In many species, the offspring of related parents suffer reduced reproductive success, a phenomenon known as inbreeding depression. In humans, the importance of this effect has remained unclear, partly because reproduction between close relatives is both rare and frequently associated with confounding social factors. Here, using genomic inbreeding coefficients ($F_{ROH}$) for >1.4 million individuals, we show that $F_{ROH}$ is significantly associated ($p < 0.0005$) with apparently deleterious changes in 32 out of 100 traits analysed. These changes are associated with runs of homozygosity (ROH), but not with common variant homozygosity, suggesting that genetic variants associated with inbreeding depression are predominantly rare. The effect on fertility is striking: $F_{ROH}$ equivalent to the offspring of first cousins is associated with a 55% decrease [95% CI 44–66%] in the odds of having children. Finally, the effects of $F_{ROH}$ are confirmed within full-sibling pairs, where the variation in $F_{ROH}$ is independent of all environmental confounding.
Given the pervasive impact of purifying selection on all populations, it is expected that genetic variants with large deleterious effects on evolutionary fitness will be both rare and recessive. However, precisely because they are rare, most of these variants have yet to be identified and their recessive impact on the global burden of disease is poorly understood. This is of particular importance for the nearly one billion people living in populations where consanguineous marriages are common, and the burden of genetic disease is thought to be disproportionately due to increased homozygosity of rare, recessive variants. Although individual recessive variants are difficult to identify, the net directional effect of all recessive variants on phenotypes can be quantified by studying the effect of inbreeding, which gives rise to autozygosity (homozygosity due to inheritance of an allele identical-by-descent).

Levels of autozygosity are low in most of the cohorts with genome-wide data and consequently very large samples are required to study the phenotypic impact of inbreeding. Here, we meta-analyse results from 119 independent cohorts to quantify the effect of inbreeding on 45 commonly measured complex traits of biomedical or evolutionary importance, and supplement these with analysis of 55 more rarely measured traits included in UK Biobank.

Continuous segments of homozygous alleles, or runs of homozygosity (ROH), arise when identical-by-descent haplotypes are inherited down both sides of a family. The fraction of each autosomal genome in ROH > 1.5 Mb (FROH) correlates well with pedigree-based estimates of inbreeding. We estimate FROH using standard methods and software for a total of 1,401,776 individuals in 234 uniform sub-cohorts. The traits measured in each cohort vary according to original study purpose, but together cover a comprehensive range of human phenotypes (Fig. 1). The five most frequently contributed traits (height, weight, body mass index, systolic and diastolic blood pressure) are measured in >1,000,000 individuals; a further 16 traits are measured >500,000 times.

We find that FROH is significantly associated with apparently deleterious changes in 32 out of 100 traits analysed. Increased FROH is associated with reduced reproductive success (decreased number and likelihood of having children, older age at first sex and first birth, decreased number of sexual partners), as well as reduced risk-taking behaviour (alcohol intake, ever-smoked, self-reported risk taking) and increased disease risk (self-reported overall health and risk factors including grip strength and heart rate). We show that the observed effects are predominantly associated with rare (not common) variants and, for a subset of traits, differ between men and women. Finally, we introduce a within-siblings method, which confirms that social confounding of FROH is modest for most traits. We therefore conclude that inbreeding depression influences a broad range of human phenotypes through the action of rare, recessive variants.

Results
Cohort characteristics. As expected, cohorts with different demographic histories varied widely in mean FROH. The within-cohort standard deviation of FROH is strongly correlated with the mean (Pearson’s r = 0.82; Supplementary Fig. 3), and the most homozygous cohorts provide up to 100 times greater per-sample statistical power than cosmopolitan European-ancestry cohorts (Supplementary Data 5). To categorise cohorts, we plotted mean FROH against FIS (Fig. 2). FIS measures inbreeding as reflected by non-random mating in the most recent generation, and is calculated as the mean individual departure from Hardy–Weinberg equilibrium (FNSP; see Methods). Cohorts with high rates of consanguinity lie near the FROH = FIS line, since most excess SNP homozygosity is caused by ROH. In contrast, cohorts with small effective population sizes, such as the Amish and Hutterite isolates of North America, have high average FROH, often despite avoidance of mating with known relatives, since identical-by-descent haplotypes are carried by many couples, due to a restricted number of possible ancestors.

Traits affected by FROH. To estimate the effect of inbreeding on each of the 100 phenotypes studied, trait values were regressed on FROH within each cohort, taking account of covariates including...
autozygosity also decreases the likelihood of having children in the subset of individuals who are, or have been, married, (OR0.0625 = 0.71 ± 0.09, \( p = 3.8 \times 10^{-5} \)) it appears that the cause is a reduced ability or desire to have children, rather than reduced opportunity. Consistent with this interpretation, we observe no significant effect on the likelihood of marriage (OR0.0625 = 0.94 ± 0.07, \( p = 0.12 \)) (Fig. 3b). All effect size, odds ratios and 95% CI are stated as the difference between \( F_{\text{ROH}} = 0 \) and \( F_{\text{ROH}} = 0.0625 \).

The effects on fertility may be partly explained by the effect of \( F_{\text{ROH}} \) on a second group of traits, which capture risky or addictive behaviour. Increased \( F_{\text{ROH}} \) is associated with later age at first sex \( (\beta_{0.0625} = 0.83 ± 0.19 \text{ years, } p = 5.8 \times 10^{-7}) \) and fewer sexual partners \( (\beta_{0.0625} = −1.38 ± 0.38, \ p = 2.0 \times 10^{-12}) \) but also reduced alcohol consumption \( (\beta_{0.0625} = −0.66 ± 0.12 \text{ units per week, } p = 1.3 \times 10^{-22}) \), decreased likelihood of smoking \( (\beta_{0.0625} = 0.79 ± 0.05, \ p = 5.9 \times 10^{-13}) \), and a lower probability of being a self-declared risk-taker \( (\beta_{0.0625} = 0.84 ± 0.06, \ p = 3.4 \times 10^{-5}) \) or exceeding the speed limit on a motorway \( (p = 4.0 \times 10^{-8}) \). Conservatism beliefs are likely to affect these traits, and are known to be confounded with \( F_{\text{ROH}} \) in some populations\(^4\), however, fitting religious participation as a covariate in UKB reduces, but does not eliminate the reported effects (Supplementary Fig. 10b, Supplementary Data 20). Similarly, fitting educational attainment as an additional covariate reduces 16 of 25 significant effect estimates, but actually increases 9, including age at first sex and number of children (Supplementary Fig. 10a, Supplementary Data 20). This is because reduced educational attainment is associated with earlier age at first sex and increased number of children, which makes it an unlikely confounder for the effects of \( F_{\text{ROH}} \), which are in the opposite directions.

