nature genetics

ARTICLES

https://doi.org/10.1038/s41588-019-0403-1

Maternal and fetal genetic effects on birth weight and their relevance to cardio-metabolic risk factors

Birth weight variation is influenced by fetal and maternal genetic and non-genetic factors, and has been reproducibly associated with future cardio-metabolic health outcomes. In expanded genome-wide association analyses of own birth weight ($n=321{,}223$) and offspring birth weight ($n=230{,}069$ mothers), we identified 190 independent association signals (129 of which are novel). We used structural equation modeling to decompose the contributions of direct fetal and indirect maternal genetic effects, then applied Mendelian randomization to illuminate causal pathways. For example, both indirect maternal and direct fetal genetic effects drive the observational relationship between lower birth weight and higher later blood pressure: maternal blood pressure-raising alleles reduce offspring birth weight, but only direct fetal effects of these alleles, once inherited, increase later offspring blood pressure. Using maternal birth weight-lowering genotypes to proxy for an adverse intrauterine environment provided no evidence that it causally raises offspring blood pressure, indicating that the inverse birth weight–blood pressure association is attributable to genetic effects, and not to intrauterine programming.

Birth weight is an important predictor of newborn and infant survival, is a key indicator of pregnancy outcomes, and is observationally associated with future risk of adult cardio-metabolic diseases in offspring. These observational associations are often assumed to reflect adaptations made by a developing fetus in response to an adverse intrauterine environment—a concept termed the developmental origins of health and disease (DOHaD)[1]. Support for DOHaD is primarily from animal models (reviewed in ref. [2]). Observational studies of famine-exposed populations support prenatal programming in relation to type 2 diabetes (T2D), but not other cardio-metabolic health measures (reviewed in ref. [3]). However, DOHaD cannot provide a complete explanation for the relationship between lower birth weight and increased risk of cardio-metabolic disease. Other probable contributing factors are (1) environmental confounding, leading to phenotypic associations across the life-course[4] and (2) shared genetic effects operating at the population level[5]. Genetic associations between birth weight and later cardio-metabolic diseases may arise from the direct effects of the same inherited genetic variants at different stages of the life-course[6]. However, consideration of an individual's own genotype in isolation cannot exclude potential confounding by any indirect effects of the correlated maternal genotype ($r=\sim0.5$) on the intrauterine, and possibly postnatal, environment. Evidence for indirect maternal effects on birth weight and later offspring disease risk could implicate the intrauterine environment in later-life disease etiology.

To date, 65 genetic loci have been associated with birth weight in genome-wide association studies (GWASs), implicating biological pathways that may underlie observational associations with adult disease[5,7–9]. However, most of these studies did not distinguish between maternal and fetal genetic influences. Evidence from monogenic human models[10] and variance component analyses[11] demonstrates that birth weight is influenced by genotypes inherited by the fetus and by maternal genotypes that influence the intrauterine environment. To date, GWASs of own birth weight[5] and maternal GWASs of offspring birth weight[7] have produced overlapping signals due to the correlation between maternal and fetal genotypes. Identified birth weight variants might have (1) a direct fetal effect only; (2) an indirect maternal effect only; or (3) some combination of the two. Performing separate GWAS analyses of own or offspring birth weight precludes full resolution of the origin of the identified genetic effects.

To address these issues, we performed greatly expanded GWASs of own ($n=321{,}223$) and offspring birth weight ($n=230{,}069$ mothers) using data from the Early Growth Genetics (EGG) Consortium and the UK Biobank (2017 release). We applied a structural equation model that we recently developed to partition genetic effects on birth weight into maternal and fetal components at genome-wide-significant loci[7,12]. We then extended the method to estimate maternal- and fetal-specific genetic effects across the genome in a computationally efficient manner, and used the results for downstream analyses. Our ability to resolve maternal and fetal genetic contributions provides substantial insights into the underlying biological regulation of birth weight, as well as the origins of observational relationships with T2D and blood pressure.

