 0.49 (SE 0.02) for height and 0.27 (SE 0.02) for BMI (Supplementary Figure 3), consistent with previous estimates [2,14]. We then repeated the estimation of h^2_{SNP} by mimicking a SNP-array plus imputation strategy and by stratifying

imputed SNPs according to MAF and LD using the GREML-LDMS approach implemented in GCTA [10]. For the LD annotation, we preferred a SNP-specific LD metric rather than a segment-based metric used previously [10] because the former was shown to be more powerful [15]. We tested the two LD metric in our data and reached the same conclusion (Online Methods, Supplementary Figure 4). In the analysis to mimic the strategy of SNP-array genotyping followed by imputation, we extracted the genotypes of variants represented on three arrays (i.e., Illumina HumanCore24, GSA 24 and Affymetrix Axiom) and then imputed the genotype data to the Haplotype Reference Consortium (HRC) reference panels (Online Methods, Supplementary Table 2 and Supplementary Table 3) [16]. Our estimates of h^2_{SNP} increased to 0.58-0.67 (SE 0.07-0.08) for height, and 0.24-0.27 (SE 0.07-0.09) for BMI, depending on the array (Figure *1*). Minor differences in estimates between arrays can be explained by the imputation process generating different number of imputed SNPs in the lower MAF bins between arrays (Supplementary Figure 5). These results from GREML-LDMS analysis on imputed SNPs are consistent with previous reports [10]. By selecting SNPs in common between TOPMed and the imputed SNPs for each array and stratifying by MAF and LD, we evaluated the influence of imputation errors on the variance estimates. Although in most cases h^2_{SNP} slightly increased using WGS over imputed data (Supplementary Figure 6), estimates remained very close between the two data sets. These results indicate that imputation recovers almost as much of the variance as WGS data given the same set of SNPs. Note that the coverage of this set of SNPs is lower than that of the TOPMed WGS data. The loss of information due to imputation is either negligible or, alternatively, imputation and sequencing errors have similar effect on variance component estimation.

We then used all sequence variants to estimate and partition additive genetic variance according to MAF and LD (Supplementary Table 4, Supplementary Figure 7), using the GREML-LDMS partitioning method [10,15]. This resulted in a large increase in heritability estimates. When correcting for the first 20 principal components (PCs), we estimated WGS-based heritability (h^2_{WGS}) to be 0.79 (SE 0.09) for height and 0.40 (SE 0.09) for BMI (Figure *2*). These estimates are close to the pedigree estimates of 0.7-0.8 and 0.4-0.6 for height and BMI, respectively [3,17] and suggest that WGS data fully recovers the total additive genetic variance. The estimates for the rank-transformed BMI data were similar but consistently smaller when compared to those from the untransformed data (Supplementary Figure 8).

The additional variance explained by WGS variants over and above that by the variants from the array plus imputation approach is predominantly from rare variants, in particular for rare variants in low LD with nearby SNPs. For the variants with MAF < 0.1, 0.38 of the phenotypic variance for height was accounted for by variants in the low-LD group but only 0.05 of the variance by variants in the high-LD group (fitting 20 PCs). For BMI, the rare variants with 0.0001 < MAF < 0.001 contributed to 0.15 (SE 0.09) of the phenotypic variance (fitting 20 PCs). This large contribution of rare variants with low LD metric could only be detected using WGS data as these variants are not present on SNP arrays and their imputation is not accurate due to differences in LD structure with imputation panels and the variant coverage of the imputation reference [18]. The proportion of variance explained by rare variants corresponds to most of the difference in estimates of heritability from previous SNP-based studies (0.56 (SE 0.02) and 0.27 (SE 0.02) for height and BMI respectively [10]) and pedigree estimates (0.8 and 0.4-0.6 for height and BMI respectively [3,17]), and confirms evidence for negative selection [8,9] (Supplementary Figure 9).

We performed a number of additional analyses to test the robustness of the estimates. First, we corrected for up to 280 PCs in the analysis. This decreased the estimates to 0.75 (SE 0.09)

for height and 0.39 (SE 0.10) for BMI in the most extreme adjustment (280 PCs), which may have removed real additive genetic variance for the trait (Figure *2*).

Second, we explored if there was any bias due to a specific LD or MAF structure in the TOPMed dataset by using a different sequenced dataset for SNP stratification. We used the UK10K dataset [19] and re-analysed the TOPMed data using the MAF and LD stratification from the UK10K data. After running a QC on UK10K dataset (Online methods), we selected the variants in common with both TOPMed and UK10K datasets. There were 13.9M variants seen in both datasets, with most of the rare variants present only in TOPMed because of the smaller sample size (n = 3781) of the UK10K set. We once again defined 7 MAF bins with 2 LD bins each from the two datasets, using only the 13.9M variants in common (Online methods). We estimated and partitioned additive genetic variance based on these 13.9M variants, using either the MAF and LD annotation from UK10K or TOPMed, fitting 20 PCs from HM3 SNPs. The estimates were highly consistent between the two analyses (Figure *3*). For height, the estimates were 0.62 (SE 0.06) when using TOPMed annotation and 0.60 (SE 0.06) using SNP stratification from UK10K, and the corresponding estimates for BMI were 0.32 (SE 0.06) and 0.30 (SE 0.06), respectively. The higher estimates from TOPMed could reflect a better accuracy of genotypes due to higher sequencing depth (30x) compared to the UK10K data set (7x). These estimates were lower than when using the full set of 47.1M variants, which is expected given the large proportion of h^2_{WGS} explained by rare variants, most of which were missing in this comparison using 13.9M variants. The similarity between the estimates from the two references for MAF and LD stratification suggests that our inference from TOPMed annotation is not biased by using MAF and LD stratification from the same data.

Third, we compared GREML-LDMS estimates by selecting a subset 20.9M of high-quality variants identified with a classifier trained using a support vector machine algorithm (SVM) and additional hard filters (Online Methods) and by randomly selecting a similar number of variants with similar MAF and LD properties from the whole set of 47.1M variants (Supplementary Figure 10). Differences between GREML-LDMS estimates calculated using these two sets of variants (the high quality SVM variants and the random ones) were small, at 0.032 for height (0.71 (SE 0.09) from high quality variants and 0.75 (SE 0.08) from random variants) and 0.002 for BMI (0.44 (SE 0.08-0.09) in both cases for BMI). With similar estimates between high quality variants and the rest of the data set, we can rule out any potential upward bias of estimates due to sequencing errors or batch effects in the sequenced data.

