DOI:
10.1038/nature14618

Document Version
Peer reviewed version

Published in final edited form as:
Nature. 2015 July 23; 523(7561): 459–462. doi:10.1038/nature14618.

Directional dominance on stature and cognition in diverse human populations

A full list of authors and affiliations appears at the end of the article.
[#] These authors contributed equally to this work.

Abstract

Homozygosity has long been associated with rare, often devastating, Mendelian disorders[1] and Darwin was one of the first to recognise that inbreeding reduces evolutionary fitness[2]. However, the effect of the more distant parental relatedness common in modern human populations is less well understood. Genomic data now allow us to investigate the effects of homozygosity on traits of public health importance by observing contiguous homozygous segments (runs of homozygosity, ROH), which are inferred to be homozygous along their complete length. Given the low levels of genome-wide homozygosity prevalent in most human populations, information is required on very large numbers of people to provide sufficient power[3,4]. Here we use ROH to study 16 health-related quantitative traits in 354,224 individuals from 102 cohorts and find statistically significant associations between summed runs of homozygosity (SROH) and four complex traits: height, forced expiratory lung volume in 1 second (FEV1), general cognitive ability (*g*) and educational attainment (nominal $p<1 \times 10^{-300}$, 2.1×10^{-6}, 2.5×10^{-10}, 1.8×10^{-10}). In each case increased homozygosity was associated with decreased trait value, equivalent to the offspring of first cousins being 1.2 cm shorter and having 10 months less education. Similar effect sizes were found across four continental groups and populations with different degrees of genome-wide homozygosity, providing convincing evidence for the first time that homozygosity, rather than confounding, directly contributes to phenotypic variance. Contrary to earlier reports in substantially smaller samples[5,6], no evidence was seen of an influence of genome-wide homozygosity on blood pressure and low density lipoprotein (LDL) cholesterol, or ten other cardio-metabolic traits. Since directional dominance is predicted for traits under directional evolutionary selection[7], this study provides evidence that increased stature and cognitive function have been positively selected in human evolution, whereas many important risk factors for late-onset complex diseases may not have been.

Correspondence and requests for materials should be addressed to jim.wilson@ed.ac.uk.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature

Competing financial interests: GP is a co-founder of CAVADIS B.V. SWvL is a former employee of CAVADIS B.V. BMP serves on the Data and Safety Monitoring Board of a clinical trial funded by the manufacturer (Zoll Lifecor) and on the Yale Open Data Access Project funded by Johnson & Johnson. NPo has received financial support and consultancy fees from several pharmaceutical companies that manufacture either blood-pressure-lowering or lipid-lowering agents or both. PS has received research awards from Pfizer. No other authors declared a conflict of interest.

Europe PMC Funders Author Manuscripts

Inbreeding influences complex traits through increases in homozygosity and corresponding reductions in heterozygosity, most likely resulting from the action of deleterious (partially) recessive mutations[8]. For polygenic traits, a systematic association with genome-wide homozygosity is not expected when dominant alleles at some loci increase the trait value while others decrease it. Rather, dominance must be biased in one direction on average over all causal loci, for instance to decrease the trait. Such directional dominance is expected to arise in evolutionary fitness-related traits due to directional selection[8]. Studies of genome-wide homozygosity thus have the potential to reveal the non-additive allelic architecture of a trait and its evolutionary history. Historically inbreeding has been measured using pedigrees[9]. However, such techniques cannot account for the stochastic nature of inheritance, nor are they practical for the capture of the distant parental relatedness present in most modern day populations. High density genome-wide single nucleotide polymorphism (SNP) array data can now be used to assess genome-wide homozygosity directly, using genomic runs of homozygosity (ROH). Such runs are inferred to be homozygous-by-descent and are common in human populations[10-11]. SROH is the sum of the length of these ROH, in megabases of DNA. F_{ROH} is the ratio of SROH to the total length of the genome. Like pedigree-based F (with which it is highly correlated[3]), F_{ROH} estimates the probability of being homozygous at any site in the genome. F_{ROH} has been shown to vary widely within and between populations[12] and is a powerful method of detecting genome-wide homozygosity effects[13].

We found marked differences by geography and demographic history in both the population mean SROH and the relationship between SROH and NROH (the numbers of separate runs of homozygosity). (Fig. 1). As observed previously[3,12,14], isolated populations have a higher burden of ROH whereas African heritage populations have the least homozygosity.

We studied β_{FROH}, defined as the effect of F_{ROH} on 16 complex traits of biomedical importance (Fig. 2). For height, FEV1 (forced expiratory volume in one second, a measure of lung function), educational attainment (EA) and *g* (a measure of general cognitive ability derived from scores on several diverse cognitive tests), we found the effect sizes were greater than two intra-sex standard deviations (SD), with p-values all less than 10^{-5}. Thus the associations could not plausibly be explained by chance alone (Table 1; see Extended data Figs. 1-4 for Forest plots of individual traits; Supplementary Table 1 for SD). To ensure that the results were not driven by a few outliers, we repeated the analysis excluding extreme sub-cohort trait results. In all cases the effect sizes and their significance remained similar or increased (see Supplementary Table 2 for comparisons with and without outliers). After exclusion of outliers, these effect sizes translate into a reduction of 1.2 cm in height and 137 ml in FEV1 for the offspring of first cousins, and into a decrease of 0.3 SD in *g* and 10 months less educational attainment.

We performed a number of analyses to exclude confounding. Whilst SROH is wholly a genetic effect, its inheritance is entirely non-additive. Therefore, unlike in genome-wide association, an association with population genetic structure or co-segregation of additive genome-wide polygenic effects and SROH (as opposed to SNPs in a GWAS) are not expected as a matter of course, except in the case of siblings. However, confounding could still theoretically arise as discussed below. We therefore assessed this by conducting

stratified and covariate analyses. We found effects of similar magnitude and in the same direction for all four traits across isolated and non-isolated European, Finnish, African, Hispanic, East Asian and South and Central Asian populations (Extended Data Fig. 5a, Supplementary Table 3). We further tested whether the effect sizes were similar when cohorts were split into more and less homozygous groups. The effect sizes were very similar even though the degree of homozygosity (and variation in homozygosity) varied 3-10-fold between the two strata (depending on which cohorts contributed to the trait; Extended Data Fig. 5b). This suggests a broadly linear relationship with SROH. In general confidence intervals overlap for stratified estimates, suggesting differences arose due to sampling variance. Larger confidence intervals for some estimates reflect the lower power of some strata, in turn reflecting the sample size and degree of homozygosity of those strata (e.g. the wider confidence intervals for estimates of Educational Attainment β_{FROH} for Finnish and African strata). Finally, we fitted educational attainment as a proxy for potential confounding by socio-economic status; this covariate was available in sufficient (47) cohorts to maintain power. The estimated effect sizes for height, FEV1 and *g* all reduced (17%, 18% and 35%, Fig. Extended Data Fig. 5c), but this might have been expected given the known covariance between these three traits and EA, and the association we found between educational attainment and F_{ROH}. We found very small differences (3-11% reductions) in estimated β_{FROH} (Extended Data Fig. 6, supplementary table 4), when comparing the fitting of polygenic mixed models as opposed to fixed-effect-only models, again suggesting that confounding (in this case due to polygenic effects due to recent common ancestry) was not substantially affecting the results.

Despite the observed 17-35% reductions in estimated effect sizes for F_{ROH} on height, FEV1 and *g*, when fitting educational attainment as a covariate, the persistence of an effect suggests that most of the signals we observe are genetic. The consistency of effects with and without fitting relatedness and in particular in populations with very different degrees of homozygosity, all appear inconsistent with confounding due to environmental or additive genetic effects. As does the broad similarity in effect sizes across continents, although the relatively smaller numbers of cohorts of non-European descent meant we had limited power to detect inter-continental differences in effect sizes.

It is also interesting to consider the potential influence of assortative mating, which is commonly observed for human stature, cognition and education. The phenotypic extremes could be more genetically similar to each other and hence the offspring more homozygous, even if the highly polygenic trait architectures reduce this effect. However, at least in its simplest balanced form, the increase in genetic similarity would be equal at both ends of the phenotypic distribution, leading to no linear association between such genetic similarity and the trait; both tall and short people would be more homozygous. Furthermore, humans also mate assortatively on body mass index (BMI), for which we see no effect. A more complex possibility, a form of reverse-causality, could arise when subjects from one trait extreme (e.g. more educated people) are on average more geographically mobile, and thus have less homozygous offspring, with those offspring in turn inheriting the trait extreme concerned[15]. We do not think that this mechanism can account for our results, since it does not readily explain the constancy of our results under different models, especially the similarity in

β_{FROH} for either more or less homozygous populations. Moreover, we observe similar effects in multiple single village cohorts, and the Amish and Hutterites, where there is no geographic structure and/or no sampling of immigrants, hence such confounding by differential migration cannot occur.

Our estimate for the effect of homozygosity in height is consistent with previous work: genomic[4] and pedigree[16] studies have shown genome-wide homozygosity effects on stature with similar effect sizes (0.01 increase in F decreases height by 0.037 SD[16] versus 0.029 SD in the present study). We speculate that homozygosity is acting on a shared endophenotype of torso size which we detect in the height and FEV1 traits. The fact that the FEV1/FVC (forced vital capacity) ratio is not associated with ROH points to the effect being on lung/chest size rather than airway calibre. The cognition effects cannot be wholly generated by height as an intermediate cause, given the greater proportion of variance explained for cognition, although we note that the correlation between height and cognition is 0.16 (SE 0.01), and the genetic correlation (the correlation in additive genetic values) is 0.28 (SE 0.09; ref 17). Height is the canonical human complex trait, highly heritable and polygenic, with 697 genome-wide significant variants in 423 loci explaining 20% of the heritability and all common variants predicted to explain 60% of the heritability[18]. Most of the genetic architecture appears to be additive in nature, however ROH analysis reveals a distinct directional dominance component.

Our genomic confirmation of directional dominance for *g* and discovery of genome-wide homozygosity effects on educational attainment in a wide range of human populations adds to our knowledge of the genetic underpinnings of cognitive differences, which are currently thought to be largely due to additive genetic effects[19]. Our findings go beyond earlier pedigree-based analyses of recent consanguinity to demonstrate that the observed effect of genome-wide homozygosity is not a result of confounding and influences demographically diverse populations across the globe. The estimated effect size is consistent with pedigree data (0.01 increase in F decreases *g* by 0.046 SD in our analysis and 0.029-0.048 SD in pedigree-based studies)[20]. It is germane to note that one extreme of cognitive function, early onset cognitive impairment, is strongly influenced by deleterious recessive loci[21], so we can speculate that an accumulation of recessive variants of weaker effect may influence normal variation in cognitive function. Although increasing migration and panmixia have generated a secular trend in decreasing homozygosity[22], the Flynn effect, wherein succeeding generations perform better on cognitive tests than their predecessors[23], cannot be explained by our findings, because the intergenerational change in cognitive scores is much larger than the differences in homozygosity would predict. Likewise, the genome-wide homozygosity effect on height cannot explain a significant proportion of the observed inter-generational increases[24].

Inbreeding depression, which arises from the effect of genome-wide homozygosity, is ubiquitous in plants and is seen for numerous fitness-related traits in animals[25], but we observed no effect for the 12 other mainly cardio-metabolic traits in which variation is strongly age-related. This suggests that previous reports in ecological studies or substantively smaller studies using pedigrees or relatively small numbers of genetic markers may have been false positives[5,6]. The lack of directional dominance on these traits does not,

however, rule out a recessive component, as recessive variants acting in different directions will cancel out. Dominance variance is predicted to be greater for late-onset fitness traits[26], so the lack of genome-wide homozygosity effects in the cardio-metabolic traits may be due to lack of directional dominance. ROH analyses within specific genomic regions are warranted to map recessive effects even when there is no genome-wide directional dominance. Such recessive effects have been observed for a subset of cardiovascular risk factors[27] and expression traits[28].

We have demonstrated the existence of directional dominance on four complex traits (stature, lung function, cognitive ability and educational attainment) whilst showing any effect on the other 12 health-related traits is at least almost an order of magnitude smaller or non-existent. This directional dominance implies that size and cognition (like schizophrenia protective alleles[29]) have been positively selected in human history – or at least that some variants increasing these traits contribute to fitness. However, the lack of any evidence for an association between many late onset cardiovascular disease risk factors and ROH is perhaps surprising and suggests testing directly for an association between ROH and disease outcome. The magnitude of genome-wide homozygosity effects is relatively small in all cases, thus Darwin's supposition[30] of "any evil [of inbreeding] being very small" is substantiated.

METHODS

Outline

Our aim was to look for an association between a genetic effect (SROH) and 16 complex traits. Our approach followed best practice genome-wide association meta-analysis (GWAMA) protocols, where applicable, except we had only one genetic effect to test.

Cohorts were invited to join based on known previous participation in GWAMA and willingness to participate. 159 sub-cohorts were created from 102 population-based or case-control genetic studies, by separating different genotyping arrays, cases and controls or ethnic sub-groups to ensure each sub-cohort was homogeneous. Within each of the 159 sub-cohorts we measured the association between SROH and trait using the following model. Where a sub-cohort had been ascertained on the basis of a disease status associated with a particular trait, that sub-cohort was excluded from the corresponding trait analysis.

Phenotype was regressed on genetic effect and known relevant covariates within each cohort, under the model specified in Equation 1. The estimated genetic effect of SROH was then meta-analysed using inverse variance meta-analysis.

$$Y=\mu+b_1\,\mathrm{SROH}+b_2\,\mathrm{age}+b_3\,\mathrm{sex}+b_4\,\mathrm{PC1}+b_5\,\mathrm{PC2}+b_6\,\mathrm{PC3}+e \quad \text{Equation (1)}$$

Where Y is the vector of trait values, μ the intercept, b_1 the effect of SROH and $b_{2\text{-}6}$ the effect of covariates. PC1 – PC3, the post quality control within-cohort principal components of the cohort's relationship matrix and e the residual. Relationship matrices were determined genomically by each cohort using genome wide array data. In addition, any other cohort-specific covariates known to be associated with the trait, including further principal

components, and any trait-specific covariates and stratifications, such as medication and smoking status, were fitted as specified below. SROH was the sum of ROH called, with a length of at least 1.5 Mb using PLINK[31].

As is routine in GWAMA, for family-based studies only, we also fitted an additional term to account for additive genetic values and relatedness, using grammar+ type residuals and full hierarchical mixed modeling using GenABEL[32] and hglm[33], as specified in equation 2.

$$Y=\mu+b_1\,\mathrm{SROH}+b_2\,\mathrm{age}+b_3\,\mathrm{sex}+b_4\,\mathrm{PC1}+b_5\,\mathrm{PC2}+b_6\,\mathrm{PC3}+\mathrm{Z}a \quad \text{Equation (2)}$$

Where *a* is the additive genetic value of each individual. Var(*a*) is assumed to be proportionate to the Genomic Relationship matrix (GRM) (a pedigree relationship matrix was used in the Framingham Heart Study) . Z is the identity matrix.