Results

Meta-analyses of fetal and maternal GWASs. We conducted GWAS meta-analyses of own (fetal) genetic variants on own birth weight (Supplementary Fig. 1 and Supplementary Tables 1 and 2) and maternal genetic variants on offspring birth weight (Supplementary Fig. 2 and Supplementary Tables 3 and 4) in individuals of European ancestry. We then performed approximate conditional and joint multiple single-nucleotide polymorphism (SNP) analyses (COJO[13]) and a trans-ethnic meta-analysis to identify further independent SNPs (Methods). The GWAS meta-analysis of own birth weight ($n=321{,}223$) identified 146 independent SNPs at genome-wide significance ($P<6.6\times10^{-9}$; Supplementary Figs. 3, 4 and 5a, Supplementary Table 5a and Methods). The GWAS meta-analysis of offspring birth weight ($n=230{,}069$ mothers) identified 72 independent SNPs ($P<6.6\times10^{-9}$; Supplementary Figs. 3, 4 and 5b, Supplementary Table 5a and Methods). Applying the more lenient significance threshold used previously ($P<5\times10^{-8}$)[5,7], 211 and 105 SNPs reached significance for own and offspring birth weight, respectively (Supplementary Table 5b).

SNPs at 30 genome-wide-significant loci (within 500 kilobases (kb) and linkage disequilibrium (LD) $r^2\geq0.1$) were identified in the GWASs of both own and offspring birth weight. Of these, 9 loci had

A full list of authors and affiliations appears at the end of the paper.

the same lead SNP and 21 loci had correlated fetal and maternal lead SNPs ($r^2 \geq 0.1$). Colocalization analysis indicated that 19/21 of these correlated lead SNP pairs were probably tagging the same birth weight signal (posterior probability > 0.5). Therefore, we identified a total of 190 independent association signals, represented by 209 lead SNPs (Supplementary Fig. 4 and Supplementary Table 5a). Of the 209 lead SNPs, 146 are novel (representing 129 independent association signals), 3 are rare (minor allele frequency (MAF) < 1%) and 13 are low frequency (1% ≤ MAF < 5%). The three rare variants (at the *YKT6/GCK*, *ACVR1C* and *MIR146B* loci) alter birth weight by more than double the effect (>100 g per allele) of the first common variants identified[9]. In the independent Norwegian Mother and Child Cohort Study (MoBa–HARVEST; $n = 13{,}934$ mother–offspring duos), the lead SNPs explained 7% of the variance in birth weight, calculated as the sum of variances explained by the fetal genotype (6%) and maternal genotype (2%), plus twice the covariance (−0.5%). Maternal genome-wide complex trait analysis (M-GCTA[11]), which estimates SNP heritability, and partitions this quantity into maternal and fetal components, estimated that 39.8% of the variance in birth weight could be explained by tagged fetal genetic variation (28.5%), tagged maternal genetic variation (7.6%) and twice the covariance (3.7%).

We integrated data from several sources to highlight possible causal genes underlying the identified associations, including gene-level expression data across 43 tissues (from GTEx version 6p[14]), placental expression quantitative trait loci (eQTL)[15], topologically associating domains (TADs) identified in human embryonic stem cells[16,17] and non-synonymous SNPs (see Supplementary Table 5a,b and Methods). Several genes were highlighted by multiple approaches; however, further functional studies are required to confirm causality.

Structural equation model to estimate maternal and fetal effects. Next, we partitioned the 209 lead SNPs into 5 categories based on their maternal and/or fetal genetic contributions to birth weight. To achieve this, we used structural equation modeling (SEM), which accounts for the correlation between fetal and maternal genotypes and thereby provides unbiased estimates of the maternal and fetal genetic effects on birth weight[12] (see Methods and Supplementary Fig. 6a for details). The results are presented in Fig. 1, Supplementary Figs. 4 and 7 and Supplementary Table 5. Using the confidence intervals (CIs) around the SEM-adjusted maternal and fetal effect estimates, we identified 64 SNPs with fetal-only effects, 32 SNPs with maternal-only effects, 27 SNPs with directionally concordant fetal and maternal effects and 15 SNPs with directionally opposing fetal and maternal effects (Supplementary Fig. 8). For example, rs10830963 at *MTNR1B* was identified in both the own-birth-weight ($P = 2.8 \times 10^{-11}$) and offspring-birth-weight GWAS ($P = 9.1 \times 10^{-39}$), but the SEM analysis revealed that its effect was exclusively maternal ($P_{\text{SEMfetal}} = 0.7$; $P_{\text{SEMmaternal}} = 4.6 \times 10^{-19}$). In contrast, rs560887 at *G6PC2* was identified only in the GWAS of offspring birth weight ($P = 1.2 \times 10^{-14}$), but was found to have directionally opposing maternal and fetal effects ($\beta_{\text{SEMfetal}} = -0.03$; $P_{\text{SEMfetal}} = 2.8 \times 10^{-8}$; $\beta_{\text{SEMmaternal}} = 0.04$; $P_{\text{SEMmaternal}} = 5.4 \times 10^{-14}$). At present, these categories are suggestive, as the current sample size has insufficient statistical power to detect small genetic effects, particularly maternal effects. Some SNPs that were classified as fetal only may have had a small undetected maternal effect. In addition, 71 SNPs remained unclassified. Asymptotic power calculations showed that we had 80% power to detect fetal (maternal) effects that explained 0.006% (0.008%) of the variance in birth weight ($\alpha = 0.05$). However, there was strong consistency with traditional conditional linear regression modeling in $n = 18{,}873$ mother–offspring pairs (Supplementary Table 6 and Methods), and overall, the method gave a clear indication as to which genetic associations are driven by the maternal or fetal genomes.