Finally, we investigated the influence of calculating the GRM estimator A_{jk}using the ratio of total SNP covariance and total SNP heterozygosity over loci ("ratio of averages") [20] instead of the average ratio over loci ("average of ratio") [13] on the GREML-LDMS estimates (Online Methods). Note that the default GCTA method (average of ratios) assumes an inverse relationship between MAF and variant effect size whereas the ratio of averages assumes no relationship between MAF and variants effect sizes. While the estimates for height remained similar between the two GRM calculation methods fitting 14 GRMs (0.79 (SE 0.09) using the average of ratios and 0.75 (SE 0.08) using the ratio of averages), they diverged notably for BMI (0.40 (SE 0.09) using the average of ratios and 0.32 (SE 0.08) using the ratio of averages). Most of the difference was due to the low LD $0.0001 < MAF < 0.001$ bin not contributing to phenotypic variance as much as previously estimated. To understand this difference, we further divided the aforementioned bin (Online Methods) to more flexibly model the relationship between variant effect size and MAF, and the estimates converged as

we increased the number of bins (BMI estimates of 0.42 (SE 0.02) using the average of ratios and 0.40 (SE 0.09) using the ratios of averages by fitting 15 bins, and of 0.43 (SE 0.02) using the average of ratios and 0.43 (SE 0.09) using the ratios of averages by fitting 16 bins) (Supplementary Figure 11). The convergence to similar estimates from the two ways of calculating the GRM implies robustness of the analysis.

Functional annotation

Having estimated and partitioned additive genetic variance from WGS data and verified that inference from SNP-array variants and arrays plus imputation are consistent with previous reports, we then explored whether the variance could further be partitioned by functional annotations. To investigate the specific contribution of low-LD variants with MAF < 0.1 to heritability in greater detail, we partitioned the low-MAF and low-LD variants bins further according to the putative effect of a variant on protein coding using SnpEff [21]. The protein-affecting group comprises loss of function and non-synonymous variants whereas the remaining variant set comprises synonymous, regulatory or non-coding variants (Online methods, Supplementary Table 5). The proportion of protein-affecting variants was different across LD and MAF groupings, with an increase in low frequency bins (Supplementary Figure 7). When running a GREML-LDMS analysis with these 20 bins, the total estimates remained the same for height, 0.78 (SE 0.10), and increased for BMI, 0.53 (SE 0.11) (Figure *4*). Interestingly, the average variance explained per variant was larger for bins with protein affecting variants (low-LD) compared to bins with non-protein-altering variants (low-LD) or high-LD variants (Figure *4*).

Discussion

We have used the largest sample to date with both WGS data and phenotypes to estimate the heritability for height and BMI captured by the genome in a sample of 21,620 unrelated individuals from the TOPMed consortium. We recover the heritability estimated from pedigree data for both traits and show that the additional variance detected over and above SNP arrays or imputation is due to rare variants, in particular rare protein altering variants in low LD with other genomic variants.

To assess the robustness of these results, we conducted several follow-up analyses. We estimated the variance explained by correcting with a large number of principal components (up to 280) which would rule out any bias due to population stratification. We used a LD and MAF reference from another European dataset with whole genome sequences, and furthermore tested a subset of high quality variants. All three of these analyses confirmed the validity of the high estimates from the TOPMed WGS data. We also estimated heritability for height and BMI using another GRM estimator and confirmed the robustness our statistical framework. Finally, using a subset of the dataset and imputing it again using HRC reference panel, we showed that heritability estimates were higher by considering individual-SNP-based LD score over the segment-based score, validating the role of LD partitioning in estimating heritability. This method also allowed us to confirm that most of the heritability due to very rare variants ($0.0001 < MAF < 0.001$) was missing when using imputed data but was revealed by using WGS data. We evaluated the loss of information on variance component estimates due to imputation compared to a similar variant coverage of WGS data and found negligible differences in the estimates of genetic variance. We investigated further the enrichment in heritability for different types of variant (high or low impact on the protein) and showed that for low-LD variants with $MAF < 0.1$, non-synonymous and protein truncating variants on average contributed much more to the heritability estimates than synonymous or non-coding variants, as it might be expected biologically.

Our estimates of heritability from WGS data have standard error of about 8%. Since standard errors are approximately inversely proportional to sample size [22], doubling the sample size to 42,000 would narrow errors to ~4% and would allow further and more precise partitioning of genetic variation. Until now the question of the contribution from rare variants to the missing heritability could only be investigated through imputing genotypes from WGS reference panels that was subject to imperfect tagging. Our results quantify this contribution and allows for recovery of most of the remaining missing heritability. It would be interesting to further partition the genome (by variant type, predicted variant deleteriousness [23] and more LD/MAF groupings), but standard errors of the estimates would be too large given our current sample size. Similarly, with a larger sample size, contribution to the heritability from assortative mating could be quantified [24]. The contribution of rare variants to narrow-sense heritability, larger than expected under a neutral model, also reinforces previous observations that height and BMI have been under negative selection, although population expansion could also lead to an increase in heritability from rare variants [25]. Once again, a larger sample size would allow us to draw stronger conclusions on the selective pressure on the genetic variants associated with the two traits.

These results have important implications for the still missing heritability of many traits and diseases [3]. Indeed, the ratio of SNP to pedigree heritability for diseases is lower than for height and BMI, leading to potentially more discovery from rare variants contributing to diseases using WGS data. These results also are important for polygenic risk scores as using WGS data could lead to predictors with larger prediction accuracy for many polygenic

diseases. With the cost for sequencing still much higher than array genotyping, large scale WGS data acquisition is currently limited to national initiatives such as TOPMed and other programs but is bound to expand in the next decade. Large cohorts of genotyping array data will still prove useful for gene discovery or predictions of common diseases and should complement WGS data for a broader understanding of genetic architecture. In the future, WGS programs for specific diseases on large cohorts could lead to a large increase of low MAF variants identified. Sample sizes required to detect such variants from genome-wide association studies using sequence data are of the same order of magnitude as current well-powered GWAS on common SNPs, i.e. hundreds of thousands of samples.