We then meta-analysed the regression coefficients (b_1) of traits on SROH for the 159 subcohorts.

Cohort Recruitment

Data from 102 independent genetic epidemiology studies of adults were included. All subjects gave written informed consent and studies were approved by the relevant research ethics committees. Homogeneous sub-cohorts were created for analysis on the basis of ethnicity, genotyping array or other factors. Where a cohort had multiple ethnicities, sub-cohorts for each separate ethnicity were created and analysed separately. In all cases European-, African-, South or Central Asian-, East Asian- and Hispanic-heritage individuals were separated. In some cases sub-categories such as Ashkenazi Jews were also distinguished. Ethnic outliers were excluded, as were the second of any monozygotic twins and pregnant subjects. Continental ancestry of cohorts participating in each trait study is presented in Extended data Table 1. Cohort genotyping and summary information are shown in Supplementary Table 6, with age, sex, trait and homozygosity summary statistics given in Supplementary Tables 9,10,, and 11.For case-control and trait extreme studies, patients or extreme-only sub-cohorts were analysed separately to controls. Where case status was associated with the trait under analysis the sub-cohort was excluded from that study (see below).

Subjects within a sub-cohort were genotyped using the same SNP array, or where two very similar arrays were used (e.g. Illumina OmniExpress and IlluminaOmni1), the intersection of SNPs on both arrays – provided the intersection exceeded 250,000 SNPs. Where a study used two different genotyping arrays, separate subcohorts were created for each array, and analysis was done separately. Paediatric cohorts were not included.

Genotyping

All subjects were genotyped using high density genome-wide (>250,000 SNP) arrays, from Illumina, Affymetrix or Perlegen. Custom arrays were not included. Each study's usual array-specific genotype quality control standards for genome-wide association were used and are shown in Supplementary Table 6. Only autosomal data were analysed.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Phenotyping

We studied 16 quantitative traits which are widely available and represent different domains related to health, morbidity and mortality: height, body mass index (BMI), waist : hip ratio (WHR), diastolic and systolic blood pressure (DBP, SBP), fasting plasma glucose (FPG), fasting insulin (FI), Haemoglobin A1c (HbA1c), total-, HDL- and LDL-cholesterol, triglycerides, forced expiratory volume in 1 second (FEV1), ratio of FEV1 to forced vital capacity (FVC), general cognitive ability (*g*) and years of educational attainment (EA). Phenotypic QC was performed locally to assess the accuracy and distribution of phenotypes and covariates. Further covariates were included when the relevant GWAS consortium also included them. The trait categories were anthropometry, blood pressure, glycaemic traits, classical lipids, lung function, cognitive function and educational attainment, following models in the GIANT[34], ICBP[35], MAGIC[36], CHARGE[37], Spirometa[38] and SSGAC[39] consortia. The model for FEV1 did not include height as a covariate. Effect sizes for FEV1 therefore include size effects that also underpin height. Studies assembled files containing study traits and the following covariates: sex, age, first three principal components of ancestry, lipid-lowering medication, ever-smoker status, anti-hypertensive medication, diabetes status and year of birth (YOB). Educational attainment was defined in accordance with the ISCED 1997 classification (UNESCO), leading to seven categories of educational attainment that are internationally comparable[39]. LDL values estimated using Friedewald's equation were accepted. Cohorts without fasting samples did not participate in the LDL-cholesterol, triglycerides, fasting insulin or fasting plasma glucose analyses. Cohorts with semi-fasting samples fitted a categorical or quantitative fasting time variable as a covariate. Subjects with less than 4 hours fasting were not included.

Where subjects were ascertained, for example, on the basis of hypertension, that sub-cohort was excluded from analysis of traits associated with the disorder, for example blood pressure. The traits excluded from meta-analysis are as follows: ascertainment on type-2-diabetes, thus fasting insulin, HbA1c, fasting plasma glucose excluded; ascertainment on hypertension, thus blood pressures excluded; ascertainment on venous thrombosis or coronary artery disease, thus blood lipids excluded; ascertainment on obesity or the metabolic syndrome, thus blood lipids, body mass index, waist-hip ratio, fasting insulin and fasting plasma glucose excluded.

Somewhat unusually for a large consortium meta-analysis, the majority of the analysis after initial genotype and phenotype QC was performed by a pipeline of standardised R and shell scripts, to ensure uniformity and reduce the risk of errors and ambiguities (available at https://www.wiki.ed.ac.uk/display/ROHgen/Analysis+Plan+production+release+3.0). The pipeline was used for all stages from this point onwards.

Calling Runs of Homozygosity

SNPs with more than 3% missingness across individuals or with a minor allele frequency less than 5% were removed. ROH were defined as runs of at least 50 consecutive homozygous SNPs spanning at least 1500 kb, with less than a 1000 kb gap between adjacent ROH and a density of SNP coverage within the ROH of no more than 50 kb/SNP, with one heterozygote and 5 no calls allowed per window, and were called using PLINK[31], with the

following settings --homozyg-window-snp 50 --homozyg-snp 50 --homozyg-kb 1500 --homozyg-gap 1000 --homozyg-density 50 --homozyg-window-missing 5 --homozyg-window-het 1. The same criteria were used by McQuillan *et al.*[3], except SNP density has been relaxed to avoid regions of sparser coverage (still including 50 SNPs) being missed. The sum of runs of homozygosity was then calculated (SROH) . F was calculated as SROH/ (3×10^9 ROH) reflecting the length of the autosomal genome. Copy number variants (CNV) are known to influence cognition[40]; however, prior calling of CNV and ROH in one of our cohorts reduced the SROH by only 0.3%[3], making it implausible that deletions called as ROH influence our findings.

ROH called from different genotyping arrays

We show that SROH called with these parameters is relatively insensitive to the density and type of array used (Extended data Fig. 7). We used 2.5 million SNPs available for 851 HapMap and 1000 Genomes Project[41] samples from multiple continents to investigate the effect of array when using our ROH-calling parameters in plink. The dataset included samples of African, European, admixed American, South and East Asian heritage. By subsampling SNPs from the 2.5 million we created array data for the commonly used Illumina CNV370 and OmniExpress beadchips and the Affymetrix6 array for each individual (see Supplementary Table 7 for details of the SNP numbers). The correlation in SROH using different arrays on the same individuals was 0.93-0.94 for all pairwise chip comparisons.

Trait association with SROH

The association between trait and SROH was calculated using a linear model in accordance with equation 1. Additional covariates were fitted for some analyses (shown below) or for some cohorts where analysts were aware of study specific effects (e.g. study centre). For BMI, WHR, FEV1, FEV1/FVC and *g*, trait residuals were calculated for the model excluding SROH, these residuals were then rank-normalised and the effect of SROH on these rank-normalised residuals estimated. Triglycerides and fasting insulin were natural log transformed. Additional covariates were as follows: age^2 was included as a covariate for all traits apart from height and *g*. BMI was included as a covariate for WHR, SBP, DBP, FPG, FI and HbA1c. YOB was included as a covariate for educational attainment and ever-smoking for FEV1 and FEV1/FVC. Where a subject was known to be taking lipid-lowering medication, total cholesterol was adjusted by dividing by 0.8. Similarly, where a subject was known to be taking anti-hypertensive medication, SBP and DBP measurements were increased by 15 and 10 mm Hg, respectively.

Where the cohort was known to have significant kinship, genetic relatedness was also fitted, using the mixed model, in accordance with equation 2. The polygenic model was fitted in GenABEL using the fixed covariates and the genomic relationship matrix[32]. GRAMMAR+ (GR+) (ref. 42) residuals were then fitted to SROH as well as the full mixed model being fitted simultaneously, using GenABEL's hierarchical generalised linear model (HGLM) function[33]. Populations with kinship thus potentially had three estimates of β_{FROH}: using fixed effects only, and using the mixed model approaches, (GR+ and HGLM) for SROH.

To investigate potential confounding, where available, EA was added as an ordinal covariate and all models rerun, giving revised estimates of β_{FROH}. This is potentially an over adjustment for *g* due to the phenotypic and genetic correlations with EA[43]. However it must be recognised that EA does not capture all potential environmental confounding.

Cohort phenotypic means and standard deviations were checked visually for inter-cohort consistency, with apparent outliers then being corrected (e.g. due to units or incorrectly specified missing values), explained (e.g. due to different population characteristics) or excluded. Individual sub-cohort trait means and standard deviations are tabulated in Supplementary Table 9 and age and gender information is in Supplementary Table 10.

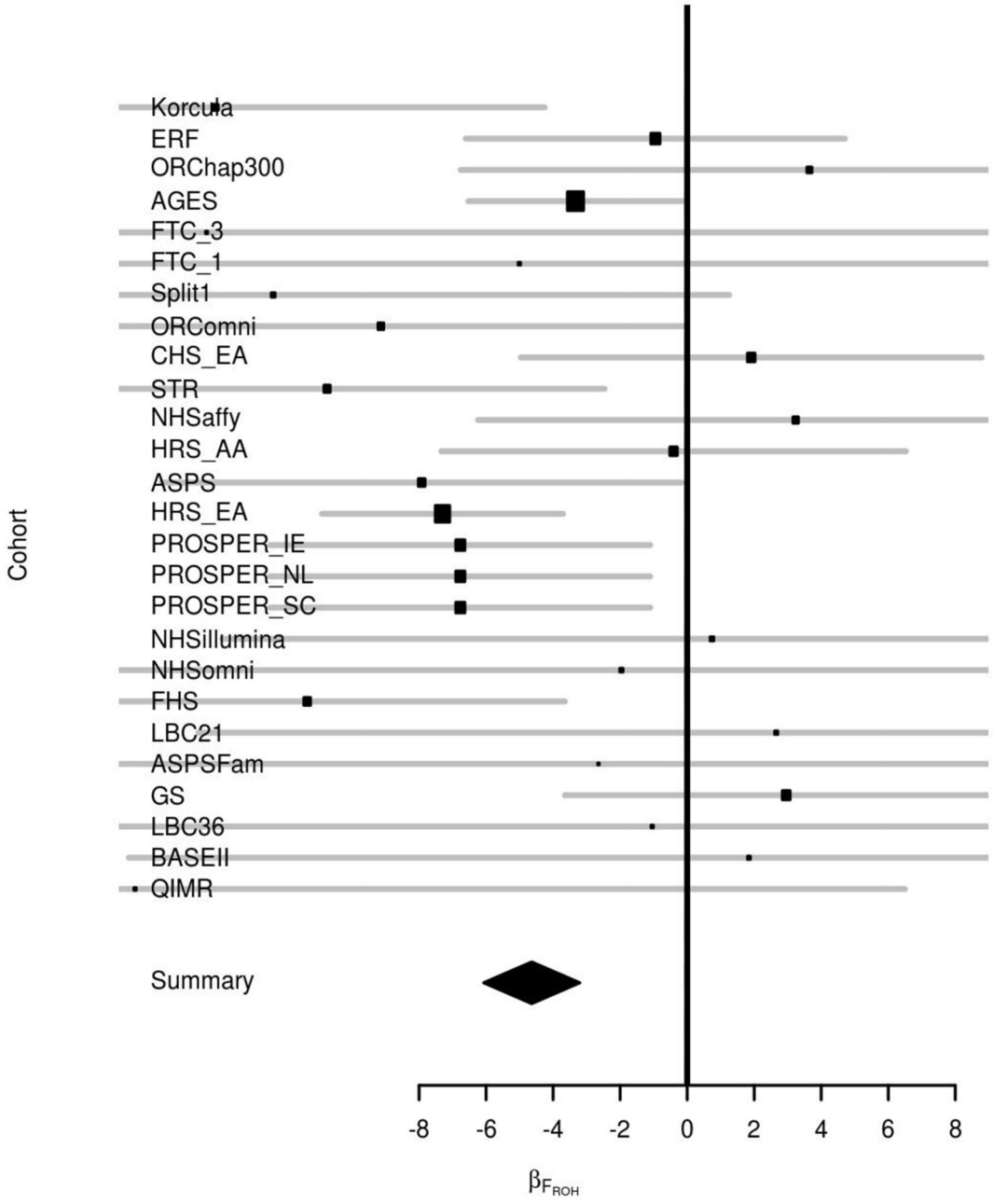
Meta-analysis

Again as is routine in GWAMA, analysis was performed within homogeneous sub-populations and only meta-analysis of the estimated (within population) effect sizes was used to combine results between populations, avoiding any confounding effects of inter-population differences in trait or genetic effect distributions. Inverse-variance meta-analysis of all sub-cohorts' effect estimates was performed using Rmeta, on a fixed effect basis (Supplementary Table 5 compares random effects meta-analysis). In the principal analyses, for cohorts with relatedness, HGLM estimates of β_{FROH} were preferred, however where HGLM had failed to converge, results using GRAMMAR+ were included. These results were combined with those for unrelated cohorts on a fixed model only basis. Result outliers were defined as individual cohort by trait results, which failed the hypothesis, cohort (β_{FROH}) = pre-QC meta-analysis (β_{FROH}), with a t-test statistic >3. Analyses were performed with and without outliers for β_{FROH} in phenotypic units and in intra-sex phenotypic standard deviations (Supplementary Table 8). The principal results we present are for F_{ROH} with outliers included for the hypothesis tests (which turns out to be more conservative), but with outliers excluded when estimating β_{FROH} (ref. 44). Meta-analysis was performed using inverse variance meta-analysis in the R package Rmeta, with β_{FROH} taken as a fixed effect and alternatively as a random effect. The principal results are on a fixed effects basis, with Supplementary Table 5 showing comparison with the random effects analysis.

Meta-analyses were rerun for various subsets, according to geographic and demographic features of the cohorts. Cohorts were divided into more homozygous and less homozygous strata with the boundary being set so each within-stratum meta-analysis had equal statistical power.