To extend the estimates of adjusted maternal and fetal effects genome wide, we developed a weighted linear model (WLM) (see Methods), which yields a good approximation to the SEM with equivalent estimates for the 209 lead SNPs (Supplementary Fig. 9). This was necessary because the SEM is too computationally intensive to fit across the genome. The adjusted fetal and maternal genotype effect estimates on birth weight from the WLM are hereafter referred to as WLM-adjusted estimates. Using linkage disequilibrium score regression[18], we observed that the genetic correlation between the WLM-adjusted maternal and fetal effects ($r_g = 0.10$; $P = 0.12$) was substantially lower than that between the unadjusted effects from the original GWAS ($r_g = 0.82$; $P < 0.01$), indicating that the WLM largely accounts for the underlying correlation between fetal and maternal genotypes. No additional novel loci were identified, but we used the WLM-adjusted estimates in downstream analyses to identify fetal- and maternal-specific mechanisms that regulate birth weight, and to investigate the genetic links between birth weight and adult traits.

Maternal- and fetal-specific tissues and mechanisms underlying birth weight regulation. Tests of global enrichment of birth weight SNP associations across tissues sampled from the GTEx project[14] using LD-SEG[19] are presented in Supplementary Fig. 10. Only enrichment for maternal-specific SNP associations for genes expressed in connective/bone tissues was detected after Bonferroni correction. Integration of epigenetic signatures defined by the Roadmap Epigenomics Project highlighted a significant enrichment of maternal-specific effects in the ovary for histone modification marks (H3K4me1) and regions of open chromatin (Supplementary Table 7); no significant enrichment was detected for other signatures. Gene set enrichment analysis implicated different fetal- (Supplementary Table 8) and maternal-specific (Supplementary Table 9) gene sets.

A major determinant of birth weight is the duration of gestation. Using linkage disequilibrium score regression[18], we found a substantial genetic correlation between published maternal genetic effects on gestational duration[20] and the WLM-adjusted maternal effects on offspring birth weight ($r_g = 0.63$; $P = 2.1 \times 10^{-5}$; Supplementary Table 10 and Methods), but not with the WLM-adjusted fetal effects on own birth weight ($r_g = -0.10$; $P = 0.34$). Gestational duration was unavailable for >85% of individuals in the birth weight GWAS analyses, so it is possible that some identified association signals influence birth weight primarily by altering gestational duration. We looked up the 209 lead birth-weight-associated SNPs in the published maternal GWAS of gestational duration[20] (Supplementary Table 11) and followed up 7 associated SNPs ($P < 2.4 \times 10^{-4}$, corrected for 209 tests; Methods) in 13,206 mother–child pairs. Meta-analysis with summary data from 23andMe[20] strengthened associations with gestational duration at five of seven loci (*EBF1*, *AGTR2*, *ZBTB38*, *KCNAB1* and *KLHL25/AKAP13*; Supplementary Table 12). The precise causal relationship between fetal growth and gestational duration at these loci requires further investigation; however, the majority of associations with birth weight do not appear to be driven by associations with gestational duration.

Maternal- and fetal-specific genetic correlations between birth weight and adult traits. The 209 lead birth-weight-associated SNPs were associated with other phenotypes in previously published GWASs and the UK Biobank (Supplementary Table 13 and Methods). At the genome-wide level, we previously reported genetic correlations between own birth weight and several adult cardiometabolic traits[5], but were unable to distinguish the direct fetal genotype contribution from the indirect contribution of maternal genotype. To understand these distinct contributions, we calculated genetic correlations using linkage disequilibrium score regression[18] between WLM-adjusted fetal and maternal SNP effect estimates