Main figures

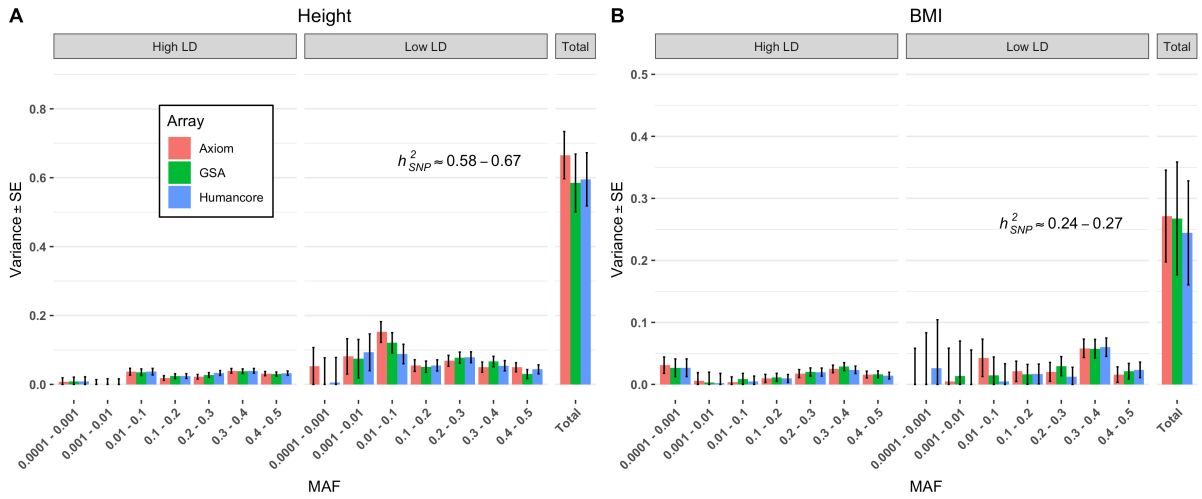

Figure 1: GREML-LDMS estimates with 14 bins (2 LD bins for each of the 7 MAF bins) and 20 PCs correction (calculated from HM3 SNPs) after imputing SNPs from Illumina HumanCore24, GSA 24 and Affymetrix Axiom arrays using Haplotype Reference Consortium reference panels. (A) Estimates of h^2_{SNP} for height are between 0.58-0.67 (SE 0.07-0.08). (B) Estimates for BMI are between 0.24-0.27 (SE 0.07-0.09). The large SEs of the estimates for variants with MAF between 0.0001 to 0.001 can be explained by the large number of imputed variants in this MAF bin because the sampling variance of a SNP-based heritability estimate is proportional to the effective number of independent variants [22]. Between ~15.5M and ~20.7M variants in total are included in the analysis. The number of variants in each of the 7 MAF bins (twice the number in each LD bin) can be found in Supplementary Figure 5.

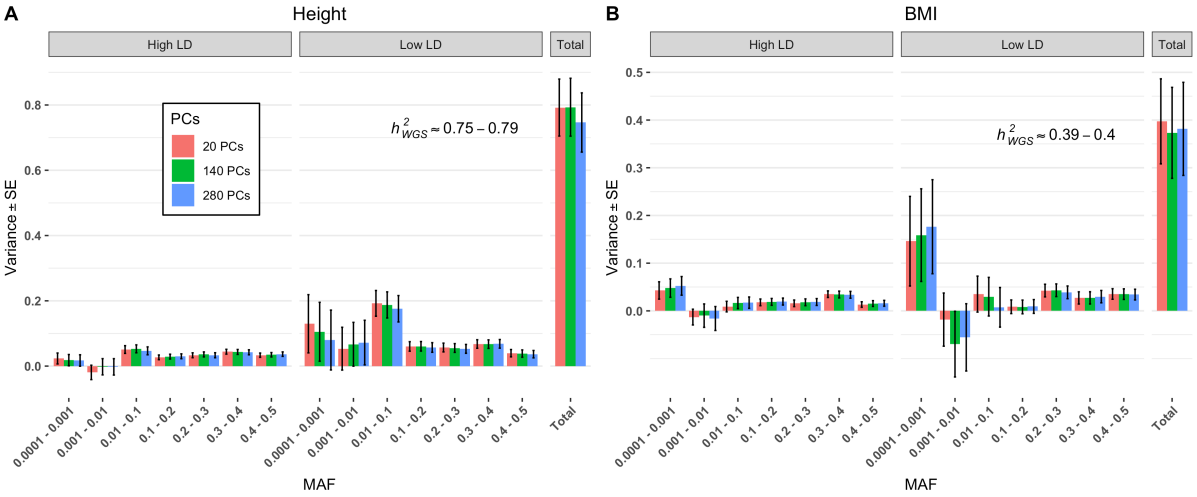

Figure 2: GREML-LDMS estimates from WGS data (~47.1M variants) stratified in 14 bins (7 MAF bins in 2 LD bins) with correction for 20 PCs (on HM3 SNPs), 140 PCs (20*7 MAF bins) and 280 PCs (20*14 bins). (A) Estimates for height with h^2_{WGS} at ~0.75 - ~0.79 (SE ~0.09). (B) Estimates for BMI with h_{WGS}^2 at ~0.39 - ~0.40 (SE 0.09 – 0.10). These estimates are consistent with previous pedigree estimates for height and BMI, with a large contribution from variants with MAF < 0.1. The number of variants in each of the 7 MAF bins (twice the number in each LD bin) is respectively, from the lowest to highest MAF bins, of 28.5M, 8.6M, 5.3M, 1.7M, 1.3M, 1.1M, 1.0M (Supplementary Table 4 and Supplementary Figure 7).

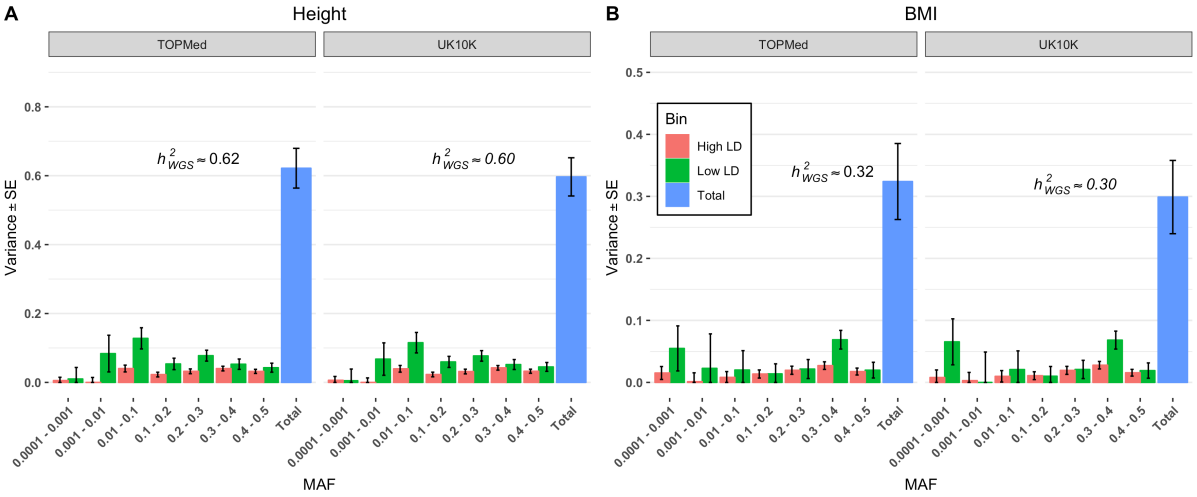

Figure 3: GREML-LDMS estimates from variants in common (~13.9M variants) between TOPMed and UK10K datasets stratified in 14 bins according to variants MAF and LD properties correcting for the first 20 PCs (computed from HM3 SNPs) using either TOPMed or UK10K LD and MAF reference. (A) Estimates of h^2_{WGS} s for height (~0.60 – 0.62 (SE 0.06)) or (B) BMI (~0.30 – 0.32 (SE 0.06)) are similar and independent of the LD and MAF reference. The number of variants in each of the 7 MAF bins (twice the number in each LD bin) is, respectively, from the lowest to highest MAF bins, of 3.6.M, 3.0M, 3.3M, 1.4M, 1.0M, 0.9M, 0.8M.