A third group of traits relates to cognitive ability. As previously reported, increased autozygosity is associated with decreased general cognitive ability,\(^6\, ^{15}\) and reduced educational attainment.\(^6\) Here, we also observe an increase in reaction time \( (\beta_{0.0625} = 11.6 ± 3.9 \text{ ms, } p = 6.5 \times 10^{-9}) \), a correlate of general cognitive ability (Fig. 3a, Supplementary Data 10).

A fourth group relates to body size. We replicate previously reported decreases in height and forced expiratory volume\(^5\) (Supplementary Data 21) and we find that increased \( F_{\text{ROH}} \) is correlated with a reduction in weight \( (\beta_{0.0625} = 0.86 ± 0.12 \text{ kg, } p = 3.4 \times 10^{-28}) \) and an increase in the waist to hip ratio \( (\beta_{0.0625} = 0.004 ± 0.001, \ p = 1.4 \times 10^{-11}) \).

The remaining effects are loosely related to health and frailty; higher \( F_{\text{ROH}} \) individuals report significantly lower overall health and slower walking pace, have reduced grip strength \( (\beta_{0.0625} = −1.24 ± 0.19 \text{ kg, } p = 6.9 \times 10^{-24}) \), accelerated self-reported facial ageing, and poorer eyesight and hearing. Increased \( F_{\text{ROH}} \) is also associated with faster heart rate \( (\beta_{0.0625} = 0.56 ± 0.24 \text{ bpm, } p = 5.9 \times 10^{-6}) \), lower haemoglobin \( (\beta_{0.0625} = 0.81 ± 0.24 \text{ g/L, } p = 1.6 \times 10^{-11}) \), lymphocyte percentage, and total cholesterol \( (\beta_{0.0625} = −0.05 ± 0.015 \text{ mmol L}^{-1}, \ p = 5.2 \times 10^{-10}) \).

Sex-specific effects of \( F_{\text{ROH}} \). Intriguingly, for a minority of traits (13/100), the effect of \( F_{\text{ROH}} \) differs between men and women (Fig. 3c, Supplementary Data 12). For example, men who are the offspring of first cousins have 0.10 mmol L\(^{-1} \) [95% CI 0.08–0.12] lower total cholesterol on average, while there is no significant effect in women; LDL shows a similar pattern. More generally, for these traits, the effect in men is often of greater magnitude than the effect in women, perhaps reflecting differing relationships between phenotype and fitness.

Associations most likely caused by rare, recessive variants. The use of ROH to estimate inbreeding coefficients is relatively new in inbreeding research\(^1\, ^{16–19} \). Earlier frequency-based estimators...
such as \(F_{SNP}\) and \(F_{GRM}\), made use of excess marker homozygosity\(^{21-23}\) and did not require physical maps. We performed both univariate and multivariate regressions to evaluate the effectiveness of \(F_{ROH}\) against these measures. The correlations between them range from 0.13 to 0.99 and are strongest in cohorts with high average inbreeding (Supplementary Data 6, Supplementary Fig. 6). Significantly, univariate regressions of traits on both \(F_{SNP}\) and \(F_{GRM}\) show attenuated effect estimates relative to \(F_{ROH}\) (Supplementary Data 13). This attenuation is greatest in low autozygosity cohorts, suggesting that \(F_{ROH}\) is a better estimator of excess homozygosity at the causal loci (Fig. 4c).

To explore this further, we first fit bivariate models with \(F_{ROH}\) and \(F_{GRM}\) as explanatory variables. For all 32 traits that were significant in the univariate analysis, we find that \(\hat{\beta}_{F_{ROH}/F_{GRM}}\) is of greater magnitude than \(\hat{\beta}_{F_{GRM}/F_{ROH}}\) in the conditional analysis (Fig. 4b, Supplementary Data 22). This suggests that inbreeding depression is predominantly caused by rare, recessive variants made homozygous in ROH, and not by the chance homozygosity of variants in strong LD with common SNPs (Fig. 4d, Supplementary Note 5). We also find that ROH of different lengths have similar effects per unit length (Fig. 4a, Supplementary Fig. 11a), consistent with their having a causal effect on traits and not with confounding by socioeconomic or other factors, as shorter ROH arise from deep in the pedigree are thus less correlated with recent consanguinity.

**Quantifying the scope of social confounding.** Previous studies have highlighted the potential for \(F_{ROH}\) to be confounded by non-genetic factors\(^{5,24}\). We therefore estimated the effect of \(F_{ROH}\) within various groups, between which confounding might be expected either to differ, or not be present at all.

For example, the effect of \(F_{ROH}\) on height is consistent across seven major continental ancestry groups (Supplementary Fig. 1, Supplementary Data 18), despite differing attitudes towards consanguinity, and consequently different burdens and origins of ROH. Similarly, grouping cohorts into consanguineous, more cosmopolitan, admixed and those with homozygosity due to ancient founder effects also shows consistent effects (Supplementary Fig. 2, Supplementary Data 19). Equally, categorising samples into bins of increasing \(F_{ROH}\) shows a dose-dependent response of the study traits with increased \(F_{ROH}\) (Supplementary Data 17 and
Fig. 4 Inbreeding depression caused by ROH. a Effect of different ROH lengths on height, compared with the effect of SNP homozygosity outside of ROH. The effects of shorter (<5 Mb) and longer (>5 Mb) ROH per unit length are similar and strongly negative, whereas the effect of homozygosity outside ROH is much weaker. The pattern is similar for other traits (Supplementary Fig. 1a; Supplementary Data 14). b \( F_{\text{ROH}} \) is more strongly associated than \( F_{\text{GRM}} \) in a bivariate model of height. Meta-analysed effect estimates, and 95% confidence intervals, are shown for a bivariate model of height (Height \( \sim F_{\text{ROH}} + F_{\text{GRM}} \)). The reduction in height is more strongly associated with \( F_{\text{ROH}} \) than \( F_{\text{GRM}} \), as predicted if the causal variants are in weak LD with the common SNPs used to calculate \( F_{\text{GRM}} \) (Supplementary Note 5). The pattern is similar for other traits (Supplementary Fig. 15a, b; Supplementary Data 22). c \( F_{\text{ROH}} \) is a lower variance estimator of the inbreeding coefficient than \( F_{\text{GRM}} \). The ratio of \( \beta \)_{\text{GRM}} : \beta_{\text{ROH}} \) is plotted against var(\( F_{\text{ROH}} \)/var(\( F_{\text{GRM}} \)) for all traits in all cohorts. When the variation of \( F_{\text{GRM}} \) which is independent of \( F_{\text{ROH}} \) has no effect on traits, \( \beta \)_{\text{GRM}} is downwardly biased by a factor of var(\( F_{\text{ROH}} \)/var(\( F_{\text{GRM}} \)) (Supplementary Note 4). A linear maximum likelihood fit, shown in red, has a gradient consistent with unity (1.01; 95% CI 0.84–1.18), as expected when the difference between \( F_{\text{GRM}} \) and \( F_{\text{ROH}} \) is not informative about the excess homozygosity at causal variants (Supplementary Note 5). d \( F_{\text{ROH}} \) is a better predictor of rare variant homozygosity than \( F_{\text{GRM}} \). The excess homozygosities of SNPs, extracted from UK Biobank imputed genotypes, were calculated at seven discrete minor allele frequencies (\( F_{\text{GRM}} \)), and regressed on two estimators of inbreeding in a bivariate statistical model (see Supplementary Note 5). The homozygosity of common SNPs is better predicted by \( F_{\text{GRM}} \), but rare variant homozygosity is better predicted by \( F_{\text{ROH}} \). The results from real data (Fig. 4b, Supplementary Figs 15a, b and Supplementary Data 22) are consistent with those simulated here, if the causal variants are predominantly rare. All errors bars represent 95% confidence intervals.