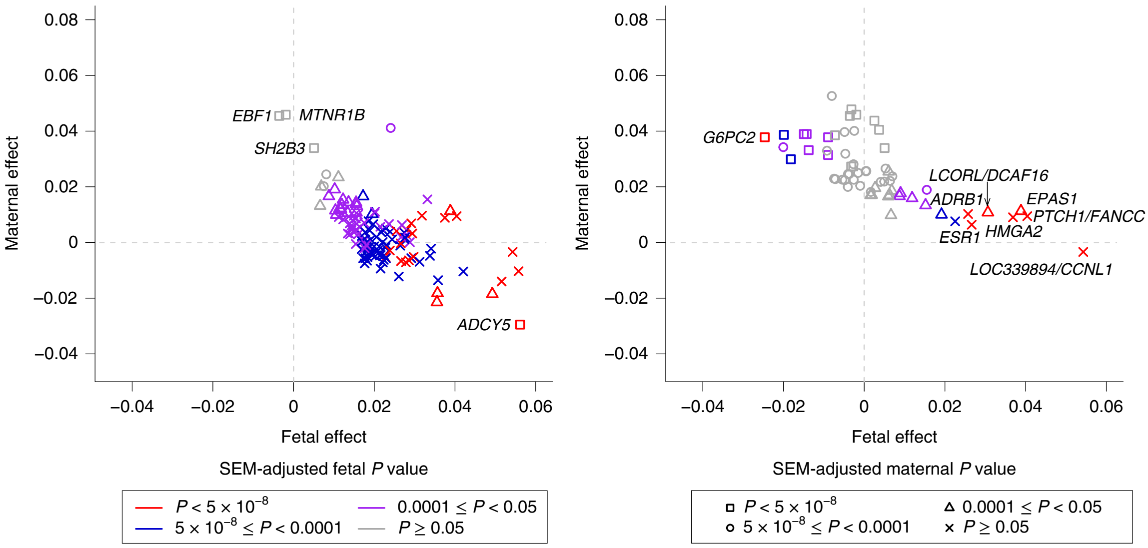

Fig. 1 | SEM-adjusted fetal and maternal effects for the 193 lead SNPs that were identified in the GWASs of either own birth weight or offspring birth weight with a MAF greater than 5%. Left, own birth weight. Right, offspring birth weight. The SEM included 85,518 individuals from the UK Biobank with both their own and their offspring's birth weight, and 178,980 and 93,842 individuals from the UK Biobank and EGG Consortium with only their own birth weight or only their offspring's birth weight, respectively. The color of each point indicates the SEM-adjusted fetal effect on the own-birth-weight association *P* value and the shape of each point indicates the SEM-adjusted maternal effect on the offspring-birth-weight association *P* value. *P* values for the fetal and maternal effect were calculated using a two-sided Wald test. SNPs labeled with the name of the closest gene are those that were identified in the GWAS of own birth weight but whose effects are mediated through the maternal genome (left) and SNPs that were identified in the GWAS of offspring birth weight but whose effects are mediated through the fetal genome (right). SNPs are aligned to the birth-weight-increasing allele from the GWAS.

and GWAS estimates for a range of traits (Fig. 2, Supplementary Table 10 and Methods). For many traits (for example, adult height), the fetal-specific genetic correlation was similar to the maternal-specific genetic correlation, but for some traits, the fetal- and maternal-specific genetic correlations were different in magnitude (for example, systolic blood pressure (SBP)) or even in direction (for example, T2D). For several glycemic traits (for example, fasting glucose), the genetic correlations estimated using the WLM-adjusted effects were substantially different from those estimated using the unadjusted effects, showing the importance of accounting for the maternal–fetal genotype correlation.

Using genetics to estimate causal effects of intrauterine exposures on birth weight. The separation of direct fetal genotype effects from indirect maternal genotype effects on birth weight offers the novel opportunity to estimate unconfounded causal influences of intrauterine exposures by using Mendelian randomization analyses. The principle of Mendelian randomization is similar to that of a randomized controlled trial: parental alleles are randomly transmitted to offspring and are therefore generally free from confounding[21,22]. Consequently, an association between a maternal genetic variant for an exposure of interest, and offspring birth weight, after accounting for fetal genotype, provides evidence that the maternal exposure is causally related to offspring birth weight (Fig. 3a). Previous attempts to estimate causal effects of maternal exposures on offspring birth weight were limited by an inability to adjust for fetal genotype in adequately powered samples[23], which can now be overcome by using WLM-adjusted estimates. We applied two-sample Mendelian randomization[24] to estimate causal effects of maternal exposures on offspring birth weight, focusing on height, glycemic traits and blood pressure. We selected SNPs known to be associated with each exposure, and regressed the WLM-adjusted maternal effects on birth weight for those SNPs against the effect estimates for the maternal exposure, weighting by the inverse of the variance of the maternal exposure effect estimates. In the same way, we used the WLM-adjusted fetal effects to estimate the casual effect of the offspring's genetic potential on their own birth weight, and compared the results with the estimated maternal causal effects.