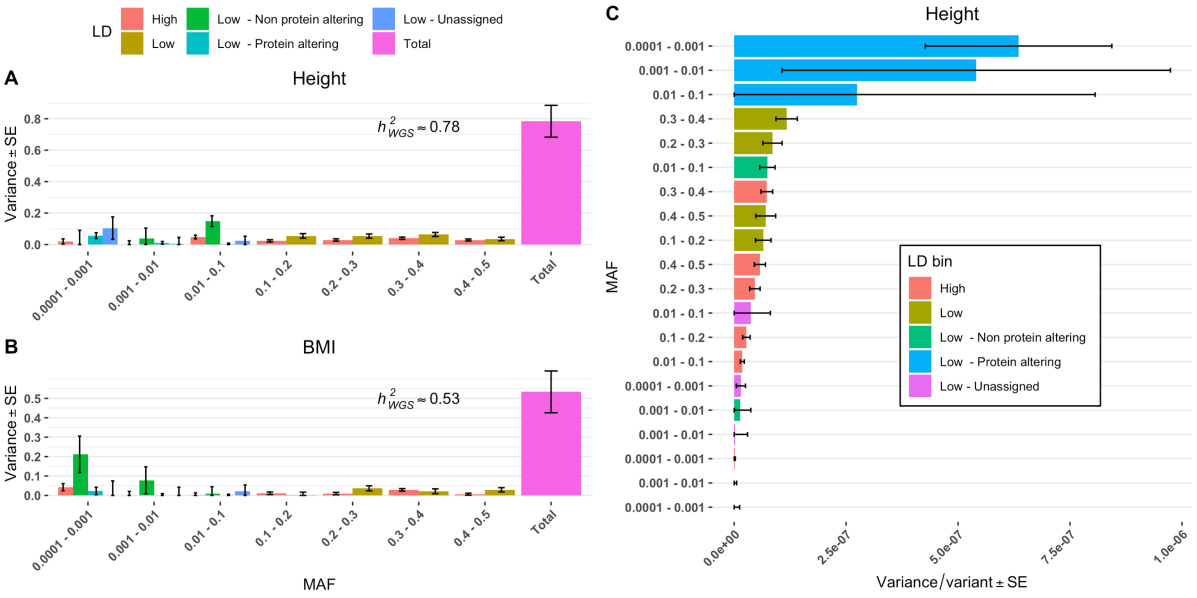

Figure 4: Estimates of h^2_{WGS} (~47.1M variants) by GREML-LDMS for height and BMI with the low-LD and low-MAF (<0.01) variants partitioned into 3 distinct categories according to the SnpEff putative effect of the variant (protein altering vs. non-protein altering, and a bin of unassigned variants), correcting for 20 PCs. There is a total of 17 genetic components in this analysis. (A) Results for height with an estimate of h^2_{WGS} at ~0.78 (SE 0.10). (B) Results for BMI with an estimate of h^2_{WGS} at ~0.53 (SE 0.11). (C) Variance explained per variant (the estimate of genetic variance divided by the variant number in each bin) of the low-MAF variants in the GREML-LDMS analysis on height. There is an apparent enrichment of heritability in the protein altering groupings (low LD) over non-protein altering (low LD) or high LD variants.

URLs

GCTA, http://cnsgenomics.com/software/gcta/#Overview. Plink, https://www.cog-genomics.org/plink2. UK10K, https://www.uk10k.org/. NHLBI, www.nhlbiwgs.org. TOPMed methods, https://www.nhlbiwgs.org/topmed-whole-genome-sequencing-project-freeze-5b-phases-1-and-2.

Online Methods

Data collection

In this study, we used Whole-Genome Sequence (WGS) data from the Trans-Omics for Precision Medicine (TOPMed) Program. The TOPMed program collects WGS data from different studies and centers, in the United States and elsewhere, in partnership with the National Heart, Lung, and Blood Institute (NHLBI, see URLs). The "freeze 5b" version of the data includes 54,499 samples containing ~470M SNPs and indels in the called variants files (as BCF, binary variant call format). These variants have been called on genome assembly GRCh38 as human genome reference (see URLs for methods). Participant consent was obtained for each of the 16 studies (Supplementary Table 1, Supplementary Table 6) containing Europeans samples in the freeze 5b as well as the associated phenotypes for height and BMI.

Quality control

We selected freeze 5b samples with height and BMI phenotypes available (N=38,466). After removing individuals under 18 years old, we had 38,291 adults left. For each sex within each of the 16 different studies included in this analysis (each cohort), we regressed the height and BMI according to their age and kept the residuals. Moreover, to remove differences in mean and standard deviation between sexes and among cohorts, we standardized the residuals by the standard deviation of each sex and cohort. The standardized residuals on height and BMI of each gender group of each cohort followed a distribution N(0,1) (Supplementary Figure 2). We also applied a rank-based inverse normal transformation on the BMI (BMI_{RINT}) after adjustment for age and sex. On the genotypes, we performed additional quality control steps on the data by excluding variants with genotypes missingness rate >0.05, Hardy-Weinberg equilibrium test P value $<1 \times 10^{-6}$, or with a minor allele frequency <0.0001 using PLINK v1.9 (see URLs, [26]). We also excluded individuals with sample missingness rate >0.05. Using HapMap phase 3 reference panels (HM3) variants and HapMap populations' references, we selected TOPMed samples within +/-6 standard deviation of the mean in principal component 1 and 2 for HapMap CEU population (Supplementary Figure 1). We also controlled for other type of stratification in our sample, by sequencing center (Supplementary Figure 1), study and sex. On the remaining samples with European ancestry (28,561 individuals), we built a genetic relatedness matrix based on pairwise genetic relationships using variants on HM3 reference panels and observed values ranging from -0.023 to 0.31, indicating that individuals with some degree of relatedness were in our dataset. We removed one of each pair of individuals with estimated genetic relatedness > 0.05. At the end of all the quality control steps, we retained 21,620 unrelated individuals of European ancestry and 47.1 million variants. MAF and LD distribution of the variants are shown in Supplementary Figure 7.