Europe PMC Funders Author Manuscripts

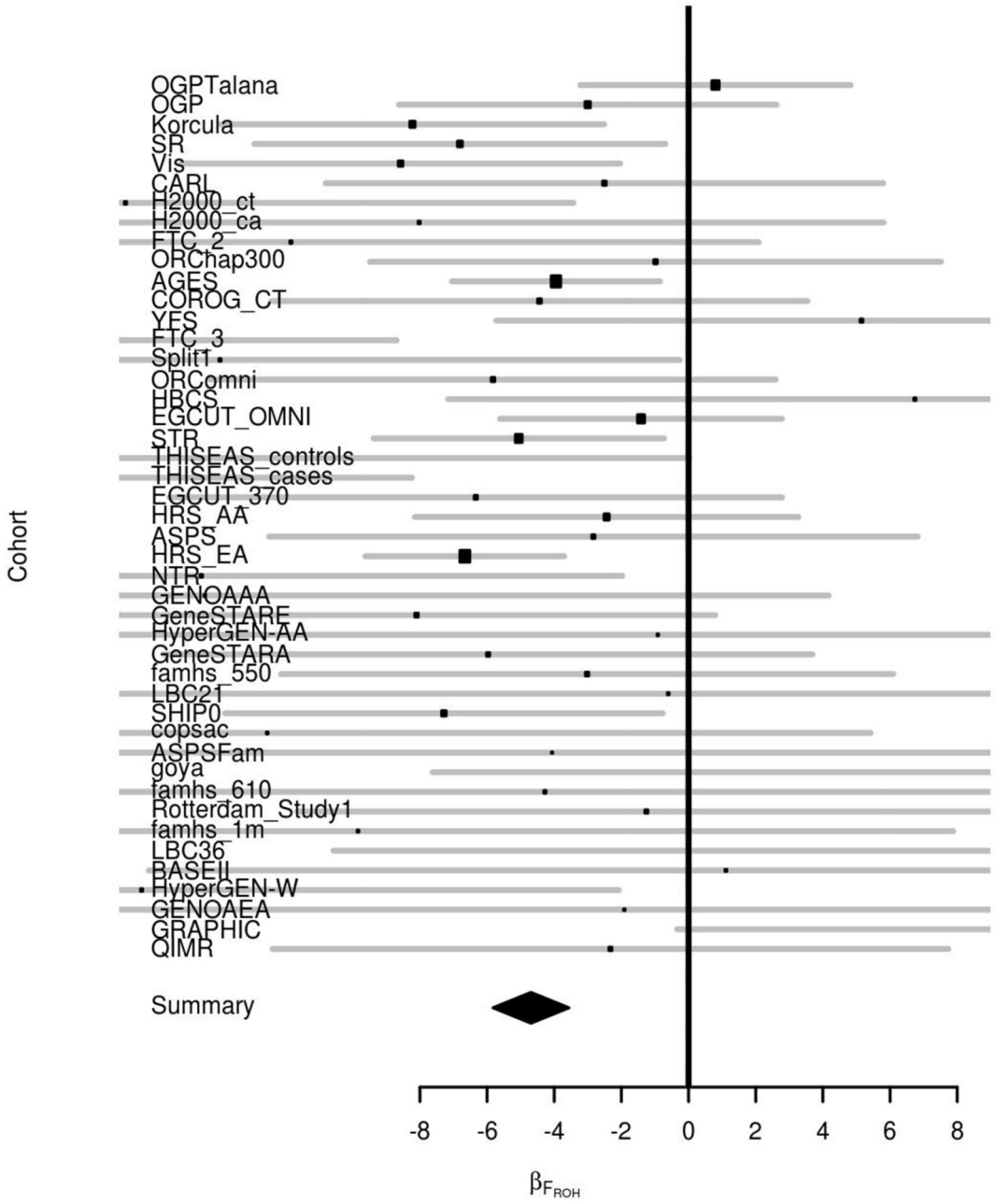
Extended Data



Extended Data Figure 1. Forest plot for cognitive *g*

Individual sub-cohort estimates of effect size and the standard error are plotted. Sub-cohorts are ordered from top to bottom according to their weight in the meta-analysis, so larger or more homozygous cohorts appear towards the top. The scale of beta F_{ROH} is in intra-sex standard deviations. The meta-analytical estimate is displayed at the bottom. Sub-cohort names follow the conventions detailed in Supplementary Table 6 and the Supplementary

Table 11 legend. Sample sizes, effect sizes and P values for association are given in Table 1. This trait was rank transformed.



Extended Data Figure 2. Forest plot for educational attainment

Individual sub-cohort estimates of effect size and the standard error are plotted. Subcohorts are ordered from top to bottom according to their weight in the meta-analysis, so larger or more homozygous cohorts appear towards the top. The scale of beta F_{ROH} is in intra-sex standard deviations. The meta-analytical estimate is displayed at the bottom. Sub-cohort

Europe PMC Funders Author Manuscripts

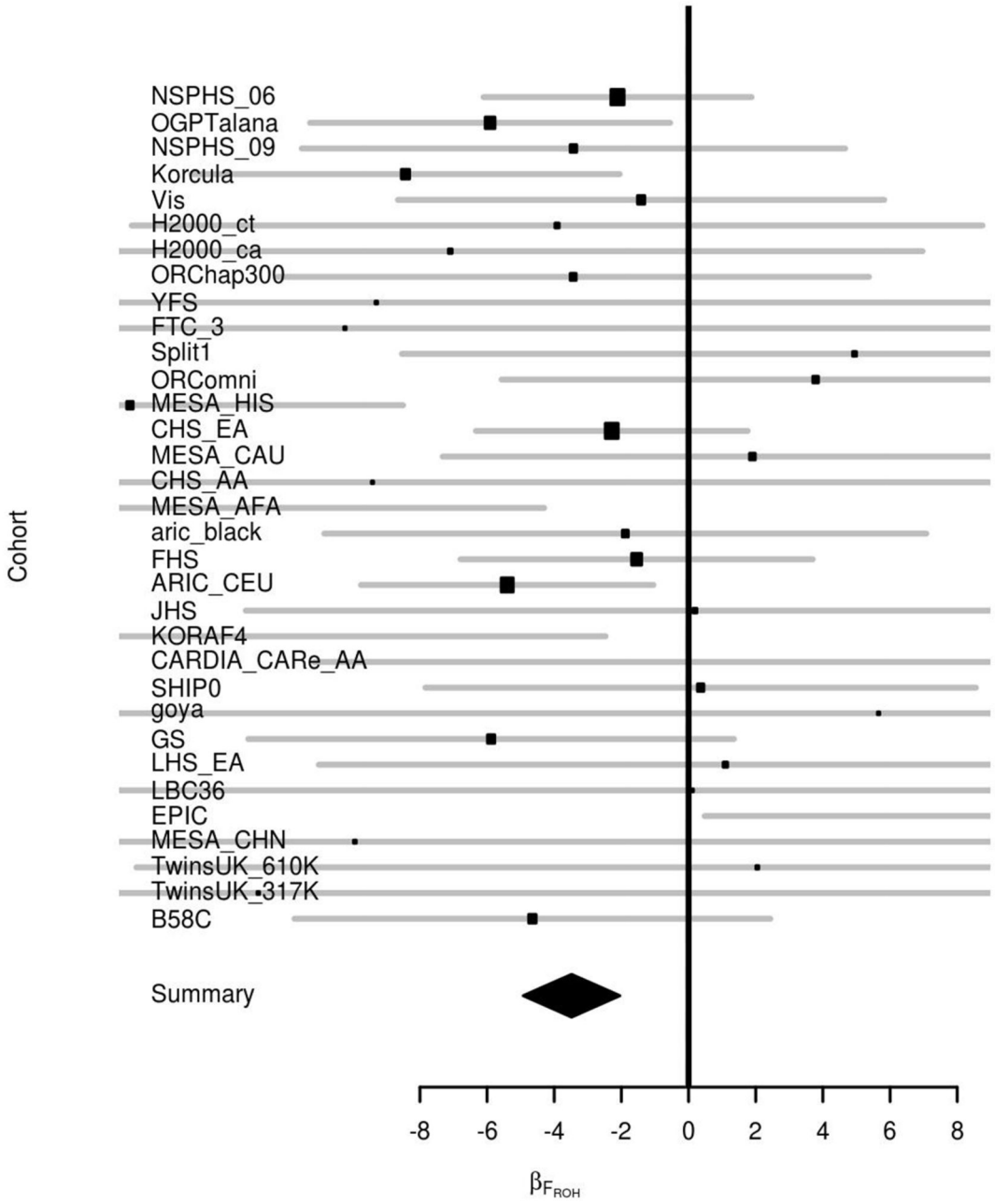
names follow the conventions detailed in Supplementary Table 6 and the Supplementary Table 11 legend. Sample sizes, effect sizes and P values for association are given in Table 1.

Extended Data Figure 3. Forest plot for height

Individual sub-cohort estimates of effect size and the standard error are plotted. Subcohorts are ordered from top to bottom according to their weight in the meta-analysis, so larger or more homozygous cohorts appear towards the top. The scale of beta F_{ROH} is in intra-sex standard deviations. The meta-analytical estimate is displayed at the bottom. Sub-cohort

Europe PMC Funders Author Manuscripts

names follow the conventions detailed in Supplementary Table 6 and the Supplementary Table 11 legend. Sample sizes, effect sizes and P values for association are given in Table 1.

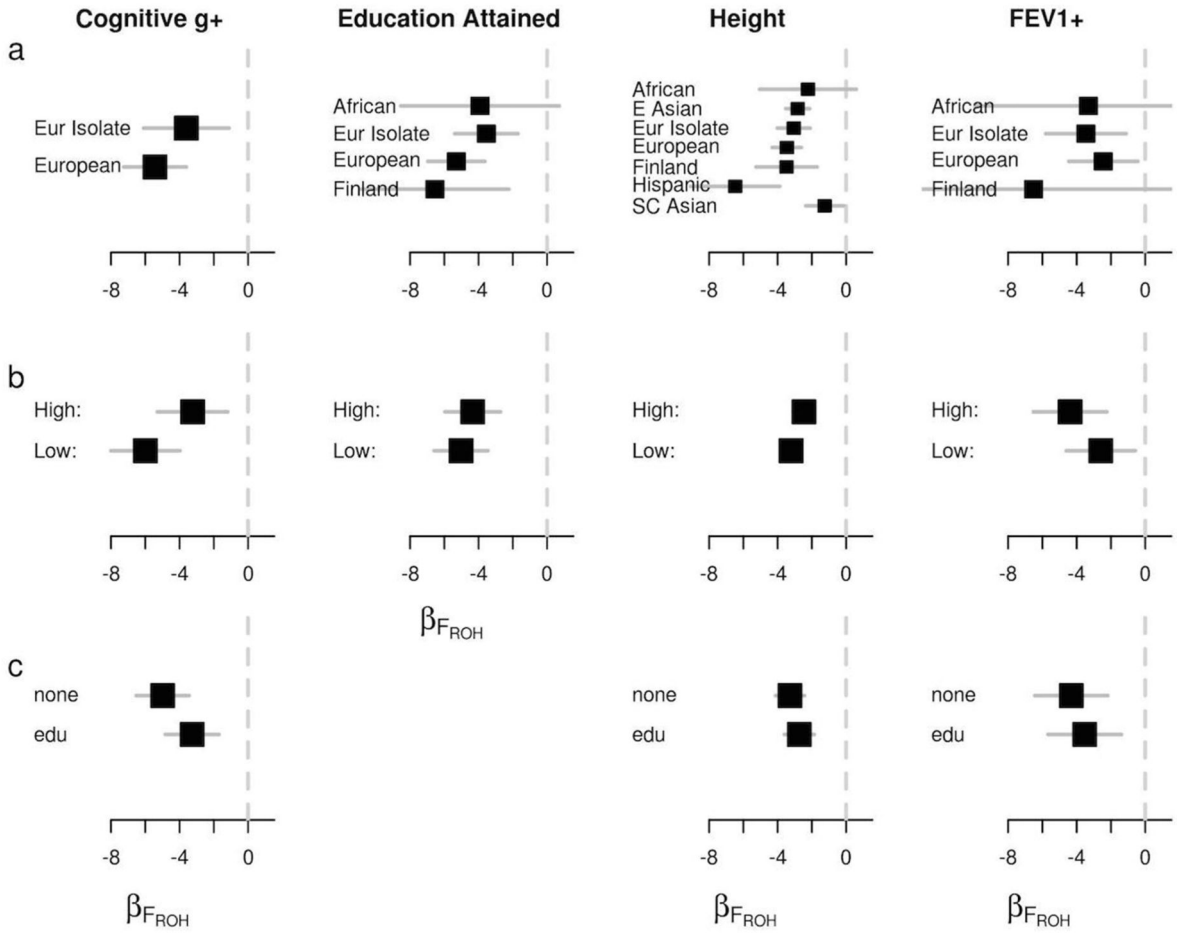


Extended Data Figure 4. Forest plot for forced expiratory lung volume in one second

Individual sub-cohort estimates of effect size and the standard error are plotted. Subcohorts are ordered from top to bottom according to their weight in the meta-analysis, so larger or more homozygous cohorts appear towards the top. The scale of beta F_{ROH} is in intra-sex standard deviations. The meta-analytical estimate is displayed at the bottom. Sub-cohort

Europe PMC Funders Author Manuscripts

names follow the conventions detailed in Supplementary Table 6 and the Supplementary Table 11 legend. Sample sizes, effect sizes and P values for association are given in Table 1. This trait was rank transformed.



Extended Data Figure 5. Signals of directional dominance are robust to stratification by geography or demographic history or inclusion of educational attainment as covariate