Fig. 5a, b show the response for height and ever having children; Supplementary Figs 9a–f for all significant traits). The proportionality of these effects is consistent with a genetic cause, while it is difficult to envisage a confounder proportionally associated across the entire range of observed \( F_{\text{ROH}} \). In particular, the highest \( F_{\text{ROH}} \) group (\( F_{\text{ROH}} > 0.18 \), equivalent to the offspring of first-degree relatives, are found to be, on average, 3.4 [95% CI 2.5–4.3] cm shorter and 3.1 [95% CI 2.5–3.7] times more likely to be childless than an \( F_{\text{ROH}} = 0 \) individual.

Next, we estimated \( \beta_{\text{ROH}} \) for 7153 self-declared adopted individuals in UK Biobank, whose genotype is less likely to be confounded by cultural factors associated with the relatedness of their biological parents. For all 26 significant traits measured in this cohort, effect estimates are directionally consistent with the meta-analysis and 3 (height, walking pace and hearing acuity) reach replication significance (\( p < 0.004 \)). In addition, a meta-analysis of the ratio \( \beta_{\text{ROH,adopted}} : \beta_{\text{ROH}} \) across all traits differs significantly from zero (Fig. 5c; average = 0.78, 95% CI 0.56–1.00, \( p = 2 \times 10^{-12} \)).

Finally, the effect of \( F_{\text{ROH}} \) was estimated in up to 118,773 individuals in sibships (full-sibling pairs, trios, etc.: \( \beta_{\text{ROH,sib}} \)). \( F_{\text{ROH}} \) differences between siblings are caused entirely by Mendelian segregation, and are thus independent of any reasonable model of confounding. The variation of \( F_{\text{ROH}} \) among siblings is a small fraction of the population-wide variation (Supplementary Data 5); nevertheless, 23 out of 29 estimates of \( \beta_{\text{ROH,sib}} \) are directionally consistent with \( \beta_{\text{ROH}} \), and two (self-reported overall health and ever having children) reach replication.
Fig. 5 Evidence ROH effects are un-confounded.  

a Linear decrease in height with increasing $F_{\text{ROH}}$. Average heights (in metres) is plotted in bins of increasing $F_{\text{ROH}}$. The limits of each bin are shown by red dotted lines, and correspond to the offspring of increasing degree unions left-to-right. The overall estimate of $\beta_{\text{ROH}}$ is shown as a solid black line. Subjects with kinship equal to offspring of full-sibling or parent-child unions are significantly shorter than those of avuncular or half-sibling unions who in turn are significantly shorter than those of first-cousin unions. 

b Linear decrease in odds of ever having children with increasing $F_{\text{ROH}}$. Linear model approximations of ln(Odds-Ratio) for ever having children ($1 = $ parous, $0 = $ childless) are plotted in bins of increasing $F_{\text{ROH}}$. A strong relationship is evident, extending beyond the offspring of first cousins. 

c ROH effects are consistent in adoptees. The ratios of effect estimates, $\beta_{\text{ROH}}$ between adoptees and all individuals are presented by trait. All traits are directionally consistent and overall show a strongly significant difference from zero (average = 0.78, 95% CI 0.56–1.00, $p = 2 \times 10^{-12}$). FEV1 forced expiratory volume in one second. 

d ROH effects are consistent in full siblings. The ratios of effect estimates within full siblings to effects in all individuals (\(\beta_{\text{ROH}}^{\text{wSibs}}: \beta_{\text{ROH}}\)) are presented by trait. Twenty-three of 29 estimates are directionally consistent and overall show a significant difference from zero (average = 0.78, 95% CI 0.53–1.04, $p = 7 \times 10^{-10}$). BMI body mass index. All errors bars represent 95% confidence intervals.
significance. A meta-analysis of the ratio $\hat{\beta}_{F_{ROH}\text{-risk}} / \hat{\beta}_{F_{ROH}}$ for all traits is significantly greater than zero (Fig. 5d; average = 0.78, 95% CI 0.53–1.04, p = 7 × 10^{-10}), indicating a substantial fraction of these effects is genetic in origin. However, for both adoptees and siblings, the point estimates are less than one, suggesting that non-genetic factors probably contribute a small, but significant, fraction of the observed effects.

Discussion

Our results reveal inbreeding depression to be broad in scope, influencing both complex traits related to evolutionary fitness and others where the pattern of selection is less clear. While studies of couples show optimal fertility for those with distant kinship25,26, fewer have examined reproductive success as a function of individual inbreeding. Those that did are orders of magnitude smaller in size than the present study, suffer the attendant drawbacks of pedigree analysis, and have found mixed results27–29. Our genomic approach also reveals that in addition to socio-demographic factors and individual choice, recessive genetic effects have a significant influence on whether individuals reproduce. The discordant effects on fertility and education demonstrate that this is not just a result of genetic correlations between the two domains30.

The effects we see on fertility might be partially mediated through a hitherto unknown effect of autozygosity on decreasing the prevalence of risk-taking behaviours. Significant effects of autozygosity are observed for self-reported risk taking, speeding on motorways, alcohol and smoking behaviour, age at first sexual intercourse and number of sexual partners. Independent evidence for a shared genetic architecture between risk-taking and fertility traits comes from analysis of genetic correlations using LD-score regression in UKB (Supplementary Table 1). The core fertility traits, ever had children and number of children, are strongly genetically correlated ($r_{\text{G}} = 0.93; p < 10^{-100}$). Genetic correlations with ever-smoking and self-reported risk-taking are lower, but also significant: 0.23–0.27, p < 10^{-10}. Age at first sex is strongly genetically correlated both with the fertility traits, ($r_{\text{G}} = 0.53–0.57$), and number of sexual partners, ever-smoking and risk-taking30 ($r_{\text{G}} = 0.42–0.60$).