Statistical framework of the GREML analysis

The GREML analysis is based on the idea to fit the effects of all the SNPs as random effects by a mixed linear model (MLM),

$$\boldsymbol{Y} = \boldsymbol{XB} + \boldsymbol{g} + \boldsymbol{\varepsilon}$$

where $\boldsymbol{XB}$ is a vector of fixed effects such as age, sex and in our case the principal components of each subset of SNPs, $\boldsymbol{g}$ is an n x 1 vector of the total genetic effects captured by all the sequenced variants of all the individuals (with n being the sample size) and $\boldsymbol{\varepsilon}$ is a vector of residuals effects. From MLM, $\boldsymbol{g}$ follows a distribution $\boldsymbol{g} \sim N(0, \boldsymbol{A}\sigma_g^2)$ where $\boldsymbol{A}$ is a GRM interpreted as the genetic relationship between individuals. We estimate σ_g^2 using the REML algorithm [13].

The genetic relationship between individuals j and k (A_{jk}) was estimated by the following equation:

$$A_{jk} = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_{ij} - 2p_i)(x_{ik} - 2p_i)}{2p_i(1 - p_i)}$$

where x_{ij} the minor allele count for SNP i in individual j, N the number of SNPs, and p is the sample minor allele frequency. We refer to this method of estimating pairwise relationships as the average of ratios, where the ratio is the SNP covariance divided by SNP heterozygosity. By using the sample allele frequencies, A_{jk} does not represent a measure of kinship between two individuals, although the GRM should be highly correlated with the kinship matrix if we were to have full and accurate pedigree data on the entire sample [20].We calculated multiple GRMs based on subset of SNPs (stratified by MAF, LD, annotations, etc) and fit them as random effects according to a more general model:

$$\boldsymbol{Y} = \boldsymbol{XB} + \sum_{i=1}^{r}\boldsymbol{g}_i + \boldsymbol{\varepsilon}$$

where the phenotypic variance σ_P^2 is the sum of the residual variance and the variance of each of the i^{th} genetic factors (each with a corresponding GRM).

To compare the methods to calculate the generic relationship between individuals j and k we also used the ratio of total SNP covariance and the total SNP heterozygosity over loci (the ratio of averages) [20] from the following equation:

$$A_{jk} = \frac{\sum_{i=1}^{N}(x_{ij} - 2p_i)(x_{ik} - 2p_i)}{2\sum_{i=1}^{N}p_i(1 - p_i)}$$

Proportion of genetic variation captured by imputation

Previous REML estimates based on rare variants were conducted by imputing SNP chip data on a reference panel such as 1000 Genomes [10]. To check the consistency of our data set with previous estimates, we mimicked SNP chip data by selecting SNPs in our dataset present on Illumina HumanCore24 v1.0, GSA 24 V1.0 and Affymetrix Axiom UKB WCSG 34 arrays. We downloaded the list of SNPs of these arrays, the UCSC to Ensembl reference and the GRCh37 reference. With those, we selected subsets of the SNPs in common between

TOPMed dataset and each of the SNP arrays and used LiftOver to convert the datasets from GRCh38 to GRCh37 reference. To prepare the files for imputation, using BCFtools we migrated from the UCSC to Ensembl style chromosomes names, fixed the forward strand convention and flipped or removed SNPs not matching on the reference genome. After this merging and cleaning process, we retained the majority of the SNPs present on each array (Supplementary Table 2). We then imputed our dataset on the Sanger imputation server [16]. Imputation was performed in two stages. Prior to imputation, each chromosome was phased against HRC.r1.1 [16] reference using EAGLE2 v2.0.5 [27]. In a second stage, data were imputed using Positional Burrows-Wheeler Transform. On the imputed datasets, we extracted variants with imputation info score > 0.3 and performed another QC filtering removing variants with individual missingness rate > 0.05, Hardy-Weinberg equilibrium test *P*-value $<1 \times 10^{-6}$, MAF < 0.0001 or variant missingness rate > 0.05 After imputation and filtering, we had between ~15 and 20M SNPs left on each imputed dataset, with notably fewer variants from the Axiom array, most likely due to the higher proportion of low MAF variants on this array (Supplementary Table 3 and Supplementary Figure 5). On each imputed dataset, we stratified SNPs into 7 MAF bins (0.0001<MAF<0. 001, 0.001<MAF<0.01, 0.01<MAF<0.1, 0.1<MAF<0.2, 0.2<MAF<0.3, 0.3<MAF<0.4, and 0.4<MAF<0.5). For each of these 7 MAF bins, we independently calculated the LD score of the variants within a MAF bin on a sliding window of 10Mb using GCTA software [13]. We performed two types of LD binning, selecting variants based on their individual LD values or on their segment-based LD value (segment length = 200Kb) (Supplementary Figure 4). Each of the 7 MAF bins was divided into 2 more bins, one for variants with LD scores above the median value of the bin (high LD bin) and one for variants with LD score below median (low LD bin) (Supplementary Table 4). We then used GCTA to perform a GREML-LDMS analysis with the first 20 PCs calculated using HM3 SNPs from the WGS data set fitted as fixed effects and the variants in the 14 MAF and LD bins as 14 random-effect components (Figure 1). To assess the influence of imputation errors on variance estimates, we selected, for each of the 3 imputed data sets, the SNPs that were in both the imputed data set and the TOPMed WGS data set. We had 13.1M, 17.3M and 17.6M SNPs found in both TOPMed and Affymetrix Axiom UKB WCSG 34, Illumina HumanCore24 v1.0 and GSA 24 V1.0 imputeds data sets respectively. On each of these 6 subset data sets (3 different arrays and 3 different subsets from WGS data), we partitioned SNPs in 14 groupings, according to their MAF and LD scores, similarly to the previous analysis. We then ran a GREML-LDMS analysis on height and BMI for each data set with 20 principal components calculated from Hapmap3 SNPs fitted as fixed covariates.