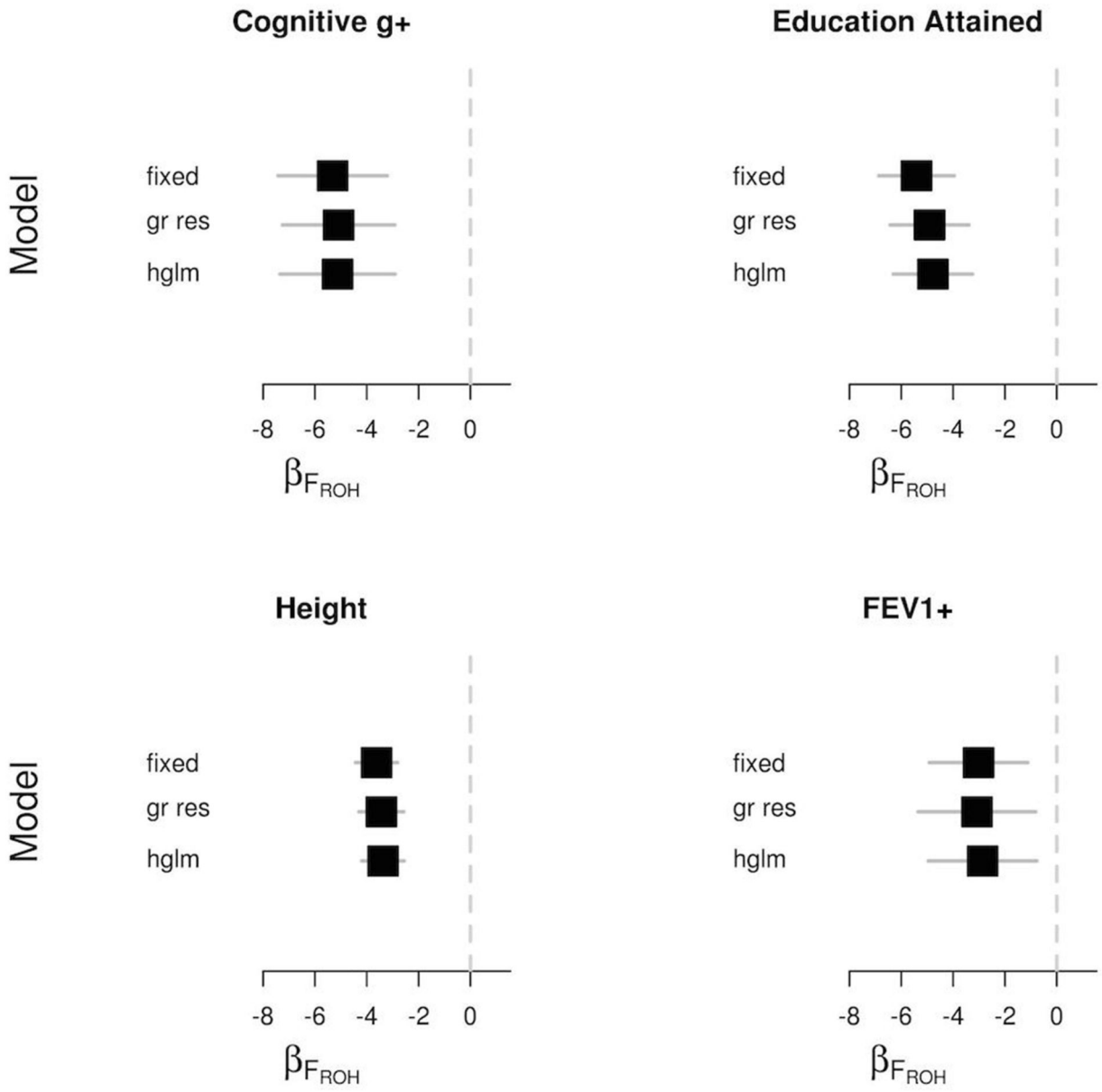
(**a**) Cohorts are divided by continental biogeographic ancestry (African (15 sub-cohorts), East Asian (5), South & Central Asian (10), Hispanic (3)), with Europeans being divided into Finns (13), other European isolates (self-declared, 23), and (non-isolated) Europeans (90). Meta-analysis was carried out for all subsets with 2000 or more samples available. Sample numbers are as follows: cognitive *g*, Eur isolate 6638, European 44,153; educational attainment, African 4811, Eur isolate 8032, European 55,549, Finland 9068; height, African 21,500, E Asian 30,011, Eur isolate 23,116, European 228,813, Finland 30,427, Hispanic 5469, SC Asian 13,523; FEV1, African 6604, Eur isolate 4837, European 49,223, Finland 2340. β_{FROH} is consistent across geography and in both isolates and more cosmopolitan populations. (**b**) Cohorts were divided into High and Low ROH strata of equal power and meta-analysis repeated – the effects are consistent across strata for all four traits. The mean SROH for the high and low strata are 13.4 and 4.3 Mb for cognitive *g*; 28.1 and 5.1 Mb for

Europe PMC Funders Author Manuscripts

education attained; 31.9 and 10.8 Mb for height; and 41.4 and 4.5 Mb for FEV1. (**c**) To assess the potential for socio-economic confounding, where available, educational attainment was included in the regression model (edu) and compared to a model without educational attainment (none) in the same subset of cohorts. The signals reduce slightly when the education covariate is included; the analysis is not possible for educational attainment as a trait. For cognitive *g*, numbers are 36847 and 36023 for edu and none; for height 131,614 and 120,945; and for FEV1, 15717 and 15425. The numbers differ because of missing individual educational data within cohorts. + indicates phenotype was rank transformed. FEV1, forced expiratory lung volume in one second; *g* is the general cognitive component (first unrotated principal component of test scores across diverse tests of cognition); SC Asian is South & Central Asian, E Asian is East Asian, trait units are intra-sex standard deviations and the genomic measure is unpruned SROH.

Europe PMC Funders Author Manuscripts

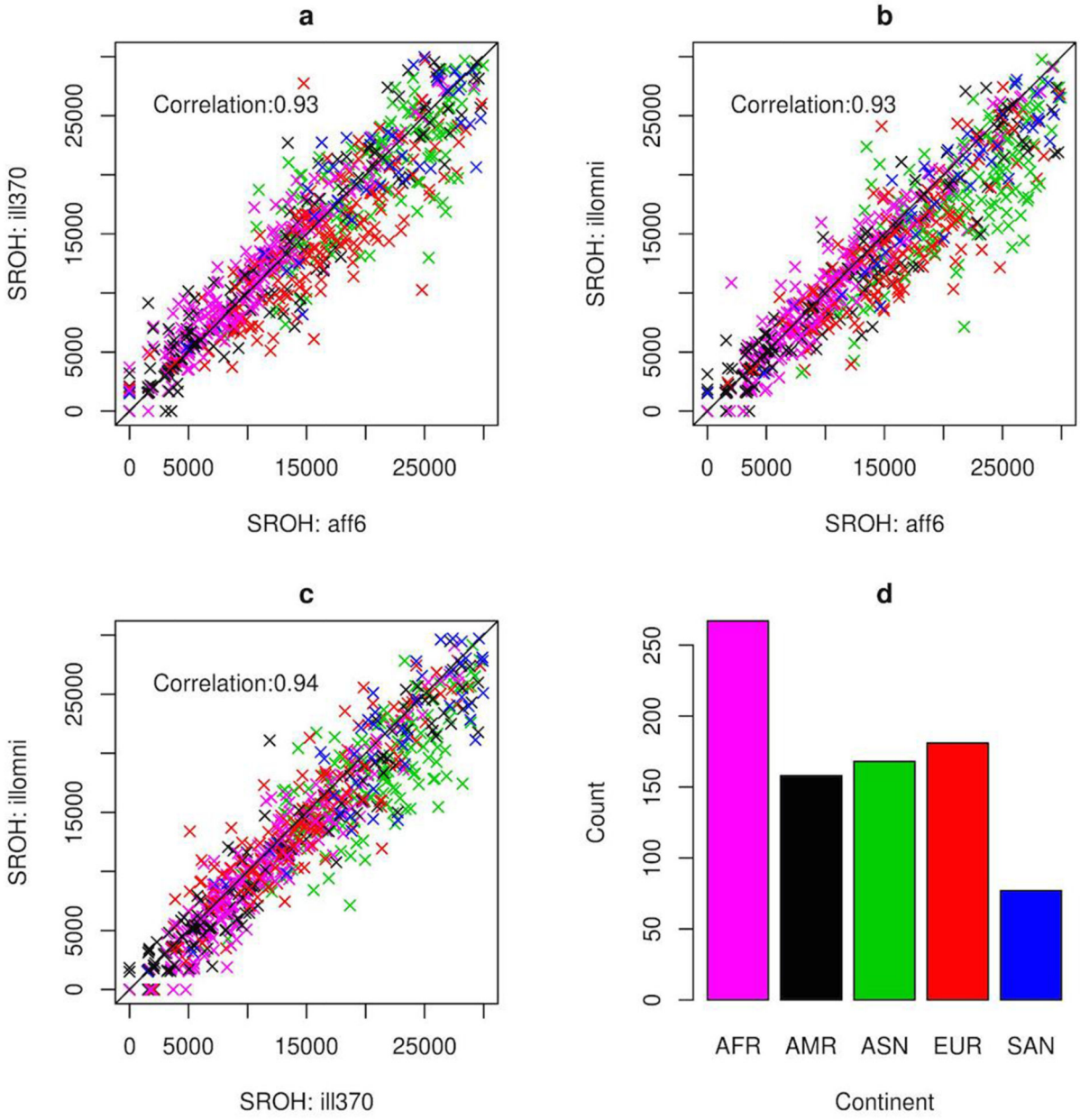
Europe PMC Funders Author Manuscripts



Extended Data Figure 6. Signals of directional dominance are robust to model choice

Meta-analytical estimates of effect size and standard errors are plotted for various models. Fixed indicates no mixed modelling was used, gr res indicates the GRAMMAR+ residuals were fitted and hglm indicates the full hierarchical generalised linear mixed model was used. + indicates the phenotype was rank transformed; FEV1 is forced expiratory lung volume in one second; Cognitive *g* is the general cognitive factor. 15,355 subjects were used for cognitive *g*, 36,060 for educational attainment, 89,112 for height and 15,262 for FEV1.

Europe PMC Funders Author Manuscripts



Extended Data Figure 7. Correlation in SROH for different genotyping arrays using HapMap populations

In panels (a) – (c), X and Y axes show SROH (sum of runs of homozygosity) from 0-30 Mb (30,000 kb). ill370: Illumina CNV370, aff6: Affymetrix6, illomni: Illumina OmniExpress. The graphs are shown for the specific plink call parameters used. (d) Sample numbers per continent are presented in a bar chart. AFR: African, AMR: Mixed American, ASN: East Asian, EUR: European, SAN: South Asian. Only samples with SROH below 30 Mb are plotted, to be conservative to the effect of outliers, which have very strongly correlated

estimates of SROH (r = 0.96-0.97 for comparisons including such very homozygous individuals). In these plots, the correlation between SROH called by the two arrays, r = 0.93-0.94.

Extended data Table 1

Continental ancestry of cohorts participating in each trait study

The first number in each cell is the number of participants with that continental ancestry. The second number is the number of sub-cohorts. BP is blood pressure; FEV1 is forced expiratory lung volume in one second; FVC is forced vital lung capacity; FP is fasting plasma; HbA1c is haemoglobin A1c; HDL/LDL are High/low-density lipoprotein**;** *g* is the general cognitive factor (first unrotated principal component of test scores across diverse domains of cognition). S/C Asian is South & Central Asian.

	African	East Asian	European	Hispanic	S/C Asian	All
BMI	21689/15	29009/5	279400/117	7836/3	13464/10	351398/150
Cognitive *g*	1539/1	NA/NA	49559/22	-	-	51098/23
Diastolic BP	17074/12	24200/5	204742/85	7284/3	12876/9	266176/114
Education Attained	4811/4	NA/NA	79576/42	-	338/1	84725/47
Fasting Insulin	6895/8	1603/1	72006/49	-	6303/5	86807/63
FEV1	6604/5	617/1	58089/27	825/1	-	66135/34
FEVl/FVC	6565/5	616/1	57888/27	822/1	-	65891/34
FP Glucose	8942/9	1615/1	122368/74	1938/1	6921/5	141784/90
HbAlc	6629/4	694/1	92732/31	4038/2	7509/4	111602/42
HDL Cholesterol	15099/13	10478/5	215621/92	4426/3	12508/9	258132/122
Height	20300/14	30011/5	281369/114	5469/2	13523/10	350672/145
LDL Cholesterol	13375/11	2503/2	172245/77	4340/3	11186/8	203649/101
Systolic BP	17023/12	24424/5	205253/85	7225/3	12859/9	266784/114
Total Cholesterol	15130/13	20187/5	209421/91	4491/3	11674/8	260903/120
Triglycerides	13886/12	2542/2	181526/84	2745/2	10688/7	211387/107
Waist-hip ratio	8182/7	2549/2	171753/73	1446/1	12598/9	196528/92

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Peter K. Joshi[#1], Tonu Esko[#2,3,4,5], Hannele Mattsson[6,7], Niina Eklund[6], Ilaria Gandin[8], Teresa Nutile[9], Anne U. Jackson[10], Claudia Schurmann[11,12], Albert V. Smith[13,14], Weihua Zhang[15,16], Yukinori Okada[17,18], Alena Stančáková[19], Jessica D. Faul[20], Wei Zhao[21], Traci M. Bartz[22], Maria Pina Concas[23], Nora Franceschini[24], Stefan Enroth[25], Veronique Vitart[26], Stella Trompet[27], Xiuqing Guo[28,29], Daniel I. Chasman[30], Jeffery R. O'Connel[31], Tanguy Corre[32,33], Suraj S.

Nongmaithem[34], Yuning Chen[35], Massimo Mangino[36,37], Daniela Ruggiero[9], Michela Traglia[38], Aliki-Eleni Farmaki[39], Tim Kacprowski[40], Andrew Bjonnes[41], Ashley van der Spek[42], Ying Wu[43], Anil K. Giri[44], Lisa R. Yanek[45], Lihua Wang[46], Edith Hofer[47,48], Cornelius A. Rietveld[49], Olga McLeod[50], Marilyn C. Cornelis[51,52], Cristian Pattaro[53], Niek Verweij[54], Clemens Baumbach[55,56,57], Abdel Abdellaoui[58], Helen R. Warren[59,60], Dragana Vuckovic[8], Hao Mei[61], Claude Bouchard[62], John R.B. Perry[63], Stefania Cappellani[64], Saira S. Mirza[42], Miles C. Benton[65], Ulrich Broeckel[66], Sarah E. Medland[67], Penelope A. Lind[67], Giovanni Malerba[68], Alexander Drong[69], Loic Yengo[70], Lawrence F. Bielak[21], Degui Zhi[71], Peter J. van der Most[72], Daniel Shriner[73], Reedik Mägi[2], Gibran Hemani[74], Tugce Karaderi[69], Zhaoming Wang[75,76], Tian Liu[77,78], Ilja Demuth[79,80], Jing Hua Zhao[63], Weihua Meng[81], Lazaros Lataniotis[82], Sander W. van der Laan[83], Jonathan P. Bradfield[84], Andrew R. Wood[85], Amelie Bonnefond[70], Tarunveer S. Ahluwalia[86,87,88], Leanne M. Hall[89], Erika Salvi[90], Seyhan Yazar[91], Lisbeth Carstensen[92], Hugoline G. de Haan[93], Mark Abney[94], Uzma Afzal[15,16], Matthew A. Allison[95], Najaf Amin[42], Folkert W. Asselbergs[96,97,98], Stephan J.L. Bakker[99], R. Graham Barr[100], Sebastian E. Baumeister[101], Daniel J. Benjamin[102,103], Sven Bergmann[32,33], Eric Boerwinkle[104], Erwin P. Bottinger[11], Archie Campbell[105], Aravinda Chakravarti[106], Yingleong Chan[3,4,5], Stephen J. Chanock[75], Constance Chen[107], Y.-D. Ida Chen[28,29], Francis S. Collins[108], John Connell[109], Adolfo Correa[61], L. Adrienne Cupples[35,110], George Davey Smith[74], Gail Davies[111,112], Marcus Dörr[113], Georg Ehret[106,114], Stephen B. Ellis[11], Bjarke Feenstra[92], Mary F. Feitosa[46], Ian Ford[115], Caroline S. Fox[110,116], Timothy M. Frayling[85], Nele Friedrich[117], Frank Geller[92], Generation Scotland[105], Irina Gillham-Nasenya[36], Omri Gottesman[11], Misa Graff[118], Francine Grodstein[52], Charles Gu[119], Chris Haley[26,120], Christopher J. Hammond[36], Sarah E. Harris[105,112], Tamara B. Harris[121], Nicholas D. Hastie[26], Nancy L. Heard-Costa[110,122], Kauko Heikkilä[123], Lynne J. Hocking[124], Georg Homuth[40], Jouke-Jan Hottenga[58], Jinyan Huang[125], Jennifer E. Huffman[26], Pirro G. Hysi[36], M. Arfan Ikram[42,126], Erik Ingelsson[69,127], Anni Joensuu[6,7], Åsa Johansson[25,128], Pekka Jousilahti[129], J. Wouter Jukema[130], Mika Kähönen[131], Yoichiro Kamatani[18], Stavroula Kanoni[82], Shona M. Kerr[26], Nazir M. Khan[44], Philipp Koellinger[49], Heikki A. Koistinen[132,133,134], Manraj K. Kooner[16], Michiaki Kubo[135], Johanna Kuusisto[136], Jari Lahti[137,138], Lenore J. Launer[121], Rodney A. Lea[65], Benjamin Lehne[15], Terho Lehtimäki[139], David C.M. Liewald[112], Lars Lind[140], Marie Loh[15], Marja-Liisa Lokki[141], Stephanie J. London[142], Stephanie J. Loomis[143], Anu Loukola[123], Yingchang Lu[11,12], Thomas Lumley[144], Annamari Lundqvist[145], Satu Männistö[129], Pedro Marques-Vidal[146], Corrado Masciullo[38], Angela Matchan[147], Rasika A. Mathias[45,148], Koichi Matsuda[149], James B. Meigs[150], Christa Meisinger[56], Thomas Meitinger[151,152], Cristina Menni[36], Frank D. Mentch[84], Evelin Mihailov[2], Lili Milani[2], May E. Montasser[31], Grant W. Montgomery[153], Alanna Morrison[104], Richard H. Myers[154], Rajiv Nadukuru[11], Pau Navarro[26], Mari Nelis[2], Markku S. Nieminen[155], Ilja M. Nolte[72], George T. O'Connor[110,156], Adesola Ogunniyi[157], Sandosh Padmanabhan[158], Walter R. Palmas[100], James S. Pankow[159], Inga Patarcic[160], Francesca Pavani[53], Patricia A. Peyser[21], Kirsi Pietilainen[7,133,161], Neil Poulter[162], Inga Prokopenko[163], Sarju Ralhan[164], Paul