Reproductive traits are understandable targets of natural selection, as might be walking speed, grip strength, overall health, and visual and auditory acuity. While we cannot completely exclude reverse causality, whereby a less risk-taking, more conservative, personality is associated with greater likelihood of consanguinous marriage, we note that the effects are consistent for ROH < 5 Mb, which are less confounded with mate choice, due to their more distant pedigree origins (Supplementary Fig. 11a). This group of traits also shows similar evidence for unconfounded effects in the analysis of adoptees and full siblings (Fig. 5c, d; Supplementary Data 16) and the signals remained after correcting for religious activity or education.

On the other hand, for some traits that we expected to be influenced by ROH, we observed no effect. For example, birth weight is considered a key component of evolutionary fitness in mammals, and is influenced by genomic homozygosity in deer31; however, no material effect is apparent here (Supplementary Data 10). Furthermore, in one case, ROH appear to provide a beneficial effect: increasing $F_{ROH}$ significantly decreases total and LDL-cholesterol in men, and may thus be cardio-protective in this regard.

Our multivariate models show that homozygosity at common SNPs outside of ROH has little influence on traits, and that the effect rather comes from ROH over 1.5 Mb in length. This suggests that genetic variants causing inbreeding depression are almost entirely rare, consistent with the dominance hypothesis1.

The alternative hypothesis of overdominance, whereby positive selection on heterozygotes has brought alleles to intermediate frequencies, would predict that more common homozygous SNPs outside long ROH would also confer an effect. The differential provides evidence in humans that rare recessive mutations underlie the quantitative effects of inbreeding depression.

Previous studies have shown that associations observed between $F_{ROH}$ and traits do not prove a causal relationship14,24. Traditional Genome-wide Association Studies (GWAS) can infer causality because, in the absence of population structure, genetic variants (SNPs) are randomly distributed between, and within, different social groups. However, this assumption does not hold in studies of inbreeding depression, where, even within a genetically homogeneous population, social groups may have differing attitudes towards consanguinity, and therefore different average $F_{ROH}$ and, potentially, different average trait values. We therefore present a number of analyses that discount social confounding as a major factor in our results. Firstly, we show that the effects are consistent across diverse populations, including those were ROH burden is driven by founder effects rather than cultural practices regarding marriage. Effects are also consistent across a 20-fold range of $F_{ROH}$: from low levels, likely unknown to the subject, to extremely high levels only seen in the offspring of first-degree relatives. Secondly, we show that the effects of ROH are consistent in direction and magnitude among adopted individuals, and also for short ROH which are not informative about parental relatedness. Finally, we introduce a within-siblings method, independent of all confounders, that confirms a genetic explanation for most of the observed effects. Variation in $F_{ROH}$ between siblings is caused entirely by random Mendelian segregation; we show that higher $F_{ROH}$ siblings experience poorer overall health and lower reproductive success, as well as other changes consistent with population-wide estimates. Nevertheless, average effect sizes from both adoptees and siblings are 20% smaller than population-wide estimates. Nevertheless, average effect sizes from both adoptees and siblings are 20% smaller than population-wide estimates, confirming the importance of accounting for social confounding in future studies of human inbreeding depression.

Our results reveal five large groups of phenotypes sensitive to inbreeding depression, including some known to be closely linked to evolutionary fitness, but also others where the connection is, with current knowledge, more surprising. The effects are mediated by ROH rather than homozygosity of common SNPs, causally implicating rare recessive variants rather than overdominance as the most important underlying mechanism. Identification of these recessive variants will be challenging, but analysis of regional ROH in particular using whole-genome sequences in large cohorts with sufficient variance in autozygosity will be the first step. Founder populations or those which prefer consanguineous marriage will provide the most power to understand this fundamental phenomenon.

see Supplementary Data.

Methods

Overview. Our initial aim was to estimate the effect of $F_{ROH}$ on 45 quantitative traits and to assess whether any of these effects differed significantly from zero. Previous work21 has shown that inbreeding coefficients are low in most human populations, and that very large samples are required to reliably estimate the genetic effects of inbreeding12. To maximise sample size, a collaborative consortium (ROHgen3) was established, and research groups administering cohorts with SNP chip genotyping were invited to participate. To ensure that all participants performed uniform and repeatable analyses, a semi-automated software pipeline was developed and executed locally by each research group. This software pipeline required cohorts to provide only quality-controlled genotypes (in plink binary format) and standardised phenotypes (in plain-text) and used standard software (R, PLINK12,32, KING33) to perform the analyses described below. Results from each cohort were returned to the central ROHgen analysts for meta-analysis.

During the initial meta-analysis, genotypes were released for >500,000 samples from the richly phenotyped UK Biobank (UKB)10. It was therefore decided to add a
further 34 quantitative traits and 21 binary traits to the ROHgen analysis. Many of these additional traits were unique to UKB, although 7 were also available in a subset of ROHgen cohorts willing to run additional analyses. In total, the effect of $F_{\text{ROH}}$ was tested on 100 traits and therefore experiment-wise significance was defined as $5 \times 10^{-4}$ ($=0.05/100$).

**Cohort recruitment.** In total, 119 independent, genetic epidemiological study cohorts were contributed to ROHgen. Of these, 118 were studies of adults and contributed multiple phenotypes, while 1 was a study of children and contributed only birth weight. To minimise any potential confounding or bias caused by within-cohort heterogeneity, studies were weighted in analyses wherever applicable. Each sub-cohort was required to use only one genotyping array and be of uniform ancestry and case-status. For example, if a study contained multiple distinct ethnicities, sub-cohorts of each ancestry were created and analysed separately. At minimum, ancestry was defined on a sub-continental scale (i.e. European, East Asian, South Asian, West Asian, Japanese, and Hispanic were always analysed separately) but more precise separation was used when deemed necessary, for example, in cohorts with large representation of Ashkenazi Jews. In case-control studies of disease, separate sub-cohorts were created for cases and controls and phenotypes associated with disease status were not analysed in the case cohort for PLINK binding file format were accepted. Each study applied their own GWAS quality controls before additional checks were made in the common

**Ethical approval.** Data from 119 independent genetic epidemiology studies were included. All subjects gave written informed consent for broad-ranging health and genetic research and all studies were approved by the relevant research ethics committees or boards. Published references are given for each study in Supplementary Data 2.