GREML estimates from WGS data

Before estimating the proportion of phenotypic variance due to additive genetic factors from WGS data we initially wanted to check for consistency with previous studies and performed a single-component GREML analysis (GREML-SC approach) in GCTA using HM3 SNPs with 0 or the 20 PCs (calculates from the same set of SNPs) fitted as fixed effects. We chose a different analysis to estimate heritability for height and BMI when using the whole dataset. It has previously been shown [10] that a GREML-SC approach can give a biased estimate of h^2 if causal variants have a different MAF spectrum. Whereas a GREML-MS analysis fits in a single model multiple GRMs (one GRM for each MAF bin). We performed this analysis using 7 MAF bin (0.0001<MAF<0. 001, 0.001<MAF<0.01, 0.01<MAF<0.1, 0.1<MAF<0.2, 0.2<MAF<0.3, 0.3<MAF<0.4, and 0.4<MAF<0.5). For each MAF bin, we performed a PCA analysis and included the first 20 eigenvectors of the bin as fixed covariates in our GREML-MS analysis (140 PCs in total). To investigate the effect of PC correction on the REML estimates, we ran the GREML-MS analysis using multiple PCs corrections: without any PC correction; correcting with 20 PCs calculated from HM3 SNPs, and with 20 PCs calculated

from the variants of each MAF bin (Supplementary Figure 3). Subsequently, after running the GREML-MS analysis and investigating the influence of PCs correction, we ran a GREML-LDMS analysis. As with the GREML-MS approach, if two variants in the same GRM have different LD properties, this can lead to biased estimates of heritability. Similarly to what was done using data imputed from array SNPs, we defined 14 bins by splitting each of the 7 MAF bin into a high LD and a low LD one, according to their median LD value calculated on all variants present on each chromosome within a sliding window of 10Mb (Figure 2). To investigate how low-quality variants would bias estimates, we also conducted an analysis by using only the 20.9M variants that passed a support vector machine (SVM) classifier and passed additional filters on excess of heterozygosity and Mendelian discordancy. The SVM classifier was trained using variants present on genotyping arrays labelled as positive controls, and variants with many Mendelian inconsistencies labelled as negative controls. We also defined a subset of the whole data set with the same number of variants, MAF and LD properties as the variants that passed the SVM classifier. On each of those two subsets of the WGS data, we ran a 14 bins GREML-LDMS analysis (7 MAF bins and 2 LD bins, LD reference calculated from the whole data set).

To investigate the robustness of the assumptions on the relationship between MAF and effect size, we ran a GREML-LDMS analysis on the previously defined 14 MAF and LD bins using either the ratio of averages over loci or the average over loci of ratios methods to compute the GRMs. We further divided the bin with low LD and $0.0001 < MAF < 0.001$ into 2 bins with a similar number of variants (7.1M variants in each bin) and into 3 bins of (4.8M variants in each bin) and calculated the GRMs using both methods. We ran a GREML-LDMS analysis using the 15 and 16 GRMs.

Comparison with UK10K dataset

To ensure the GREML-LDMS estimates were not due to population stratification, we repeated the GREML-LDMS analysis using LD and MAF reference from another dataset. We converted UK10K WGS data [19] to GRCh38 reference coordinates using LiftMap, a wrapper Python script for LiftOver [28]. There are 3781 individuals in the UK10K dataset. As with TOPMed, we performed a quality control step on the genotypes using PLINK with the following filtering thresholds: individual and variant missingness > 0.05, Hardy-Weinberg equilibrium test P value $<1 \times 10^{-6}$, minor allele frequency < 0.0001, genotype missingness rate > 0.05; and retained 42.68M variants. On these variants, we selected the ones in common with the TOPMed dataset and obtained 13.9M variants in common. Using these 13.9M variants on a subset of UK10K dataset, we defined 7 MAF bins ($0.0001<MAF<0.001$, $0.001<MAF<0.01$, $0.01<MAF<0.1$, $0.1<MAF<0.2$, $0.2<MAF<0.3$, $0.3<MAF<0.4$, and $0.4<MAF<0.5$) that we further split in 14 bins according to their LD scores (above or below the median LD score value of each MAF group), based on the LD score of individual variant calculated in a window of 10Mb. We then ran a GREML-LDMS analysis for height and BMI after calculating GRMs from the TOPMed dataset with the 14 variant bins defined from the UK10K dataset. We also defined similar MAF and LD bins on a subset of TOPMed dataset, but using bins defined from TOPMed dataset. We ran both GREML-LDMS analyses (with variants bins defined from UK10K and TOPMed) by correcting for HM3 SNPs bins and by correcting by 20 PCs in each variant bin (Figure 3).

Enrichment analysis using the variant effect consequence

Lastly, using SnpEff annotations [21] and the LD and MAF bins previously defined from the GREML-LDMS analysis on the WGS data mentioned above, we further separated the low-

LD variants in each of the 0.0001<MAF<0.001, 0.001<MAF<0.01 and 0.01<MAF<0.1 bins into each of 3 more bins according to their predicted variant effects. The SnpEff variant effect annotations were divided in 4 categories according to their predicted effects on gene expression and protein translation. The 4 categories are based on the Sequence Ontology terms used in functional annotations (Supplementary Table 5). Putative effects on proteins can be "High" (protein truncating variants, frameshift variants, stop gained, and stop lost etc), "Moderate" (mostly non-synonymous variants), "Low" (mostly synonymous variants) or "Modifier" (mostly intronic and upstream or downstream regulatory variants). We merged variants having "High" and "Moderate" impacts in a "Protein altering" bin and variants having "Low" and "Modifier" impacts in a "Non-protein altering" bin. Variants with missing predicted effect annotation were grouped in an "Unassigned" bin. We then ran a GREML-LDMS analysis with the 20 PCs calculated from HM3 SNPs fitted as fixed effects on 17 bins in total (Figure 4). To compute the variance per SNP, we divided the variance explained by each bin by the number of variants in each bin. The standard error was obtained by dividing the standard error of the heritability estimate of the bin by the number of variants in the bin.

Acknowledgements

This research was supported by the Australian Research Council (DP160102400, FT180100186 and FL180100072), the Australian National Health and Medical Research Council (1113400 and 1078037), the US National Institutes of Health (R01MH100141), and the Sylvia & Charles Viertel Charitable Foundation. This study makes use of data from UK10K project (a full list of acknowledgements to this data set can be found below).