Redmond[111], Stephen S. Rich[165], Harri Rissanen[145], Antonietta Robino[64], Lynda M. Rose[30], Richard Rose[166], Cinzia Sala[38], Babatunde Salako[157], Veikko Salomaa[129], Antti-Pekka Sarin[6,7], Richa Saxena[41], Helena Schmidt[167], Laura J. Scott[10], William R. Scott[15,16], Bengt Sennblad[50,168], Sudha Seshadri[110,122], Peter Sever[162], Smeeta Shrestha[34], Blair H. Smith[169], Jennifer A. Smith[21], Nicole Soranzo[147], Nona Sotoodehnia[170], Lorraine Southam[69,147], Alice V. Stanton[171], Maria G. Stathopoulou[172], Konstantin Strauch[57,173], Rona J. Strawbridge[50], Matthew J. Suderman[74], Nikhil Tandon[174], Sian-Tsun Tang[175], Kent D. Taylor[28,29], Bamidele O. Tayo[176], Anna Maria Töglhofer[167], Maciej Tomaszewski[89,177], Natalia Tšernikova[2,178], Jaakko Tuomilehto[132,179,180], Andre G. Uitterlinden[42,181], Dhananjay Vaidya[45,182], Astrid van Hylckama Vlieg[93], Jessica van Setten[83], Tuula Vasankari[183], Sailaja Vedantam[3,4,5], Efthymia Vlachopoulou[141], Diego Vozzi[64], Eero Vuoksimaa[123], Melanie Waldenberger[55,56], Erin B. Ware[21], William Wentworth-Shields[94], John B. Whitfield[184], Sarah Wild[1], Gonneke Willemsen[58], Chittaranjan S. Yajnik[185], Jie Yao[28], Gianluigi Zaza[186], Xiaofeng Zhu[187], The BioBank Japan Project[18], Rany M. Salem[3,4,5], Mads Melbye[92,188], Hans Bisgaard[86,87], Nilesh J. Samani[89,177], Daniele Cusi[90], David A. Mackey[91], Richard S. Cooper[176], Philippe Froguel[70,163], Gerard Pasterkamp[83], Struan F.A. Grant[84,189], Hakon Hakonarson[84,189], Luigi Ferrucci[190], Robert A. Scott[63], Andrew D. Morris[191], Colin N.A. Palmer[192], George Dedoussis[39], Panos Deloukas[82,193], Lars Bertram[78,194], Ulman Lindenberger[77], Sonja I. Berndt[75], Cecilia M. Lindgren[4,69], Nicholas J. Timpson[74], Anke Tönjes[195], Patricia B. Munroe[59,60], Thorkild I.A. Sørensen[88,196], Charles N. Rotimi[73], Donna K. Arnett[197], Albertine J. Oldehinkel[198], Sharon L.R. Kardia[21], Beverley Balkau[199], Giovanni Gambaro[200], Andrew P. Morris[2,69,201], Johan G. Eriksson[129,202,203,204,205], Margie J. Wright[206], Nicholas G. Martin[184], Steven C. Hunt[207], John M. Starr[112,208], Ian J. Deary[111,112], Lyn R. Griffiths[65], Henning Tiemeier[42,209], Nicola Pirastu[8,64], Jaakko Kaprio[7,123,210], Nicholas J. Wareham[63], Louis Pérusse[211], James G. Wilson[212], Giorgia Girotto[8], Mark J. Caulfield[59,60], Olli Raitakari[213,214], Dorret I. Boomsma[58], Christian Gieger[55,56,57], Pim van der Harst[54,97,215], Andrew A. Hicks[53], Peter Kraft[107], Juha Sinisalo[155], Paul Knekt[145], Magnus Johannesson[216], Patrik K.E. Magnusson[217], Anders Hamsten[50], Reinhold Schmidt[47], Ingrid B. Borecki[218], Erkki Vartiainen[129], Diane M. Becker[45,219], Dwaipayan Bharadwaj[44], Karen L. Mohlke[43], Michael Boehnke[10], Cornelia M. van Duijn[42], Dharambir K. Sanghera[220,221], Alexander Teumer[101], Eleftheria Zeggini[147], Andres Metspalu[2,178], Paolo Gasparini[64], Sheila Ulivi[64], Carole Ober[94], Daniela Toniolo[38], Igor Rudan[1], David J. Porteous[105,112], Marina Ciullo[9], Tim D. Spector[36], Caroline Hayward[26], Josée Dupuis[35,110], Ruth J.F. Loos[11,12,222], Alan F. Wright[26], Giriraj R. Chandak[34,223], Peter Vollenweider[146], Alan Shuldiner[31,224,225], Paul M. Ridker[30], Jerome I. Rotter[28,29], Naveed Sattar[226], Ulf Gyllensten[25], Kari E. North[118,227], Mario Pirastu[23], Bruce M. Psaty[228,229], David R. Weir[20], Markku Laakso[136], Vilmundur Gudnason[13,14], Atsushi Takahashi[18], John C. Chambers[15,16,230], Jaspal S. Kooner[16,175,230], David P. Strachan[231], Harry Campbell[1], Joel N. Hirschhorn[3,4,5], Markus Perola[2,6], Ozren Polašek[#1,160], and James F. Wilson[#1,26] **for ROHgen**

Affiliations

[1]Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Teviot Place, Edinburgh, EH8 9AG, Scotland [2]Estonian Genome Center, University of Tartu, Riia 23b, 51010, Tartu, Estonia. [3]Division of Endocrinology and Center for Basic and Translational Obesity Research, Boston Children's Hospital, Cambridge, 02141, MA, USA [4]Program in Medical and Population Genetics, Broad Institute, Cambridge Center 7, Cambridge, 02242, MA, USA [5]Department of Genetics, Harvard Medical School, 25 Shattuck St, Boston, 02115, MA, USA [6]Unit of Public Health Genomics, National Institute for Health and Welfare, P.O. Box 104, Helsinki, FI-00251, Finland [7]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, P.O. Box 20, Helsinki, FI-00014, Finland [8]Department of Medical Sciences, University of Trieste, Strada di Fiume 447 - Osp. di Cattinara, Trieste, 34149, Italy [9]Institute of Genetics and Biophysics "A. Buzzati-Traverso" CNR, via Pietro Castellino, 111, Naples, 80131, Italy [10]Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, 48109, MI, USA [11]The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, 10029, USA [12]The Genetics of Obesity and Related Metabolic Traits Program, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, 10029, USA [13]Icelandic Heart Association, Holtasmari 1, 201, Kopavogur, Iceland [14]Faculty of Medicine, University of Iceland, Reykjavik, 101, Iceland [15]Department of Epidemiology and Biostatistics, Imperial College London, Norfolk Place, London, W2 1PG, UK [16]Department of Cardiology, Ealing Hospital NHS Trust, Uxbridge Road, Southall, Middlesex, UB1 3HW, UK [17]Department of Human Genetics and Disease Diversity, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyoku, Tokyo, 113-8510, Japan [18]Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan [19]Department of Medicine, University of Eastern Finland, 70210 Kuopio, Finland [20]Institute for Social Research, University of Michigan, 426 Thompson Street, 48104, Ann Arbor, MI, USA [21]Department of Epidemiology, University of Michigan, 1415 Washington Heights, 48109, Ann Arbor, MI, USA [22]Cardiovascular Health Research Unit, Departments of Biostatistics and Medicine, University of Washington, 1730 Minor Ave, Suite 1360, Seattle, 98101, WA, USA [23]Institute of Population Genetics, National Research Council, Trav. La Crucca n. 3 – Reg. Baldinca, Sassari, 07100, Italy [24]Epidemiology, University of North Carolina, 137 E. Franklin St., Suite 306, 27599-8050, Chapel Hill, USA [25]Immunology, Genetics & Pathology, Uppsala University, Husargatan 3, Box 815, Uppsala, SE-751 08, Sweden [26]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Crewe Road, EH4 2XU, Edinburgh, UK [27]Department of Gerontology and Geriatrics, Leiden University Medical Center , PO Box 9600, Leiden, Netherlands [28]Institute for Translational Genomics and Population Sciences , Los Angeles Biomedical Research Institute, 1124 W. Carson Street, Torrance, 90502, USA [29]Department of Pediatrics, Harbor-UCLA Medical

Center, Torrance, 90502, USA [30]Division of Preventive Medicine, Brigham and Women's Hospital, 900 Commonwealth Avenue, East, Harvard Medical School, Boston, Boston, MA 02215, USA [31]Division of Endocrinology, Diabetes, and Nutrition and Program for Personalised and Genomic Medicine, Department of Medicine, University of Maryland School of Medicine, 685 Baltimore St. MSTF, Baltimore, 21201, USA [32]Department of Medical Genetics, University of Lausanne, Rue du Bugnon 27, Lausanne, 1005, Switzerland [33]Swiss Institute of Bioinformatics, Quartier Sorge - batiment génopode, Lausanne, 1015, Switzerland [34]Genomic Research on Complex Diseases (GRC) Group, CSIR-Centre for Cellular and Molecular Biology, Habshiguda, Uppal Road, Hyderabad, 500007, India [35]Department of Biostatistics, Boston University School of Public Health, 801 Massachusetts Avenue, Boston, 02118, MA, USA [36]Department of Twin Research & Genetic Epidemiology, King's College London, South Wing, Block D, 3rd Floor, Westminster Bridge Road, London, SE1 7EH, UK [37]NIHR Biomedical Research Centre , Guy's and St. Thomas' Foundation Trust, Westminster Bridge Road, London, SE1 7EH, UK [38]Division of Genetics and Cell Biology, San Raffaele Scientific Institute, Via Olgettina 58, Milano, 20132, Italy [39]Department of Nutrition and Dietetics, Harokopio University of Athens, 70, El. Venizelou Ave, Athens, 17671, Greece [40]Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Friedrich-Ludwig-Jahn-Str. 15A, Greifswald, 17475, Germany. [41]Center for Human Genetic Research , 55 Fruit Street, Massachusetts General Hospital, 2114, USA [42]Department of Epidemiology, Erasmus Medical Center, PO Box 2040, Rotterdam, 3000 CA, The Netherlands [43]Department of Genetics, University of North Carolina, Chapel Hill, 27599, NC, USA [44]Genomics and Molecular Medicine, CSIR-Institute of Genomics & Integrative Biology, Mathura Road, New Delhi, 110025, India [45]The GeneSTAR Research Program, Division of General Internal Medicine, Department of Medicine, The Johns Hopkins University School of Medicine, Baltimore, 21287, Maryland, USA [46]Department of Genetics, Washington University School of Medicine, 4444 Forest Park Boulevard, Saint Louis, 63108, MO, USA [47]Department of Neurology, Clinical Division of Neurogeriatrics, Medical University Graz, Auenbruggerplatz 22, Graz, A-8036, Austria [48]Institute for Medical Informatics, Statistics and Documentation, Medical University Graz, Auenbruggerplatz2, Graz, A-8036, Austria [49]Erasmus School of Economics, Erasmus University Rotterdam, Burgemeester Oudlaan 50, Rotterdam, 3000 DR, The Netherlands [50]Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska Institutet, CMM L8:03, Karolinska University Hospital, Solna, Stockholm, 171 76, Sweden [51]Channing Division of Network Medicine, Brigham & Women's Hospital, 181 Longwood, Boston, 02115, USA [52]Nutrition, Harvard School of Public Health, 401 Park Drive, Boston, 02215, USA [53]Center for Biomedicine, European Academy Bozen/Bolzano (EURAC), Bolzano, Italy - Affiliated Institute of the University of Lübeck, Lübeck, Germany [54]University of Groningen, University Medical Center Groningen, Department of Cardiology, Hanzeplein 1, Groningen, 9700RB, The Netherlands [55]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for