**Genotyping.** All samples were genotyped on high-density (minimum 250,000 markers), genome-wide SNP microarrays supplied by Illumina or Affymetrix. Genotyping arrays with highly variable genomic coverage (such as Exome chip, Metabochip, or Immunochip) were judged unsuitable for the ROH calling algorithm and were not permitted. Imputed genotypes were also not permitted; only called genotypes for PLINK binding file format were accepted. Each study applied their own GWAS quality controls before additional checks were made in the common analysis pipeline: SNPs with >3% missingness or MAF <5% were removed, as were individuals with >3% missing data. Only autosomal genotypes were used for the analyses reported here. Additional, cohort-specific, genotyping information is available in Supplementary Data 2.

**Phenotyping.** In total, results are reported for 79 quantitative traits and 21 binary traits. These traits were chosen to represent different domains of health and reproductive success, with consideration given to presumed data availability. Many of these traits have been the subject of existing genome-wide association meta-analyses (GWAMA), and phenotype modelling, such as inclusion of relevant covariates, was copied from the relevant consortia (GIANT for anthropometry, EGCG for blood weight, ICBP for blood pressures, MAGIC for glycaemic traits, CHARGE-Cognitive, -Inflammation & -Haemostasis working groups for cognitive function, CRP, fibrinogen, CHARGE-CKDGen for eGFR, CHARGE-RepGen for ages at menarche and menopause, Blood Cell & HaemGen for haematology, GUGC for urate, RRgen, PRIMA & QT-IGC for electrocardiography, GLGC for classical lipids, CREAM for spherical equivalent refraction, Spirometry13 for lung function, and SEFAC for educational attainment and number of children ever born). Further information about individual phenotype modelling is available in Supplementary Note 1 and Supplementary Data 8.

**ROH calling.** Runs of homozygosity (ROH) of >1.5 Mb in length were identified using published methods.\(^4,5\) In summary, SNPs with minor allele frequencies below 5% were removed, before continuous ROH SNPs were identified using PLINK with the following parameters: homozygous-window snp 50; homozygous 50; homozygous kb 1500; homozygous gap 1000; homozygous 50; heterozygosity 50; heterozygosity 50; homozygous-low 5; homozygous-window-het 1. No linkage disequilibrium pruning was performed. These parameters have been previously shown to call ROH that correspond to autozygous segments in which all SNPs (including those not present on the chip) are homozygous-by-descent, not chance arrangements of independent homozygous SNPs. The following coefficients estimates calculated for $F_{\text{ROH}}$ correlate well with pedigree-based estimates ($F_{\text{ped}}$).\(^4\) Moreover, they have also been shown to be robust to array choice.\(^8\)

### Calculating estimators of F

For each sample, two estimates of the inbreeding coefficient ($\hat{F}$) were calculated, $F_{\text{ROH}}$ and $F_{\text{SNP}}$. We also calculated three additional measures of homozygosity $F_{\text{ROH<5Mb}}$ ($F_{\text{ROH}}$), $F_{\text{ROH<5Mb,het}}$, and $F_{\text{SNP,meta}}$. $F_{\text{ROH}}$ is the fraction of each genome in ROH >1.5 Mb. For example, in a sample for which PLINK had identified $i$ ROH of length $l_i$ (in Mb), $i \in [1..n]$, then $F_{\text{ROH}}$ was then calculated as

$$F_{\text{ROH}} = \sum_{i=1}^{n} \frac{l_i}{5 \times 10^{-4}}$$

where $F_{\text{ROH<5Mb}}$ and $F_{\text{ROH<5Mb,het}}$ are the genomic fractions in ROH of length >5 Mb, and in ROH of length ≤5 Mb (but >1.5 Mb), respectively, and the length of the autosomal genome is estimated at 3 gigabases (Gb). It follows from this definition that

$$F_{\text{ROH}} = F_{\text{ROH<5Mb}} + F_{\text{ROH<5Mb,het}}$$

Single-point inbreeding coefficients can also be estimated from individual SNP homozygosity without any reference to a genetic map. For comparison with $F_{\text{ROH}}$, a method of moments estimate of inbreeding coefficient was calculated, referred to here as $F_{\text{SNP}}$ and implemented in PLINK by the command--het.

$$F_{\text{SNP}} = \frac{\text{O}(\text{HOM}) - \text{E}(\text{HOM})}{N - \text{E}(\text{HOM})}$$

where $\text{O}(\text{HOM})$ is the observed number of homozygous SNPs, $\text{E}(\text{HOM})$ is the expected number of homozygous SNPs, i.e. $\sum_i (1 - 2 \rho_{pq})$, and $N$ is the total number of non-missing genotyped SNPs. $F_{\text{ROH}}$ and $F_{\text{SNP}}$ are strongly correlated, especially in cohorts with significant inbreeding, since both are estimates of $F$. To clarify the conditional effects of $F_{\text{ROH}}$ and $F_{\text{SNP}}$, an additional measure of homozygosity, $F_{\text{SNP,meta}}$, was calculated to describe the SNP homozygosity observed outside ROH.

$$F_{\text{SNP,meta}} = \frac{\text{O}(\text{HOM}) - \text{E}(\text{HOM})}{N - \text{E}(\text{HOM})}$$

And $N_{\text{SNP,ROH}}$ is the number of homozygous SNPs found in ROH. Note that $F_{\text{SNP,meta}}$ captures $F_{\text{ROH}}$.

A further single point estimator of the inbreeding coefficient, described by Yang et al.\(^20\) as $\hat{F}_{\text{pm}}$, is implemented in PLINK by the parameter --ib (Fast3) and was also calculated for all samples.

$$F_{\text{GRM}} = \hat{F}_{\text{pm}} = \frac{1}{N} \sum_{i=1}^{N} \left[ \frac{x_i^2 - 1 - 2p_i \beta_{\text{xib}}} {2p_i(1-p_i)} \right]$$

where the number of SNPs, $p_i$, is the reference allele frequency of the ith SNP in the sample population and $x_i$ is the number of copies of the reference allele.

### Effect size estimates for quantitative traits

In each cohort of $n$ samples, for each of the quantitative traits measured in that cohort, trait values were modelled by

$$y = \beta_{\text{cons}} + F_{\text{ROH}} + X_b + e$$

where $y$ is a vector ($n \times 1$) of measured trait values, $\beta_{\text{cons}}$ is the unknown scalar effect of $F_{\text{ROH}}$ on the trait, $F_{\text{ROH}}$ is a known vector ($n \times 1$) of individual $F_{\text{ROH}}$, $b$ is a vector ($m \times 1$) of unknown fixed covariate effects (including a mean), $X$ in a known design matrix ($n \times m$) for the fixed effects, and $e$ is an unknown vector ($n \times 1$) of residuals.