Whole genome sequencing (WGS) for the Trans-Omics in Precision Medicine (TOPMed) program was supported by the National Heart, Lung and Blood Institute (NHLBI). WGS for "NHLBI TOPMed: Genetics of Cardiometabolic Health in the Amish" (phs000956.v3.p1.c999) was performed at the Broad Institute of MIT and Harvard (HHSN268201500014C). WGS for TOPMed "NHLBI TOPMed: Trans-Omics for Precision Medicine Whole Genome Sequencing Project: ARIC" (phs001211.v1.p1.c999) was performed at the Broad Institute of MIT and at the Baylor Human Genome Sequencing Center (3R01HL092577-06S1, HHSN268201500015C, 3U54HG003273-12S). WGS for "NHLBI TOPMed: The Cleveland Clinic Atrial Fibrillation Study of the CV/Arrhythmia Biobank" (phs001189.v1.p1.c999) was performed at the Broad Institute of MIT and Harvard (3R01HL092577- 06S1). WGS for "NHLBI TOPMed: The Cleveland Family Study (WGS)" (phs000954.v2.p1.c999) was performed at the University of Washington Northwest Genomics Center (3R01HL098433-05S1). WGS for "NHLBI TOPMed: Cardiovascular Health Study" (phs001368.v1.p1.c999) was performed at the Baylor Human Genome Sequencing Center (HHSN268201500015C). WGS for "NHLBI TOPMed: Genetic Epidemiology of COPD (COPDGene) in the TOPMed Program" (phs000951.v2.p2.c999) was performed at the Broad Institute of MIT and Harvard and the University of Washington Northwest Genomics Center (HHSN268201500014C). WGS for "NHLBI TOPMed: Whole Genome Sequencing and Related Phenotypes in the Framingham Heart Study" (phs000974.v3.p2.c999) was performed at the Broad Institute of MIT and Harvard (3R01HL092577-06S1). WGS for "NHLBI TOPMed: GeneSTAR (Genetic Study of Atherosclerosis Risk)" (phs001218.v1.p1.c999) was performed at the Broad Institute of MIT and Harvard, at Macrogen Corp (HHSN268201500014C) and at Illumina (HL112064).WGS for "NHLBI TOPMed: Genetics of Lipid Lowering Drugs and Diet Network (GOLDN)" (phs001359.v1.p1.c999) was performed at the University of Washington Northwest Genomics Center (3R01HL104135-04S1). WGS for "NHLBI TOPMed: Heart and Vascular Health Study (HVH)" (phs000993.v2.p2.c999) was performed at the Broad Institute of MIT and Harvard and the Baylor Human Genome Sequencing Center (3R01HL092577-06S1, 3U54HG003273-12S2). WGS for "NHLBI TOPMed: Whole Genome Sequencing of Venous Thromboembolism (WGS of VTE)" (phs001402.v1.p1.c999) was performed at the Baylor Human Genome Sequencing Center (HHSN268201500015C, 3U54HG003273-12S2). WGS for "NHLBI TOPMed: MESA and MESA Family AA-CAC" (phs001416.v1.p1.c999) was performed at the Broad Institute of MIT and Harvard (3U54HG003067-13S1, HHSN268201500014C). WGS for "NHLBI TOPMed: MGH Atrial Fibrillation Study" (phs001062.v3.p2.c999) was performed at the Broad Institute of MIT and Harvard (3R01HL092577-06S1). WGS for "NHLBI TOPMed: The Vanderbilt AF Ablation Registry" (phs000997.v3.p2.c999) was performed at the Broad Institute of MIT and Harvard (3R01HL092577-06S1). WGS for "NHLBI TOPMed: The Vanderbilt Atrial Fibrillation Registry" (phs001032.v3.p2.c999) was performed at the Broad Institute of MIT and Harvard (3R01HL092577-06S1). WGS for "NHLBI TOPMed: Women's Health Initiative (WHI)" (phs001237.v1.p1.c999) was performed at the Broad Institute of MIT and Harvard (HHSN268201500014C). Centralized read mapping and genotype calling, along with variant quality metrics and filtering were provided by the TOPMed Informatics Research Center

(3R01HL-117626-02S1). Phenotype harmonization, data management, sample-identity QC, and general study coordination, were provided by the TOPMed Data Coordinating Center (3R01HL-120393-02S1). We gratefully acknowledge the studies and participants who provided biological samples and data for TOPMed.

COHORT SPECIFIC ACKNOWLEDGEMENTS:
Amish Research Program: This research has been supported in part by NIH grants U01 HL072515, U01 HL072515, R01 AG18728, and P30 DK072488.

Atherosclerosis Risk in Communities Study: The Atherosclerosis Risk in Communities study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services, under Contract nos. (HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700005I, HHSN268201700004I). The authors thank the staff and participants of the ARIC study for their important contributions.

Cardiovascular Health Study: This research was supported by contracts HHSN268201200036C, HHSN268200800007C, HHSN268201800001C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, and grants U01HL080295, U01HL130114, and R01HL059367 from the National Heart, Lung, and Blood Institute (NHLBI), with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided by R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Cleveland Family Study (CFS): CFS was supported by National Institutes of Health grants National Institutes of Health grants R01-HL046380-15, HL113338, and KL2-RR024990-05, and R35HL135818 [5-R01-HL046380, and 5-KL2-RR024990-05.

Framingham Heart Study: The Framingham Heart Study (FHS) acknowledges the support of contracts NO1-HC-25195 and HHSN268201500001I from the National Heart, Lung and Blood Institute and grant supplement R01 HL092577-06S1 for this research. We also acknowledge the dedication of the FHS study participants without whom this research would not be possible. Dr. Vasan is supported in part by the Evans Medical Foundation and the Jay and Louis Coffman Endowment from the Department of Medicine, Boston University School of Medicine.

Massachusetts General Atrial Fibrillation Study: This work was supported by grants from the National Institutes of Health to Dr. Ellinor (1RO1HL092577, R01HL128914, K24HL105780). Dr. Ellinor is also supported by the American Heart Association (18SFRN34110082) and by the Fondation Leducq (14CVD01).

Multi-Ethnic Study of Atherosclerosis: The MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-001079, UL1-TR-001420, UL1-TR-

001881, and DK063491. Whole genome sequencing of the MESA cohort was funded through the Trans-Omics for Precision Medicine (TOPMed) Program of the National Heart, Lung, and Blood Institute. General study coordination was provided by the TOPMed Data Coordinating Center (3R01HL-120393-02S1).

The Johns Hopkins Genetic Study of Atherosclerosis Risk (GeneSTAR): was supported by grants from the National Institutes of Health through the National Heart, Lung, and Blood Institute (HL49762, HL071025, U01HL72518, HL087698, HL092165, HL099747, K23HL105897, HL112064) and the National Institute of Nursing Research (NR0224103), by a grant from the National Center for Research Resources (M01-RR000052) to the Johns Hopkins General Clinical Research Center, and by a grant from the National Center for Research Resources and the National Center for Advancing Translational Sciences (UL1 RR 025005) to the Johns Hopkins Institute for Clinical and Translational Research.