Environmental Health, Ingolstädter Landstr. 1, Neuherberg, 85764, Germany [56]Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstr. 1, Neuherberg, 85764, Germany [57]Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstr. 1, Neuherberg, 85764, Germany [58]Department of Biological Psychology, VU University Amsterdam, Van der Boechorststraat 1, 1081 BT, Amsterdam, Netherlands [59]Clinical Pharmacology, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Charterhouse Square, London, EC1M 6BQ, UK [60]NIHR Barts Cardiovascular Biomedical Research Unit, Queen Mary University of London, Charterhouse Square, London, EC1M 6BQ, UK [61]Department of Medicine, University of Mississippi Medical Center, 2500 N. State St., Jackson, 39216, MS, USA [62]Pennington Biomedical Research Center, 6400 Perkins Rd, Baton Rouge, LA 70808, USA [63]MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge Biomedical Campus, Cambridge, CB2 0QQ, UK [64]Institute for Maternal and Child Health - IRCCS “Burlo Garofolo”, via dell’Istria 65, Trieste, 34137, Italy [65]Institute of Health and Biomedical Innovation, Queensland University of Technology, 60 Musk Avenue, Kelvin Grove, GPO Box 2434, Brisbane Qld 4001, Brisbane, Australia [66]Department of Pediatrics, Medical College of Wisconsin, 8701 Watertown Plank Rd, Milwaukee, 53226, WI, USA [67]Quantitative Genetics, QIMR Berghofer Medical Research Institute, 300 Herston Rd, Herston, Brisbane, 4006, Australia [68]Dipartimento di Scienze della Vita e della Riproduzione, University of Verona, Strada Le Grazie 15, Verona, 37134, Italy [69]Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford, OX3 7BN, UK [70]CNRS UMR 8199, European Genomic Institute for Diabetes (EGID), Lille 2 University, 1 Rue du Professeur Calmette, 59000, Lille, France [71]Department of Biostatistics, University of Alabama at Birmingham, 1665 University Blvd, Birmingham, 35294, AL, USA [72]Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, P.O. box 30.001, 9700 RB, Groningen, The Netherlands [73]Center for Research on Genomics and Global Health, National Human Genome Research Institute, Building 12A/Room 4047, 12 South Dr., Bethesda, 20892, Maryland, USA [74]MRC Integrative Epidemiology Unit, University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN, UK [75]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, 9609 Medical Center Drive, Rockville, 20850, MD, USA [76]Cancer Genomics Research Laboratory, National Cancer Institute, SAIC-Frederick, Inc., Frederick National Laboratory for Cancer Research, Frederick,MD, USA [77]Center for Lifespan Psychology, Max Planck Institute for Human Development, Lentzeallee 94, Berlin, 14195, Germany [78]Vertebrate Genomics, Max Planck Institute for Molecular Genetics, Ihnestr. 72, Berlin, 14195, Germany [79]Charité Research Group on Geriatrics, Charité – Universitätsmedizin Berlin, Reinickendorferstr. 61, 13347, Berlin, Germany [80]Institute of Medical and Human Genetics, Charité – Universitätsmedizin Berlin, Augustenburger Platz 1,

Berlin, 13353, Germany [81]Division of Population Health Sciences, Medical Research Institute, University of Dundee, Ninewells hospital and School of Medicine, Dundee, DD2 4BF, Scotland. [82]William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Charterhouse Square, London, EC1M 6BQ, UK [83]Experimental Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Heidelberglaan 100, Utrecht, 3584 CX, The Netherlands [84]Center for Applied Genomics, Children's Hospital of Philadelphia, 3615 Civic Center Boulevard, Philadelphia, PA 19104, USA [85]Genetics of Complex Traits, University of Exeter Medical School, University of Exeter, Royal Devon and Exeter Hospital, Barrack Road, Exeter, EX2 5DW, UK [86]Copenhagen Prospective Studies on Asthma in Childhood, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark [87]The Danish Pediatric Asthma Center, Gentofte Hospital, The Capital Region, Copenhagen, Denmark [88]Novo Nordisk Centre for Basic Metabolic Research, Section of Metabolic Genetics, Faculty of Health and Medical Sciences, University of Copenhagen, Universitetsparken 1, Copenhagen, 2100, Denmark [89]Department of Cardiovascular Sciences, University of Leicester, BHF Cardiovascular Research Centre, Glenfield Hospital, Groby Road, Leicester, LE3 9QP, UK [90]Department of Health Sciences, University of Milan, via A. di Rudinì 8, 20142 Milan, Italy. [91]Centre for Ophthalmology and Visual Science, University of Western Australia, Lions Eye Institute, 2 Verdun St, Perth, 6009, Australia [92]Department of Epidemiology Research, Statens Serum Institut, Artillerivej 5, Copenhagen, 2300, Denmark [93]Clinical Epidemiology, Leiden University Medical Center, PO Box 9600, Leiden, 2300RC, The Netherlands [94]Department of Human Genetics, University of Chicago, 920 E. 58th Street, Chicago, IL, USA [95]Department of Family and Preventive Medicine, University of California San Diego, 9500 Gilman Drive, La Jolla, 92093, USA [96]Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Heidelberglaan 100, Utrecht, 3584 CX, The Netherlands [97]Durrer Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, Catharijnesingel 52, Utrecht, 3501 DG, The Netherlands [98]Institute of Cardiovascular Science, faculty of Population Health Sciences, University College London, Gower Street, London, WC1E 6BT, UK [99]University of Groningen, University Medical Center Groningen, Department of Internal Medicine, Hanzeplein 1, Groningen, 9700RB, The Netherlands [100]Department of Medicine, Columbia University, 622 W. 168th Street, New York, 10032, NY, USA [101]Institute for Community Medicine, University Medicine Greifswald, W.-Rathenau-Str. 48, Greifswald, 17475, Germany [102]Department of Economics, Cornell University, 480 Uris Hall, Ithaca, NY, 14853, USA [103]Department of Economics and Center for Economic and Social Research, University of Southern California, 314C Dauterive Hall, 635 Downey Way, Los Angeles, CA, 90089, USA [104]Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, 1200 Pressler St., Suite 453E, Houston, Texas, 77030, USA [105]Centre for Genomic and Experimental Medicine, University of Edinburgh, Western General Hospital, Edinburgh, EH4 2XU, UK [106]McKusick-Nathans Institute of Genetic Medicine, Johns

Hopkins University School of Medicine, Baltimore, 21205, MD, USA [107]Program in Genetic Epidemiology and Statistical Genetics, Harvard School of Public Health, 665 Huntington Ave, Boston, 02115, USA [108]Genome Technology Branch, National Human Genome Research Institute, NIH, Bethesda, 20892, MD, USA [109]College of Medicine, Dentistry and Nursing, Ninewells Hospital and Medical School, College Office, Level 10, Dundee, DD1 9SY, UK [110]National Heart, Lung, and Blood Institute's Framingham Heart Study, 73 Mt. Wayte Ave, Framingham, 01702, MA, USA [111]Psychology, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, UK [112]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, UK [113]Department of Internal Medicine B, University Medicine Greifswald, Ferdinand-Sauerbruch-Str. NK, Greifswald, 17475, Germany [114]Cardiology, Geneva University Hospitals, Rue Gabrielle-Perret-Gentil, 4, Genève 14, 1211, Switzerland [115]Robertson Centre, University of Glasgow, Boyd Orr Building, Glasgow, G12 8QQ, Scotland. [116]Division of Endocrinology, Brigham and Women's Hospital and Harvard Medical School , 75 Francis St, Boston, 02115, MA, USA [117]Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Ferdinand-Sauerbruch-Str. NK, 17475, Greifswald, Germany. [118]Epidemiology, University of North Carolina, 137 E Franklin St., Suite 306, USA [119]Division of Biostatistics, Washington University, 660 S Euclid, St Louis, 63110, MO, USA [120]Roslin Institute, University of Edinburgh, Easter Bush, Midlothian, EH25 9RG, Scotland [121]National Institutes on Aging, National Institutes of Health, Bethesda, MD, USA [122]Department of Neurology, Boston University School of Medicine, 72 E Concord St, Boston, 02118, MA, USA [123]Department of Public Health, University of Helsinki, Hjelt Institute, P.O.Box 41, Mannerheimintie 172, Helsinki, FI-00014, Finland [124]Musculoskeletal Research Programme, Division of Applied Medicine, University of Aberdeen, Foresterhill, Aberdeen, AB25 2ZD, UK [125]State Key Laboratory of Medical Genomics, Shanghai Institute of Hematology, Rui Jin Hospital Affiliated with Shanghai Jiao Tong University School of Medicine, 197 Rui Jin Er Road, Shanghai, 200025 China [126]Department of Radiology, Erasmus Medical Center, PO Box 2040, Rotterdam, 3000 CA, The Netherlands [127]Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden [128]Uppsala Clinical Research Center, Uppsala University, Uppsala, SE-75237, Sweden. [129]Department of Chronic Disease Prevention, National Institute for Health and Welfare, P.O. Box 30, Helsinki, FI-00271, Finland [130]Department of Cardiology C5-P , Leiden University Medical Center, PO Box 9600, Leiden, Netherlands [131]Department of Clinical Physiology, University of Tampere and Tampere University Hospital, P.O. Box 2000, Tampere, 33521, Finland [132]Diabetes Prevention Unit, National Institute for Health and Welfare, P.O. Box 30, FI-00271 Helsinki, Finland [133]Department of Medicine, Division of Endocrinology, Helsinki University Central Hospital, P.O.Box 340, Haartmaninkatu 4, Helsinki, FI-00029, Finland [134]Minerva Foundation Institute for Medical Research, Biomedicum 2U, Tukholmankatu 8, Helsinki, FI-00290, Finland [135]Laboratory for Genotyping Development RCfIMS, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa, 230-0045, Japan [136]Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, 70210, Finland [137]Institute of Behavioural Sciences, University of Helsinki, P.O. Box 9, FI-00014 University of Helsinki, Helsinki, Finland [138]Folkhälsan Reasearch Centre, PB 63, Helsinki, FI-00014 University of Helsinki, Finland [139]Department of Clinical Chemistry, Fimlab Laboratories and School of Medicine University of Tampere, Tampere, 33520, Finland [140]Department of Medical Sciences, University Hospital, Uppsala, 75185, Sweden. [141]Transplantation laboratory, Haartman Institute, University of Helsinki, P.O. Box 21, Helsinki, FI-00014, Finland [142]National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, RTP, NC, USA [143]Ophthalmology, Massachusetts Eye and Ear, 243 Charles St, Boston, 02114, USA [144]Department of Statistics, University of Auckland, 303.325 Science Centre, Private Bag 92019, Auckland, 1142, New Zealand [145]Department of Health, Functional Capacity and Welfare, National Institute for Health and Welfare, P.O. Box 30, Helsinki, FI-00271, Finland [146]Department of Internal Medicine, University Hospital, Route du Bugnon 44, Lausanne, 1011, Switzerland [147]Human Genetics, Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1HH, UK [148]Division of Allergy and Clinical Immunology, Department of Medicine, The Johns Hopkins University School of Medicine, Baltimore, Maryland, 21224, USA [149]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo, 108-8639, Japan [150]Division of General Internal Medicine, Massachusetts General Hospital , 50 Staniford St, Boston, 02114, MA, USA [151]Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstr. 1, Neuherberg, 85764 Germany [152]Institute of Human Genetics, Klinikum rechts der Isar, Technische Universität München, Ismaninger Str. 22, München, 81675, Germany [153]Molecular Epidemiology, QIMR Berghofer Medical Research Institute, 300 Herston Rd, Herston, Brisbane, 4006, Australia [154]Genome Science Institute, Boston University School of Medicine, 72 East Concord Street, E-304, Boston, 2118, MA, USA [155]HUCH Heart and Lung center, Helsinki University Central Hospital, P.O. Box 340, Helsinki, FI-00029, Finland [156]Pulmonary Center and Department of Medicine, Boston University School of Medicine, 72 E Concord St, Boston, 02118, MA, USA [157]Department of Medicine, University of Ibadan, Ibadan, Nigeria [158]ICAMS, University of Glasgow, 126 University Way, Glasgow, G12 8TA, UK [159]Division of Epidemiology and Community Health , University of Minnesota , 1300 S 2nd Street, Minneapolis, 55454, USA [160]Centre for Global Health and Department of Public Health, School of Medicine, University of Split, Soltanska 2, 21000 Split, Croatia [161]Obesity Research Unit, Research Programs Unit, Diabetes and Obesity, University of Helsinki, P.O.Box 63, Haartmaninkatu 8, FI-00014, Helsinki, Finland [162]International Centre for Circulatory Health, Imperial College London, London, W2 1LA, UK [163]Department of Genomics of Common Disease, School of Public Health, Imperial College London, London, SW7 2AZ, UK [164]Department of Cardiology and Cardio thoracic Surgery Hero DMC Heart Institute,

Civil Lines, 141001, Ludhiana, India [165]Department Public Health Sciences, University of Virginia School of Medicine, 3232 West Complex, Charlottesville, 22908, USA [166]Department of Psychological & Brain Sciences, Indiana University Bloomington, 1101 E. 10th St., Bloomington, IN 47405, USA [167]Institute of Molecular Biology and Biochemistry, Medical University Graz, Harrachgasse 21, Graz, A-8010, Austria [168]Science for Life Laboratory, Karolinska Institutet, Stockholm, Sweden. [169]University of Dundee, Kirsty Semple Way, Dundee, DD2 4DB, UK [170]Cardiovascular Health Research Unit, Division of Cardiology, University of Washington, 1730 Minor Ave, Suite 1360, Seattle, 98101, WA, USA [171]Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, St. Stephen's Green, Dublin 2, Ireland [172]UMR INSERM U1122; IGE-PCV "Interactions Gène-Environnement en Physiopathologie Cardio-Vasculaire", INSERM, University of Lorraine, 30 Rue Lionnois, Nancy, 54000, France [173]Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany. [174]Department of Endocrinology, All India Institute of Medical Sciences, Ansari Nagar East, New Delhi, 110029, India [175]National Heart and Lung Institute, Imperial College London, Du Cane Road, London, W12 0NN, UK [176]Department of Public Health Sciences, Stritch School of Medicine, Loyola University Chicago, Maywood, USA [177]NIHR Leicester Cardiovascular Biomedical Research Unit, University of Leicester, Glenfield Hospital, Groby Road, Leicester, LE3 9QP, UK [178]Institute of Molecular and Cell Biology, University of Tartu, Riia 23, Tartu, 51010 Estonia [179]Centre for Vascular Prevention, Danube-University Krems, 3500 Krems, Austria [180]Diabetes Research Group, King Abdulaziz University, 21589 Jeddah, Saudi Arabia [181]Department of Internal Medicine, Erasmus Medical Center, PO Box 2040, Rotterdam, 3000 CA, The Netherlands [182]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, 21205, USA [183]Finnish Lung Health Association, Sibeliuksenkatu 11 A 1, Helsinki, FI-00250, Finland. [184]Genetic Epidemiology, QIMR Berghofer Medical Research Institute, 300 Herston Rd, Herston, Brisbane, 4006, Australia [185]Diabetes Unit, KEM Hospital and Research Centre, Rasta Peth, Pune, 411011, India. [186]Renal Unit, Department of Medicine, Piazzale A. Stefani 1, Verona, 37124, Italy. [187]Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, USA [188]Department of Medicine, Stanford University, 300 Pasteur Drive, Stanford, 94305, CA, USA [189]Department of Pediatrics, Perelman School of Medicine, The University of Pennsylvania, 3615 Civic Center Boulevard, Philadelphia, PA 19104, USA [190]Translational Gerontology Branch, National institute on Aging, Baltimore, 21225, Maryland, USA [191]Jacqui Wood Cancer Centre, Medical Research Insitute, University of Dundee, Ninewells hospital and School of Medicine, Dundee, DD1 9SY, Scotland. [192]Centre for Pharmacogenetics and Pharmacogenomics, Medical Research Institute, University of Dundee, Ninewells hospital and School of Medicine, Dundee, DD1 9SY, Scotland [193]Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah, 21589, Saudi Arabia. [194]Faculty of Medicine, Imperial College London, Charing Cross Campus - St Dunstan's Road,