Height and birth weight. Classical animal experiments[25] showed that larger maternal size can support greater fetal growth. This is supported by observational human data showing that offspring height shifts from being closer to the maternal than the paternal height percentile in infancy towards mid-parental height in adulthood, the latter reflecting the predominant role of inherited genetic variation[26]. However, several observational studies have provided mixed evidence regarding correlations between maternal or paternal height and offspring birth weight. Some studies show a stronger correlation with maternal than paternal height[27,28], which would be consistent with a role for intrauterine effects, while others show that maternal and paternal height are both strongly correlated with offspring birth weight[29–31]. The Mendelian randomization analysis, using 693 height-associated SNPs[32] (Supplementary Table 14), estimated that a 1 s.d. (6 cm) higher maternal height is causally associated with a 0.11 s.d. (95% CI: 0.10 to 0.13) higher offspring birth weight (Fig. 3b), independent of the direct fetal effects. A similar estimate was obtained using the WLM-adjusted fetal effects on own birth weight (0.11 s.d. (95% CI: 0.09 to 0.13)), reflecting the role of inherited height alleles (Supplementary Table 15). Both a previous study[33] and complementary analysis using transmitted and non-transmitted height alleles in mother–offspring pairs estimated a much larger contribution of direct fetal effects than indirect maternal effects to offspring birth weight (Supplementary Table 16 and Methods), but with relatively small sample sizes. In contrast with a previous report[33], there was little supportive evidence that the maternal height effect on birth weight was via prolonged gestation ($P = 0.12$; Supplementary Table 15). These Mendelian randomization results are consistent with the hypothesis that greater maternal height causally increases birth weight, and that this effect is independent of the direct birth weight-raising effect of height alleles