The Women's Health Initiative (WHI): The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, and HHSN268201600004C. Investigators: *Program Office*: (National Heart, Lung, and Blood Institute, Bethesda, Maryland) Jacques Rossouw, Shari Ludlam, Joan McGowan, Leslie Ford, and Nancy Geller. *Clinical Coordinating Center*: (Fred Hutchinson Cancer Research Center, Seattle, WA) Garnet Anderson, Ross Prentice, Andrea LaCroix, and Charles Kooperberg. *Investigators and Academic Centers*: (Brigham and Women's Hospital, Harvard Medical School, Boston, MA) JoAnn E. Manson; (MedStar Health Research Institute/Howard University, Washington, DC) Barbara V. Howard; (Stanford Prevention Research Center, Stanford, CA) Marcia L. Stefanick; (The Ohio State University, Columbus, OH) Rebecca Jackson; (University of Arizona, Tucson/Phoenix, AZ) Cynthia A. Thomson; (University at Buffalo, Buffalo, NY) Jean Wactawski-Wende; (University of Florida, Gainesville/Jacksonville, FL) Marian Limacher; (University of Iowa, Iowa City/Davenport, IA) Jennifer Robinson; (University of Pittsburgh, Pittsburgh, PA) Lewis Kuller; (Wake Forest University School of Medicine, Winston-Salem, NC) Sally Shumaker; (University of Nevada, Reno, NV) Robert Brunner. *Women's Health Initiative Memory Study*: (Wake Forest University School of Medicine, Winston-Salem, NC) Mark Espeland

UK10K (EGA accessions: EGAS00001000108 and EGAS00001000090): was funded by the Wellcome Trust award WT091310. Twins UK (TUK): TUK was funded by the Wellcome Trust and ENGAGE project grant agreement HEALTH-F4-2007–201413. The study also receives support from the Department of Health via the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St. Thomas' NHS Foundation Trust in partnership with King's College London. Dr Spector is an NIHR senior Investigator and ERC Senior Researcher. Funding for the project was also provided by the British Heart Foundation grant PG/12/38/29615 (Dr Jamshidi). A full list of the investigators who contributed to the UK10K sequencing is available from https://www.uk10k.org/

Competing interests

Dr. Ellinor is supported by a grant from Bayer AG to the Broad Institute focused on the genetics and therapeutics of cardiovascular diseases. Dr. Ellinor has also served on advisory boards or consulted for Bayer AG, Quest Diagnostics and Novartis.

References

1 Fisher, R. A. XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh* **52**, 399-433, doi:10.1017/s0080456800012163 (1918).

2 Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* **42**, 565-569, doi:10.1038/ng.608 (2010).

3 Visscher, P. M., Brown, M. A., McCarthy, M. I. & Yang, J. Five years of GWAS discovery. *Am J Hum Genet* **90**, 7-24, doi:10.1016/j.ajhg.2011.11.029 (2012).

4 Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* **47**, 1228-1235, doi:10.1038/ng.3404 (2015).

5 Speed, D. *et al.* Reevaluation of SNP heritability in complex human traits. *Nat Genet* **49**, 986-992, doi:10.1038/ng.3865 (2017).

6 Lynch, M. & Walsh, B. *Genetics and analysis of quantitative traits*. (Sinauer, 1998).

7 MacArthur, J. *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res* **45**, D896-D901, doi:10.1093/nar/gkw1133 (2017).

8 Gazal, S. *et al.* Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. *Nat Genet* **49**, 1421-1427, doi:10.1038/ng.3954 (2017).

9 Zeng, J. *et al.* Signatures of negative selection in the genetic architecture of human complex traits. *Nat Genet* **50**, 746-753, doi:10.1038/s41588-018-0101-4 (2018).

10 Yang, J. *et al.* Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat Genet* **47**, 1114-1120, doi:10.1038/ng.3390 (2015).

11 Zuk, O., Hechter, E., Sunyaev, S. R. & Lander, E. S. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci U S A* **109**, 1193-1198, doi:10.1073/pnas.1119675109 (2012).

12 Young, A. I. *et al.* Relatedness disequilibrium regression estimates heritability without environmental bias. *Nat Genet* **50**, 1304-1310, doi:10.1038/s41588-018-0178-9 (2018).

13 Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76-82, doi:10.1016/j.ajhg.2010.11.011 (2011).

14 Yang, J. *et al.* Genome partitioning of genetic variation for complex traits using common SNPs. *Nat Genet* **43**, 519-525, doi:10.1038/ng.823 (2011).

15 Evans, L. M. *et al.* Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. *Nat Genet* **50**, 737-745, doi:10.1038/s41588-018-0108-x (2018).

16 McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* **48**, 1279-1283, doi:10.1038/ng.3643 (2016).

17 Elks, C. E. *et al.* Variability in the heritability of body mass index: a systematic review and meta-regression. *Front Endocrinol (Lausanne)* **3**, 29, doi:10.3389/fendo.2012.00029 (2012).

18 Mitt, M. *et al.* Improved imputation accuracy of rare and low-frequency variants using population-specific high-coverage WGS-based imputation reference panel. *Eur J Hum Genet* **25**, 869-876, doi:10.1038/ejhg.2017.51 (2017).

19 UK10K Consortium. The UK10K project identifies rare variants in health and disease. *Nature* **526**, 82-90, doi:10.1038/nature14962 (2015).

20 Goudet, J., Kay, T. & Weir, B. S. How to estimate kinship. *Mol Ecol* **27**, 4121-4135, doi:10.1111/mec.14833 (2018).

21 Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80-92, doi:10.4161/fly.19695 (2012).

22 Visscher, P. M. *et al.* Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated samples. *PLoS Genet* **10**, e1004269, doi:10.1371/journal.pgen.1004269 (2014).

23 Shihab, H. A. *et al.* An integrative approach to predicting the functional effects of non-coding and coding sequence variation. *Bioinformatics* **31**, 1536-1543, doi:10.1093/bioinformatics/btv009 (2015).

24 Yengo, L. *et al.* Imprint of assortative mating on the human genome. *Nature Human Behaviour* **2**, 948-954, doi:10.1038/s41562-018-0476-3 (2018).

25 Uricchio, L. H., Zaitlen, N. A., Ye, C. J., Witte, J. S. & Hernandez, R. D. Selection and explosive growth alter genetic architecture and hamper the detection of causal rare variants. *Genome Res* **26**, 863-873, doi:10.1101/gr.202440.115 (2016).

26 Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7, doi:10.1186/s13742-015-0047-8 (2015).

27 Loh, P. R. *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet* **48**, 1443-1448, doi:10.1038/ng.3679 (2016).

28 Hinrichs, A. S. *et al.* The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res* **34**, D590-598, doi:10.1093/nar/gkj144 (2006).