London, W6 8RP, UK [195]Department of Medicine, University of Leipzig, Leipzig, Germany. [196]Institute of Preventive Medicine, Bispebjerg and Frederiksberg Hospital , The Capital Region, Copenhagen, 2000, Denmark [197]Department of Epidemiology, University of Alabama at Birmingham, 1665 University Blvd, Birmingham, 35294, AL, USA [198]Department of Psychiatry, University Medical Center Groningen, University of Groningen, P.O. box 30.001, Groningen, 9700 RB, The Netherlands [199]Epidemiology of diabetes, obesity and chronic kidney disease over the lifecourse, Inserm, CESP Center for Research in Epidemiology and Population Health U1018, 16 Avenue Paul Vaillant Couturier, Villejuif, 94807, France [200]Dipartimento di Scienze Mediche, Catholic University of the Sacred Heart, Via G. Moscati 31/34, Roma, 00168, Italy [201]Department of Biostatistics, University of Liverpool, Duncan Building, Daulby Stree, Liverpool, L69 3GA, UK [202]Department of General Practice and Primary Health Care, University of Helsinki, P.O. Box 20, University of Helsinki, Helsinki, FI-00014, Finland [203]Vasa Central Hospital, Sandviksgatan 2-4, Vasa, 65130, Finland [204]Folkhälsan Reasearch Centre, PB 63, University of Helsinki, Helsinki, FI-00014, Finland [205]Unit of General Practice, Helsinki University Central Hospital, Haartmaninkatu 4, Helsinki, FI-00290, Finland [206]Neuro-Imaging Genetics, QIMR Berghofer Medical Research Institute, 300 Herston Rd, Herston, Brisbane, 4006 Australia [207]Cardiovascular Genetics Division, University of Utah, 420 Chipeta Way, Room 1160, Salt Lake City, 84117, Utah, USA [208]Alzheimer Scotland Research Centre, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, UK [209]Department of Psychiatry, Erasmus Medical Center, PO Box 2040, Rotterdam, 3000 CA, The Netherlands [210]National Institute for Health and Welfare (THL), P.O.Box 30, Mannerheimintie 166, Helsinki, FI-00271, Finland [211]Department of kinesiology, Laval University, 2300 rue de la Terrasse, Quebec, G1V 0A6, Canada [212]Department of Physiology and Biophysics, University of Mississippi Medical Center, 2500 N. State St., Jackson, 39216, MS, USA [213]Department of Clinical Physiology and Nuclear Medicine, University of Turku and Turku University Hospital, Turku, 20521, Finland [214]Research Center of Applied and Preventive Cardiovascular medicine, University of Turku, Turku, 20521, Finland [215]University of Groningen, University Medical Center Groningen, Department of Genetics, Hanzeplein 1, Groningen, 9700RB, The Netherlands [216]Department of Economics, Stockholm School of Economics, Box 6501, Stockholm, SE-113 83, Sweden [217]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Box 281, Stockholm, SE-171 77, Sweden [218]Department of Genetics and Biostatistics, Washington University School of Medicine, 4444 Forest Park Boulevard, Saint Louis, 63108, MO, USA [219]Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, 21205, Maryland, USA [220]Department of Pediatrics, University of Oklahoma Health Sciences Center, 940 Stanton Young Boulevard, Oklahoma City, 73104, OK , USA [221]Department of Pharmaceutical Sciences , University of Oklahoma Health Sceienecs Center, Oklahoma City , 73104, USA [222]The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, 10029, USA

[223]Genome Institute of Singapore, 60 Biopolis Street, #02-01 Genome, Singapore, 138672, Singapore [224]Program for Personalised and Genomic Medicine, Department of Medicine, University of Maryland School of Medicine, 685 Baltimore St. MSTF, Baltimore, 21201, USA [225]Geriatric Research and Education Clinical Center, Veterans Administration Medical Center, 685 W Baltimore MSTF, Baltimore, 21201, USA [226]BHF centre, University of Glasgow, 126 University Avenue, Glasgow, G12 8TA, Scotland [227]Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, 137 E. Franklin St., Suite 306, Chapel Hill, USA [228]Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology and Health Services, University of Washington, 1730 Minor Ave, Suite 1360, Seattle, 98101, WA, USA [229]Group Health Research Institute, Group Health Cooperative, 1730 Minor Ave, Suite 1360, Seattle, 98101, WA, USA [230]Imperial College Healthcare NHS Trust, Imperial College London, Praed Street, London, W2 1NY, UK [231]Population Health Research Institute, St George's, University of London, Cranmer Terrace, London, SW17 0RE, UK.

Acknowledgements

We thank the participants in all ROHgen studies; cohort-specific acknowledgements are detailed in Supplementary Table 6. This work was funded by a UK Medical Research Council (MRC) PhD studentship to PKJ, and JFW and OP acknowledge support from the MRC Human Genetics Unit "QTL in Health and Disease" programme. We thank W.G. Hill for discussions and comments on the manuscript and K. Lindsay for administrative assistance.

Author contributions

CHal, PN, MMe, HB, NJS, DC, DAM, RSC, PF, GP, SFG, HH, LF, RAS, ADM, CNP, GDe, PD, LB, UL, SIB, CML, NJT, ATon, PBM, TIS, CNR, DKA, AJO, SLK, BB, GGa, APM, JGE, MJW, NGM, SCH, JMS, IJD, LRG, HT, NPi, JKa, NJW, LP, JGW, GGi, MJC, OR, DDB, CGi, Pv, AAH, PKr, JS, PKn, MJ, PKM, AH, RSc, IBB, EVa, DMB, DB, KLM, MB, CMvD, DKS, ATe, EZ, AMe, PG, SU, CO, DT, GDS, IR, DJP, MC, TDS, CHay, JD, RJL, AFW, GRC, PV, ASh, PMR, JIR, NS, UG, KEN, MP, BMP, DRW, MLa, VG, ATa, JCC, JSK, DPS, HC, JNH, MP, OP, JFW designed individual studies. TN, JDF, SE, VV, STr, DIC, SSN, MMa, DR, AF, LRY, EH, CBo, JRP, SC, UB, GM, TLi, ID, JZ, JPB, ES, SY, MAA, SJB, GRB, EPB, ACa, YChan, SJC, YDIC, FSC, JC, ACo, LCu, GDa, MD, SBE, BF, MFF, IF, CSF, TMF, NFri, FGe, IGi, OG, FGr, CGu, CJH, SEH, NDH, NLH, KH, LJH, GHo, PGH, EI, ÅJ, PJ, JJ, MKa, SK, SMK, NMK, HKK, MKu, JKu, JL, RAL, TLe, DCL, LLi, MLL, ALo, TLu, ALu, SM, KM, JBM, CMei, TM, CMen, FDM, LM, GWM, RHM, RN, MN, MSN, GTO, AO, SP, WRP, JSP, IPa, KP, NPo, SRa, PR, SSR, HR, AR, LMR, RR, BSa, RMS, VS, ASa, LJS, SSe, PS, BHS, NSor, ASttn, MGS, KS, NTa, KDT, BOT, ATog, MTo, JT, AGU, AvHV, TV, SV, EVl, EVu, MW, JBW, SW, GW, CSY, GZ, XZ, MMe, HB, NJS, DC, DAM, RSC, GP, SFG, HH, LF, RAS, GDe, PD, LB, UL, SIB, GDS, NJT, ATon, PBM, TIS, CNR, DKA, AJO, SLK, BB, MKK, GGa, JGE, MJW, NGM, SCH, JMS, IJD, LRG, JKa, NJW, LP, JGW, GGi, MJC, OR, DDB, CGi, Pv, AAH, PKr, JS, PKn, MJ, PKM, AH, RSc, IBB, EVa, DMB, DB, KLM, MB, CMvD, DKS, EZ, AMe, PG, CO, DT, DJP, MC, TDS, CHay, RJL, AFW, GRC, PV, ASh, PMR, JIR, NS, UG, MP, BMP, DRW, MLa, JCC, JSK, DPS, JNH, MP, OP, JFW collected the data. STr, DIC,

Europe PMC Funders Author Manuscripts

MCC, CBo, UB, ID, MA, FWA, SJB, DJB, EB, EPB, ACc, SJC, JC, IF, TMF, CGu, CJH, TBH, NDH, MI, EI, JJ, PKo, MKu, LJL, RAL, LLi, RAM, KM, JBM, GWM, RHM, PAP, KP, SSR, RR, HS, PS, BHS, NSor, NSot, DVa, JBW, CSY, MMe, NJS, DC, DAM, RSC, PF, GP, SFG, HH, LF, GDe, PD, LB, UL, SIB, CML, ATon, PBM, CNR, DKA, AJO, SLK, BB, GGa, APM, MJW, NGM, SCH, JMS, IJD, LRG, JKa, NJW, LP, MJC, DDB, Pv, PKr, MJ, PKM, AH, RSc, IBB, DMB, DB, KLM, MB, CMvD, DKS, EZ, AMe, PG, SU, CO, IR, DJP, MC, TDS, CHay, AFW, GRC, PV, ASh, PMR, JIR, NS, UG, KEN, BMP, DRW, MLa, VG, DPS, HC, OP, JFW contributed to funding. PKJ, TE, HMa, NE, IGa, TN, AUJ, CSc, AVS, WZhan, YO, AStc, JDF, WZhao, TMB, MMC, NFra, SE, VV, STr, XG, DIC, JRO, TC, SSN, YChen, MMa, DR, MTa, AF, TKac, ABj, AvS, YW, AKG, LRY, LW, EH, CAR, OM, MCC, CP, NV, CBa, AAA, HRW, DVu, HMe, JRP, SSMi, MCB, SSMe, PAL, GM, AD, LY, LFB, DZ, PJv, DS, RM, GHe, TKar, ZW, TLi, ID, JZ, WM, LLa, SWvL, JPB, ARW, ABo, TSA, LMH, ES, SY, IMM, LCa, HGdH, MA, UA, NA, FWA, SEB, SB, ACa, YChan, CC, GDa, GE, BF, MFF, FGe, MG, SEH, JJH, JH, JEH, PGH, AJ, YK, SK, RAL, BL, MLo, SJLoo, YL, PM, AMa, CMen, FDM, EM, MEM, AMo, AO, IPa, FP, IPr, LMR, BSa, RMS, RSa, HS, WRS, CSa, CMa, BSe, SSh, SJLon, JAS, LS, RJS, MJS, STa, BOT, ATog, MTo, NTs, JvS, SV, DVo, EBW, WW, JY, GZ, NJS, RAS, ADM, CNP, SIB, NJT, APM, SCH, HT, NPi, LP, Pv, PKr, RSc, IBB, ATe, CO, MC, JD, JIR, NS, KEN, ATa, JCC, JSK, DPS analysed the data. PKJ, TE, HMa, NE, IGa, TN, AUJ, CSc, AVS, MCB, DPS performed beta-testing of scripts. PKJ and TE performed meta-analysis. PKJ, TE, OP and JFW wrote the manuscript. All authors approved the final manuscript.

References

1. Garrod A. The incidence of alkaptonuria: a study of chemical individuality. Lancet. 1902; 11:1616–1620.
2. Darwin, C. The Variation of Animals and Plants Under Domestication. Appleton: 1868.
3. McQuillan R, et al. Runs of Homozygosity in European Populations. Am. J. Hum. Genet. 2008; 83:359–372. [PubMed: 18760389]
4. McQuillan R, et al. Evidence of Inbreeding Depression on Human Height. PLoS Genet. 2012; 8:e1002655. [PubMed: 22829771]
5. Rudan I, et al. Inbreeding and the Genetic Complexity of Human Hypertension. Genetics. 2003; 163:1011–1021. [PubMed: 12663539]
6. Campbell H, et al. Effects of genome-wide heterozygosity on a range of biomedically relevant human quantitative traits. Hum. Mol. Genet. 2007; 16:233–241. [PubMed: 17220173]
7. Charlesworth D, Willis JH. The genetics of inbreeding depression. Nature Rev. Genetics. 2009; 10:783–796. [PubMed: 19834483]
8. Wright, S. Evolution and the Genetics of Populations, Vol. 3: Experimental Results and Evolutionary Deductions. University of Chicago Press; 1977.
9. Wright S. Coefficients of inbreeding and relationships. Am. Nat. 1922; 56:330–339.
10. Broman KW, Weber JL. Long Homozygous Chromosomal Segments in Reference Families from the Centre d'Étude du Polymorphisme Humain. Am. J. Hum. Genet. 1999; 65:1493–1500. [PubMed: 10577902]
11. Gibson J, Morton NE, Collins A. Extended tracts of homozygosity in outbred human populations. Hum. Mol. Genet. 2006; 15:789–795. [PubMed: 16436455]
12. Kirin M, McQuillan R, Franklin CS, Campbell H, McKeigue PM, Wilson JF. Genomic Runs of Homozygosity Record Population History and Consanguinity. PLoS ONE. 2010; 5:e13996. [PubMed: 21085596]