The $m$ fixed covariates included in each model were chosen with reference to the leading GWAMA consortium for that trait and are detailed in Supplementary Data 8. For all traits, these covariates included: age (and/or year of birth), sex, and at least the first 10 principal components of the genomic relatedness matrix (GRM). Where necessary, additional adjustments were made for study site, medications, and other relevant covariates (Supplementary Data 3).

For reasons of computational efficiency, it was decided to solve Eq. (10) in two steps. In the first step, the trait ($y$) was regressed on all fixed covariates to obtain the maximum likelihood solution of the model:

$$y = X_b + e'$$

All subsequent analyses were performed using the vector of trait residuals $e'$, which may be considered as the trait values corrected for all known covariates. In cohorts with a high degree of relatedness, mixed-modelling was used to correct for family structure, although, because ROH are not narrow-sense heritable, this was considered less essential than in Genome-Wide Association Studies. Equation (11) becomes

$$y = X_b + u + e'$$

\(^8\) Nature Communications | https://doi.org/10.1038/s41467-019-12283-6 | www.nature.com/naturecommunications
Adoptee $wSib$ was estimated by linear distribution of mean 0 and covariance matrix $\Sigma$. In these related cohorts, a GRM was calculated using PLINK v1.9 and Grammar+ residuals of Eq. (12) were estimated using GenABEL. These Grammar+ residuals ($\epsilon$) were used in subsequent analyses.

To estimate $\beta_{FROH}$ for each trait, trait residuals were regressed on $F_{ROH}$ to obtain the maximum likelihood (ML) solution of the model

$$\epsilon' = \mu + \beta_{FSNPs} \cdot F_{SNPs} + \epsilon.$$  \hspace{1cm} (13a)

The sex-specific estimates of $\beta_{FROH}$ (Supplementary Data 12) were obtained from Eq. (13) applied to the relevant sex.

For all traits, a corresponding estimates of $\beta_{FSNPs}$ and $\beta_{FROH}$ were obtained from the models

$$\epsilon' = \mu + \beta_{FSNPs} \cdot F_{SNPs} + \epsilon,$$  \hspace{1cm} (13b)

and the effects of different ROH lengths and of SNP homozygosity (Fig. 4b) were obtained from the model

$$\epsilon' = \mu + \beta_{FSNPs} \cdot F_{SNPs,assort} + \epsilon,$$  \hspace{1cm} (14)

Effect size estimates for binary traits. Binary traits were analysed by two methods. The primary estimates of $\beta_{FROH}$ (Fig. 3b and Supplementary Data 10) were obtained from full logistic models:

$$g(\hat{E}(y)) = \beta_0 + \beta_1 \cdot x.$$  \hspace{1cm} (16)

where $g$ is the link function (logit), and where $F_{ROH}$ and all applicable covariates (Supplementary Data 3, 8) were fitted simultaneously. Mixed modelling for family structure was not attempted in the logistic models since an accepted method was not apparent.

For all subsequent results, $y$ was scaled by $1/s^2$ and analysed by linear models, as for quantitative traits, including mixed-modelling where appropriate for family studies. This method of estimating binary traits with simple linear models gives asymptotically unbiased estimates of $\beta_{FROH}$ and set($\beta_{FROH}$) on the In(Odds-Ratio) scale.

For all significant binary traits, a comparison of $\beta_{FROH}$ from the full model with $\beta_{FROH, lin}$ from the linear model approximation is presented in Supplementary Fig. 8.

To give $\beta_{FROH}$ a more tangible interpretation, effect estimates are frequently quoted in the text as $\beta_{ROH, OR}$ i.e. the estimated effect in the offspring of first cousins, where 6.25% of the genome is expected to be autosyzygous.

Reliability and educational attainment as additional covariates. To assess the importance of potential social confounders, proxy measures of socio-economic status and religiosity were separately included in Eq. (13) as additional covariates.

The modified effect estimates ($\hat{\beta}_{FROH}$) were tested for significance (Supplementary Data 20) and compared to the uncorrected estimates ($\beta_{FROH}$) (Supplementary Fig. 10a, b).

Since Educational Attainment (EA) was measured in many cohorts, this was chosen as the most suitable proxy for socio-economic status. However, since $F_{ROH}$ is known to affect EA directly any change in $\beta_{FROH}$ when conditioning on EA cannot be assumed to be entirely due to environmental confounding.

The analysis of religiosity was only carried out in UKB, where a rough proxy was available. Although no direct questions about religious beliefs were included, participants were asked about their leisure activities. In response to the question Which of the following do you attend once a week or more often? (You can select more than one), 15.6% of UKB participants selected Religious Group from one of the seven options offered. In the models described, religiosity was coded as 1 for those who selected Religious Group and 0 for those who did not. Although this is likely to be an imperfect measure of actual religious belief it is currently the best available in a large dataset.

Assortative mating. Humans are known to mate assortatively for a number of traits including height and cognition, and so we sought to investigate if this could influence our results, for example, by the trait extremes being more genetically similar and thus the offspring more homozygous. We see no evidence for an effect of assortative mating on autozygosity, however. Firstly, a polygenic risk score for height (see Supplementary Note 1), which explains 18.7% of the phenotypic variance in height, was not associated with $F_{ROH}$ ($p = 0.77$; Supplementary Fig. 5).

Secondly, linear relationships between traits and autozygosity extend out to very high $F_{ROH}$ individuals (Supplementary Figs. 9a–f). Samples in the highest $F_{ROH}$ group are offspring of genetically similar parents, very likely first or second degree relatives and, for example, the height of these samples is on average 3.4 cm [95% CI 2.5–4.3] shorter than the population mean. Assortative mating would suggest this height deficit has been inherited from genetically shorter parents, but this would require an implausibly strong relationship between short stature and a propensity to marry a close relative. Thirdly, the sex-specific effects we observe could only be explained by assortative mating if the additive heritability of these traits also differed by gender.

Average trait values in groups of similar $F_{ROH}$. In each cohort individuals were allocated to one of ten groups of similar $F_{ROH}$. The bounds of these groups were the same for all cohorts, specifically: 0.00, 0.02, 0.041, 0.0667, 0.0108, 0.0186, 0.0333, 0.06, 0.10, 0.18 and 1.0. Within each group the mean trait residual ($\epsilon$) and mean $F_{ROH}$ were calculated, along with their associated standard errors. Within each cohort the expectation of $\epsilon$ is zero at the mean $F_{ROH}$, however as mean $F_{ROH}$ varies between cohorts (Fig. 2, Supplementary Data 5) it was necessary to express $\epsilon$ relative to a common $F_{ROH}$. Hence, for this analysis only, the trait residuals ($\epsilon$) were expressed relative to the $F_{ROH} = 0$ intercept, i.e. by subtracting $\mu$ from Eq. (13).

Effect of $F_{ROH}$ within adoptees. We compared $\beta_{FROH, adoptees}$ to cross-cohort $\beta_{FROH}$, not that from UKB alone, as we consider the latter to be a noisy estimate of the former; estimates in UKB are consistent with those from meta-analysis.

Effect of $F_{ROH}$ within full-sibling families. In a subset of cohorts, with substantial numbers of related individuals, further analyses were performed to investigate the effect of $F_{ROH}$ within full-sibling families. In each of these cohorts, all second-degree, or closer, relatives were identified using KING (parameters: –related–degree 2). Full-siblings were then selected as relative pairs with genomic kinship >0.175 and IBS >0.001. This definition includes monozygotic twins, who were intentionally considered as part of full-sibling families. Although monozygotic twins are expected to have identical $F_{ROH}$, they may not have identical trait values, and including additional trait measurements decreases the sampling error of the within-family variance estimate, hence increasing statistical power. Dizygotic twins were also included.

For each individual ($i$) with identified siblings, the values of $F_{ROH}$ and trait residual ($\epsilon$) were calculated relative to their family mean (and called $F_{ROH_{wSibs}}$ and $\epsilon_{wSibs}$, respectively). For individual $j$ with $n$ full-siblings $S$, where $k \in [1..n]$ analysis using the B package metafor38. Results shown in Figs. 3–5 are meta-analysed averages of the within-cohort effects.

Between-cohort meta-analysis. As is typical in genome-wide association meta-analyses (GWAMA), genetic effects were estimated within single-ethnicity subcohorts and meta-analysis of the within-cohort effect sizes was used to combine results. This established method eliminates any potential confounding caused by between-cohort associations between $F_{ROH}$ and traits.

Each cohort returned estimates and standard errors of: $\beta_{FROH, j}$, $\beta_{FROH, j, ROH}$, $\beta_{FROH, j, wSibs}$, $\beta_{FROH, j, wSibs0}$, as well as trait means ($\bar{T}$) and standard errors within each of 10 $F_{ROH}$ bins. The between-cohort means of each of these 16 estimates was then determined by fixed-effect, inverse-variance meta-analysis using the R package metafor38. Results shown in Figs. 3–5 are meta-analysed averages of the within-cohort effects.

The meta-analysis was also run for various subsets of cohorts, stratified by ancestry as defined in Supplementary Data 18. Meta-analysis estimates from these groupings are shown in Supplementary Fig. 1.

Median and 95% CI of a ratio. In the analysis of adoptees (Fig. 5c), siblings (Fig. 5d) and potential confounders (Supplementary Figs. 10a, b) we wished to compare the effect estimates ($\hat{\beta}_{j, k}$) from two different methods across a wide range of traits. The units of $\beta_{j, k}$ differ by trait so, to allow comparison across all traits, the unitless ratio of effect size estimates was calculated (for example $\beta_{FSNPs, j, ROH} / \beta_{FSNPs, j, wSibs}$). Figure 5c, d and Supplementary Figs. 10a, b show the medians and 95% CI of these ratios. These were determined empirically by bootstrap since, although formulae exist for the mean and standard error of a ratio, the assumption of normality is violated when $\beta_{j, k}$ is small.

Genetic correlations in UK Biobank. Genetic correlations were calculated using LD-Score Regression41, implemented in LDSC v1.0.0 (https://github.com/bulik/ldsc). Summary statistics were parsed using default parameters in the LDSC.
8. Pemberton, T. J. et al. Genomic patterns of homozygosity in worldwide human populations. *Nature Genet.* 10, 783–796 (2009).
9. Bittles, A. H. & Black, M. L. Consanguinity, human evolution, and complex diseases. *Proc. Natl. Acad. Sci. USA.* 107, 1779–1786 (2010).
10. Nakatsuka, N. et al. The promise of discovering population-specific disease-associated genes in South Asia. *Nat. Genet.* 49, 1403–1407 (2017).
11. Narasimhan, V. M. et al. Health and population effects of rare gene knockouts in adult humans with related parents. *Science.* 352, 474–477 (2016).
12. Saleheen, D. et al. Runs of homozygosity in European populations. *Am. J. Hum. Genet.* 91, 275–292 (2012).
13. Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M. & Wilson, J. F. Runs of homozygosity: windows into population history and trait architecture. *Nat. Rev. Genet.* 19, 220–234 (2018).
14. McQuillin, R. et al. Runs of homozygosity in European populations. *Am. J. Hum. Genet.* 83, 359–372 (2008).
15. Purcell, S. & Chang, C. *PLINK 1.9.* www.cog-genomics.org/plink/1.9/.
16. McQuillin, R. et al. Evidence of inbreeding depression on human height. *PLoS ONE.* 8, e1002655 (2012).
17. Abdellauoi, A. et al. Association between autozygosity and major depression: stratification due to religious assortment. *Behav. Genet.* 43, 455–467 (2013).
18. Bashi, J. Effects of inbreeding on cognitive performance. *Nature.* 266, 440–442 (1977).
19. Broman, K. W. & Weber, J. L. Long homoygous chromosomal segments in reference families from the centre d’Etude du polymorphisme humain. *Am. J. Hum. Genet.* 63, 1071–1075 (1998).
20. Gibbons, J. M., Morton, N. E. & Collins, A. Extended tracts of homozygosity in outbred human populations. *Hum. Mol. Genet.* 15, 789–795 (2006).
21. Lee, J. S. et al. Genomic patterns of homozygosity in many cohorts is around 0.25 (the expectation in the offspring of first-degree relatives), and the effect size estimates from these samples are consistent with the overall estimates (Fig. 5c, d and Supplementary Fig. 9a–i).
David W Clark 1, Yukinori Okada 2,3,4, Kristjan H S Moore 5, Dan Mason 6, Nicola Pirastu 1, Ilaria Gardin 7,8, Hannele Mattsson 9,10, Catriona L K Barnes 11, Kuang Lin 11, Jing Hua Zhao 12,13, Patrick Deelen 14, Rebecca Rohde 15, Claudia Schurmann 16, Xiuqing Guo 17, Franco Giuliani 18, Weihua Zhang 19,20, Carolina Medina-Gomez 21,22,23, Robert Karlsson 14,24, Yanchun Bao 24, Shengchao Alfred Li 57, Nicola Pirastu 6, Dan Mason 5, Patrick Deelen 30,24, Christian Fuchsberger 39, He Gao 19,40, Marine Germain 41,42, Scott D Gordon 43.