13. Keller MC, Visscher PM, Goddard ME. Quantification of Inbreeding Due to Distant Ancestors and Its Detection Using Dense Single Nucleotide Polymorphism Data. Genetics. 2011; 189:237–249. [PubMed: 21705750]
14. Pemberton TJ, Rosenberg NA. Population-genetic influences on genomic estimates of the inbreeding coefficient: a global perspective. Hum Hered. 2014; 77:37–48. [PubMed: 25060268]
15. Abdellaoui A, et al. Educational Attainment Influences Levels of Homozygosity through Migration and Assortative Mating. PLoS ONE. 2015; 10:e0118935. [PubMed: 25734509]
16. Neel JV, Schull WJ, Yamamoto M, Uchida S, Yanase T, Fujiki N. The effects of parental consanguinity and inbreeding in Hirado, Japan. II. Physical development, tapping rate, blood pressure, intelligence quotient, and school performance. Am. J. Hum. Genet. 1970; 22:263–83. [PubMed: 5444999]
17. Marioni RE, et al. Common genetic variants explain the majority of the correlation between height and intelligence: the generation Scotland study. Behav. Genet. 2014; 44:91–96. [PubMed: 24554214]
18. Wood AR, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. Nature Genet. 2014; 46:1173–86. [PubMed: 25282103]
19. Deary IJ, et al. Genetic contributions to stability and change in intelligence from childhood to old age. Nature. 2012; 482:212–215. [PubMed: 22258510]
20. Morton NE. Effect of inbreeding on IQ and mental retardation. Proc. Natl. Acad. Sci. USA. 1978; 75:3906–3908. [PubMed: 279005]
21. Najmabadi H, et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature. 2011; 478:57–63. [PubMed: 21937992]
22. Nalls MA, et al. Measures of Autozygosity in Decline: Globalization, Urbanization, and Its Implications for Medical Genetics. PLoS Genet. 2009; 5:e1000415. [PubMed: 19282984]
23. Flynn JR. Massive IQ gains in 14 nations: what IQ tests really measure. Psychol. Bull. 1987; 101:171–191.
24. Galton, F. Natural inheritance. MacMillan; London: 1889.
25. Hoffman JI, et al. High-throughput sequencing reveals inbreeding depression in a natural population. Proc. Natl. Acad. Sci. USA. 2014; 111:3775–3780. [PubMed: 24586051]
26. Wright A, Charlesworth B, Rudan I, Carothers A, Campbell H. A polygenic basis for late-onset disease. Trends Genet. 2003; 19:97–106. [PubMed: 12547519]
27. Weiss, L.A, Pan, L.; Abney, M.; Ober, C. The sex-specific genetic architecture of quantitative traits in humans. Nature Genet. 2006; 38:218–222. [PubMed: 16429159]
28. Powell JE, et al. Congruence of Additive and Non-Additive Effects on Gene Expression Estimated from Pedigree and SNP Data. PLoS Genet. 2014; 9:e1003502. [PubMed: 23696747]
29. Keller MC, et al. Runs of Homozygosity Implicate Autozygosity as a Schizophrenia Risk Factor. PLoS Genet. 2012; 8:e1002656. [PubMed: 22511889]
30. Darwin, C. The Effects of Crossing and Self Fertilization in the Vegetable Kingdom. John Murray; 1876.
31. Purcell S. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am. J. Hum. Genet. 2007; 81:559–575. [PubMed: 17701901]
32. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics. 2007; 23:1294–1296. [PubMed: 17384015]
33. Ronnegard L, Shen X, Alam M. hglm: A Package for Fitting Hierarchical Generalized Linear Models. The R Journal. 2010; 2:20–28.
34. Lango Allen H, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature. 2010; 467:832–838. [PubMed: 20881960]
35. Ehret GB, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature. 2011; 478:103–109. [PubMed: 21909115]
36. Scott RA, et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. Nature Genetics. 2012; 44:991–1005. [PubMed: 22885924]

37. Willer CJ, et al. Discovery and refinement of loci associated with lipid levels. Nature Genetics. 2013; 45:1274–1283. [PubMed: 24097068]
38. Soler Artigas M, et al. Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. Nature Genetics. 2011; 43:1082–1090. [PubMed: 21946350]
39. Rietveld CA, et al. GWAS of 126,559 individuals identified genetic variants associated with educational attainment. Science. 2013; 340:1467–1471. [PubMed: 23722424]
40. Stefansson H, et al. CNVs conferring risk of autism or schizophrenia affect cognition in controls. Nature. 2014; 505:361–366. [PubMed: 24352232]
41. 1000 Genomes Project. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491:56–65. [PubMed: 23128226]
42. Aulchenko YS, de Koning DJ, Haley C. Genomewide rapid association using mixed model and regression: a fast and simple method for genome-wide pedigree-based quantitative trait loci association analysis. Genetics. 2007; 177:577–85. [PubMed: 17660554]
43. Marioni RE, et al. Molecular genetic contributions to socioeconomic status and intelligence. Intelligence. 2014; 44:26–32. [PubMed: 24944428]
44. Hedges, LV.; Olkin, I. Statistical Methods for Meta-Analysis. Academic Press; New York: 1985.

Europe PMC Funders Author Manuscripts

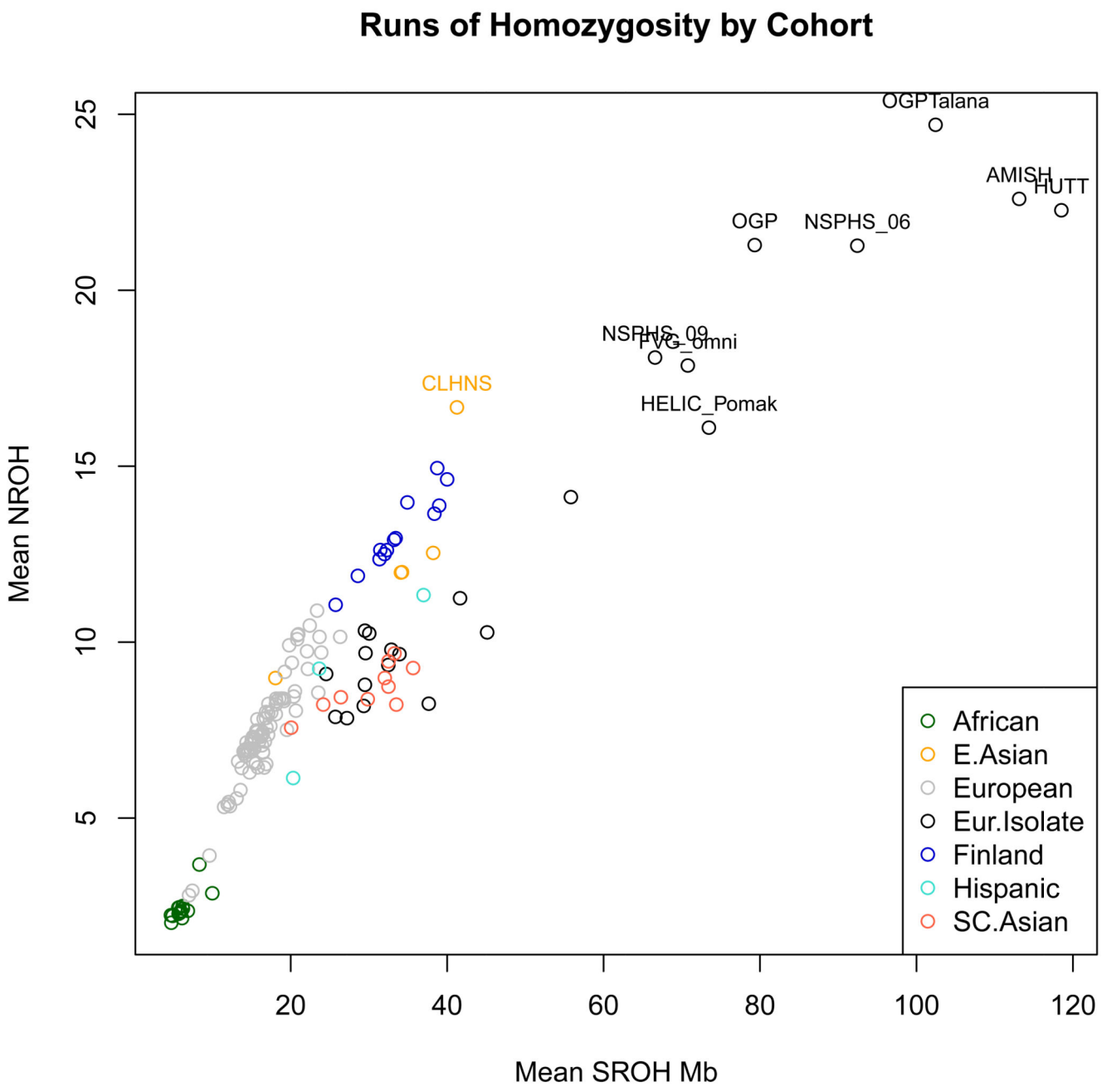


Figure 1. Runs of Homozygosity by Cohort

The sum of runs of homozygosity (SROH) and the number of runs of homozygosity (NROH) are shown by sub-cohort. . Populations differ by an order of magnitude in their mean burden of ROH. There are clear differences by continent and population type both in the mean SROH, and the relationship between SROH and NROH.. SC.Asian is South & Central Asian, E.Asian is East Asian, Eur.Isolate is European isolates. The ten most homozygous cohorts are labelled: AMISH are the Old Order Amish from Lancaster County, Pennsylvania; HUTT, S-Leut Hutterites from South Dakota; NSPHS, North Swedish Population Health Study, 06 and 09 suffixes are different sampling years from different

counties in Northern Sweden; OGP, Ogliastra Genetic Park, Sardinia, Italy; Talana is a particular village in the region; FVG, Friuli-Venezia-Giulia Genetic Park, Italy, omni and 370 suffices refer to subsets genotyped with the Illumina OmniX and 370CNV arrays; HELIC, Hellenic Isolates, Greece, from Pomak villages in Thrace, and CLHNS, Cebu Longitudinal Health and Nutrition Study in the Philippines.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

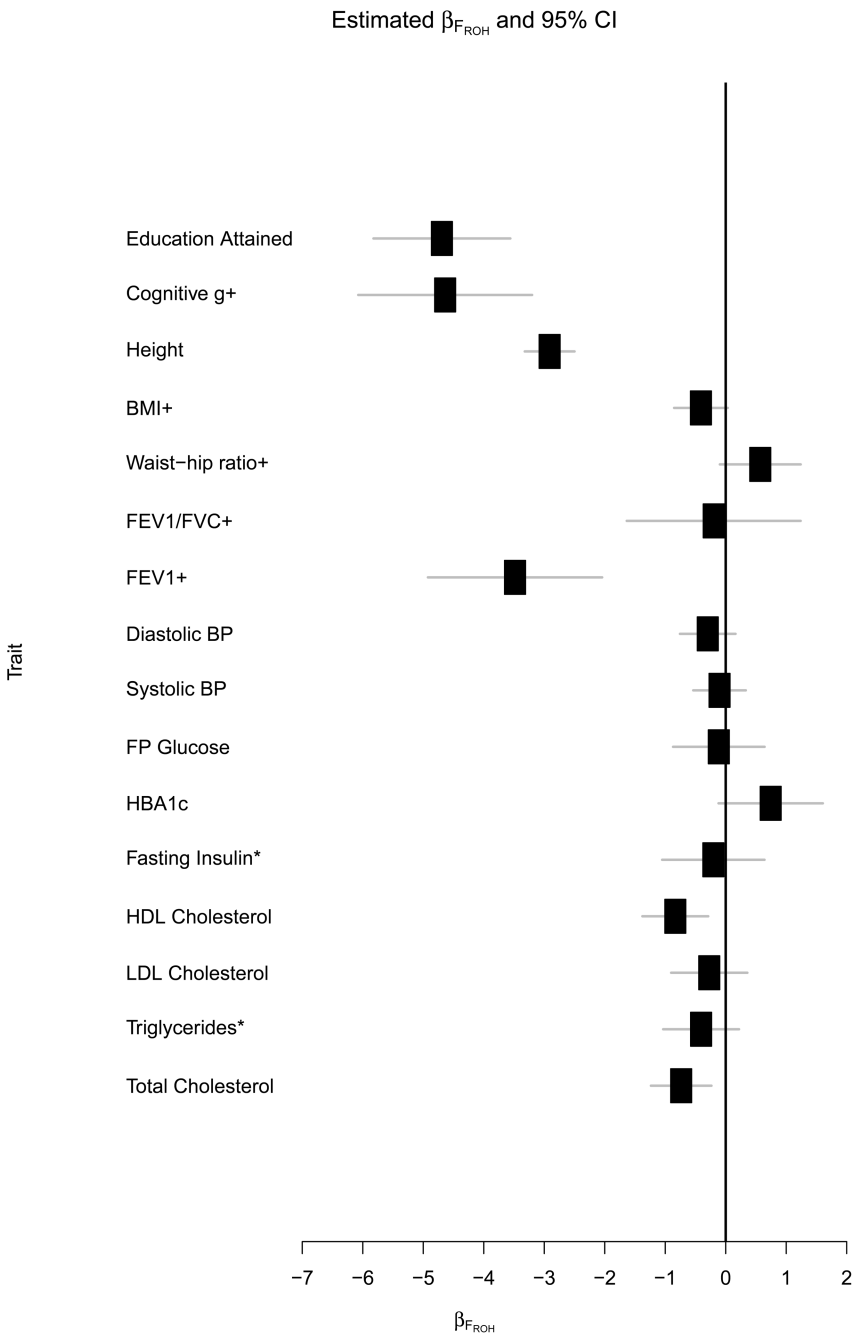


Figure 2. Effects of genome-wide homozygosity, $\beta_{FROH,}$ on 16 traits

Four phenotypes show a significant effect of burden of ROH: height (145 sub-cohorts), FEV1 (34), educational attainment (47) and general cognitive ability, *g* (23). HDL and total cholesterol are not significantly different from zero after correcting for 16 tests and no effect is observed for the other traits. To account for the different numbers of males and females in cohorts and marked effect of sex on some traits, trait units are intra-sex standard deviations. β_{FROH} is the estimated effect of F_{ROH} on the trait, where F_{ROH} is the ratio of the SROH to the total length of the genome. 95% confidence intervals (CIs) are also plotted. + indicates

phenotype was rank transformed, * indicates phenotype was log transformed. BMI, body mass index; BP, blood pressure; FP fasting plasma; HbA1c, haemoglobin A1c (glycated haemoglobin); FEV1, forced expiratory volume in one second; FVC, forced vital capacity; HDL, high density lipoprotein; LDL, low density lipoprotein.

Europe PMC Funders Author Manuscripts

Table 1

Effects of genome-wide burden of runs of homozygosity on four traits

P-association is P value for association, P-heterogeneity is P value for heterogeneity in a meta-analysis between trait and unpruned F_{ROH}, β_{FROH}-SD is the effect size estimate of F_{ROH} expressed in units of intra-sex phenotypic standard deviations and SE is the standard error. β_{FROH}-units is the effect size estimate for F_{ROH} = 1 expressed in the measurement units and SE the standard error. The P values for those traits showing evidence for association are calculated including 5 outlying cohort-specific effect size estimates (an outlier was defined as T-test statistic over 3 for the null hypothesis that the cohort effect size estimate equals the meta-analysis effect size estimate), which is conservative as the majority of these are in the opposite direction. Beta estimates however exclude these outliers, for which there is evidence of discrepancy, and should thus be more accurate. + indicates phenotype was rank transformed; FEV1 is forced expiratory lung volume in one second; g is the general cognitive factor (first unrotated principal component of test scores across diverse domains of cognition).

Phenotype	**Outliers**	**Height**	**FEV1+**	**Educational Attainment**	**Cognitive *g*+**
Subjects		354,224	64,446	84,725	53,300
P-association	Included	$<1 \times 10^{-300}$	2.1×10^{-6}	1.8×10^{-10}	2.5×10^{-10}
P-heterogeneity	Included	0.014	0.10	1.2×10^{-5}	0.071
β_{FROH}-SD	Excluded	−2.91	−3.48	−4.69	−4.64
SE β_{FROH}-SD	Excluded	0.21	0.73	0.58	0.73
β_{FROH}-units	Excluded	−0.188	−2.2	−12.9	−4.64
SE β_{FROH}-units	Excluded	0.014	0.46	1.83	0.73
Units		m	litres	years	SD
First cousin offspring effect	Excluded	−1.2	−137	−9.7	−0.29
Units		cm	ml	months	SD