Competing interests
M.L.A. is an employee of Genentech, a member of The Roche Group. D.A.L. has received support from several national and international government and charity funders, as well as Roche Diagnostics and Medtronic for work unrelated to this publication. M.M.G.: The views expressed in this article are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health. He has served on advisory panels for Pfizer, NovoNordisk, Zoe Global; has received honoraria from Merck, Pfizer, Novo-Nordisk and Eli Lilly; has stock options in Zoe Global; has received research funding from Abbvie, AstraZeneca, Boehringer Ingelheim, Eli Lilly, Janssen, Merck, Novo-Nordisk, Pfizer, Roche, Sanofi Aventis, Servier & Takeda. As of June 2019, M.M.C. is an employee of Genentech, and holds stock in Roche. T. Muka is now working as medical specialist at Novo Nordisk. O.P. is owner of Gen-infio Ltd. Gen-infio Ltd provided support in the form of salaries and financial gains for author O.P., but did not have any additional role in the selection of the journal or preparation of this manuscript. N. Poulier received financial support from several pharmaceutical companies which manufacture either blood pressure lowering or lipid lowering agents, or both, and consultancy fees. S.H. has participated in a congress trip sponsored by Novo Nordisk. P.J.S. has received research awards from Pfizer Inc. M.J.C. is Chief Scientist for Genomics England, a UK government company. B.M.P. serves on the DSMB of a clinical trial funded by Zoll LifeCor and on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson. A.R.S. is an employee of Regeneron Pharmacuetical Inc. The remaining authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-019-12283-6.

Peer review information Nature Communications thanks Jun Li, Matthew Robinson and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019.
John R.B. Perry, Tōnu Esko, Agnar Helgason, Kari Stefansson, Peter K Joshi, Michiaki Kubo & James F Wilson

1Centre for Global Health Research, Usher Institute, University of Edinburgh, Edinburgh EH8 9AG, Scotland. 2Department of Statistical Genetics, Osaka University Graduate School of Medicine, Suita, Osaka 565-0871, Japan. 3Laboratory of Statistical Immunology, Immunology Frontier Research Center (WPI-IFRCe), Osaka University, Suita, Osaka 565-0871, Japan. 4deCODE genetics/Amgen Inc., Reykjavik 101, Iceland. 5Bradford Institute for Health Research, Bradford Teaching Hospitals NHS Trust, Bradford BD96RJ, UK. 6Research Unit, Area Science Park, Trieste 34149, Italy. 7Department of Medicine, Surgery and Health Sciences, University of Trieste, Trieste, Italy. 8Unit of Public Health Solutions, National Institute for Health and Welfare, Helsinki, Finland. 9Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland. 10Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, UK. 11MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge CB2 0QQ, UK. 12Cardiovascular Epidemiology Unit, Department of Public health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK. 13Department of Genetics, University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands, Groningen, Groningen 9700 RB, The Netherlands. 14Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC 27514, USA. 15The Charles Bronfman Institute for Personalized Medicine, Ichan School of Medicine at Mount Sinai, New York, NY 10029, USA. 16Division of Genomic Outcomes, Department of Pediatrics, The Institute for Translational Genomics and Population Sciences, LABioMed at Harbor-UCLA Medical Center, Torrance, California 90502, USA. 17Division of Preventive Medicine, Brigham and Women’s Hospital, Boston, MA 02215, USA. 18Department of Epidemiology and Biostatistics, Imperial College London, London W2 1PG, UK. 19Department of Cardiology, Ealing Hospital, Middlesex, Middlesex UB1 3HW, UK. 20Department of Internal Medicine, Erasmus University Medical Center, Rotterdam 3015 CN, Netherlands. 21Department of Epidemiology, Erasmus University Medical Center, Rotterdam 3015 CN, Netherlands. 22The Generation R Study Group, Erasmus University Medical Center, Rotterdam 3015 CN, The Netherlands. 23Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm 17177, Sweden. 24Institute for Social and Economic Research, University of Essex, Colchester CO4 3SQ, UK. 25Cardiovascular Health Research Unit, Department of Statistics and Medicine, University of Washington, Seattle, WA 98101, USA. 26Research Unit of Molecular Epidemiology, Institute of Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg 85764, Germany. 27Institute of Molecular Genetics, National Research Council of Italy, Pavia 27100, Italy. 28Department of Biochemistry, University of Otago, Dunedin 9054, New Zealand. 29Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore 117549, Singapore. 30Department of Computational Biology, University of Lausanne, Lausanne 1011, Switzerland. 31Center for Primary Care and Public Health (Unisanté), University of Lausanne, Lausanne, Switzerland. 32Swiss Institute of Bioinformatics, Lausanne 1015, Switzerland. 33Division of Human Genetics, Children’s Hospital of Philadelphia, Philadelphia, PA 19104, USA. 34Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA. 35Department of Neurology, Brain Centre Rudolf Magnus, University Medical Centre Utrecht, Utrecht University, Utrecht 3584 CX, The Netherlands. 36Genomics Research Centre, School of Biomedical Sciences, Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Queensland 4059, Australia. 37Malaghan Institute of Medical Research, Wellington 6242, New Zealand. 38Institute for Biomedicine, Eurac Research, Affiliated Institute of the University of Lübeck, Bolzano 39100, Italy. 39MRC-PHE Centre for Environment and Health, Imperial College London, London W2 1PG, UK. 40INSERM UMR_S 1166, Sorbonne Universités, Paris 75013, France. 41ICAN Institute for Cardiometabolism and Nutrition, Paris 75013, France. 42QIMR Berghofer Institute of Medical Research, Brisbane, Australia. 43Department of Clinical Epidemiology, Leiden University Medical Center, Leiden 2333 ZA, The Netherlands. 44Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh EH9 9JZ, UK. 45Centre for Genetic and Experimental Medicine, Institute of Genetics & Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK. 46Clinical Division of Neurogenetics, Department of Neurology, Medical University of Graz, Graz 8036, Austria. 47Institute of Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz 8036, Austria. 48CONACyT, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico 03940, México. 49Department of Human Genetics, University of Chicago, Chicago, IL 60637, USA. 50Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Vrije Universiteit Amsterdam, Amsterdam 1081 HV, The Netherlands. 51Alzheimer Center Department of Neurology, VU University Medical Center, Amsterdam 1081HV, The Netherlands. 52Fondazione per la Ricerca su Malattie Neurologiche e Mentali (FISM), University of Milan, Milan 20133, Italy. 53Department of Neurology, University of Turin, Turin 10126, Italy. 54Department of Epidemiology, University of Turin, Turin 10126, Italy. 55Department of Public Health, University of Turin, Turin 10126, Italy. 56Department of Neurology, IRCCS Fondation Carlo Besta Neurological Institute, Milan 20133, Italy. 57Department of Pediatrics, College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA. 58Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892-5635, USA. 59Department of Epidemiology Research, Statens Serum Institut, Copenhagen DK-2300, Denmark. 60Icelandic Heart Association, Kopavogur 201, Iceland. 61Faculty of Medicine, School of Health Sciences, University of Iceland,