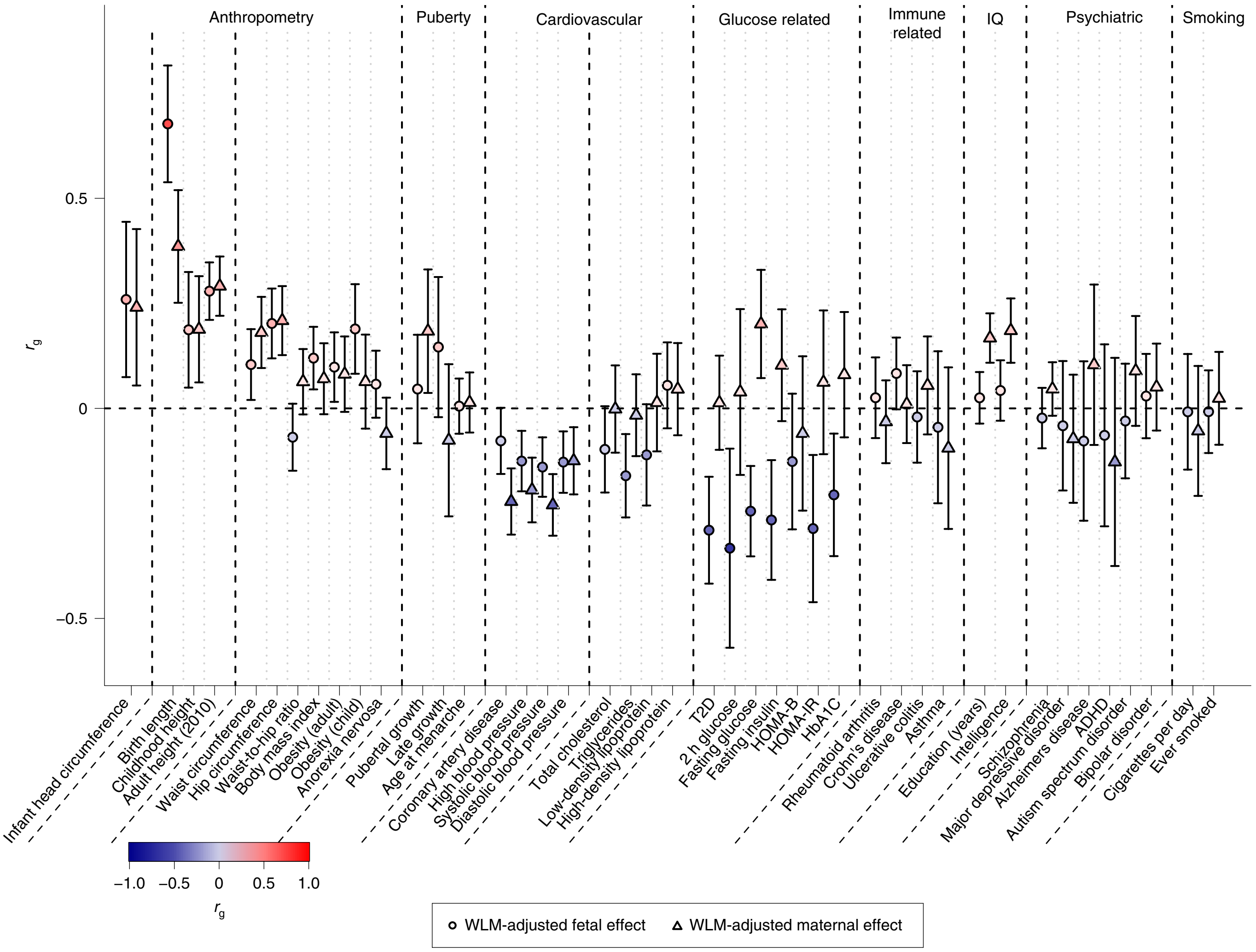

Fig. 2 | Genome-wide genetic correlation between birth weight and a range of traits and diseases in later life. Genetic correlation (r_g) between birth weight and the traits, and corresponding 95% CIs, were estimated using linkage disequilibrium score regression in LD Hub. Genetic correlations were estimated from the summary statistics of the WLM-adjusted fetal GWAS (WLM-adjusted fetal effect on own birth weight) and WLM-adjusted maternal GWAS (WLM-adjusted maternal effect on offspring birth weight). The total sample size included in the WLM-adjusted GWAS was 406,063 individuals with their own and/or their offspring's birth weight. The genetic correlation estimates are color coded according to their intensity and direction (red, positive correlation; blue, negative correlation). ADHD, attention deficit hyperactivity disorder; HbA1c, hemoglobin A1c; HOMA-B, homeostasis model assessment of β-cell function; HOMA-IR, homeostasis model assessment of insulin resistance. See Supplementary Table 10 for the references for each of the traits and diseases displayed, as well as the genetic correlation results for other traits and diseases.

inherited by the fetus. Although greater availability of space for fetal growth is a possible explanation, we cannot rule out other causal pathways. For example, causal associations between greater height and a more favorable socio-economic position[34] could enhance maternal nutritional status and result in higher offspring birth weight. We also cannot exclude the contribution of assortative mating[35] to these results: correlation between maternal and paternal height genotypes could lead to similar maternal and fetal Mendelian randomization estimates.

Glycemic traits and birth weight. Maternal glucose is a key determinant of fetal growth: it crosses the placenta, stimulating the production of fetal insulin, which promotes growth[36]. As a consequence, strong positive associations are seen between maternal fasting glucose, fetal insulin levels and offspring birth weight[37]. In a randomized clinical trial of women with gestational diabetes mellitus, glucose control was shown to reduce offspring birth weight[38]. Therefore, we anticipated detecting a positive causal effect of maternal glucose on offspring birth weight, as previously observed using Mendelian randomization in a smaller sample[23]. Indeed, the Mendelian randomization analysis using 33 fasting glucose-associated SNPs (Supplementary Table 14) estimated a 0.18 s.d. (95% CI: 0.13 to 0.23) higher offspring birth weight due to 1 s.d. ($0.4\,\mathrm{mmol\,l^{-1}}$) higher maternal fasting glucose, independent of the direct fetal effects (Supplementary Table 15 and Fig. 3c). A large part of the genetic variation underlying fasting glucose levels is implicated in pancreatic β-cell function and thus overlaps with the genetics of insulin secretion. To estimate the causal effect of insulin secretion on birth weight, we used 18 SNPs associated with the disposition index—a measure of insulin's response to glucose, adjusted for insulin sensitivity. Alleles that increase insulin secretion in the mother tend to decrease her glucose levels, which consequently reduces insulin-mediated growth of the fetus. This was reflected in the negative causal estimate from the Mendelian randomization analysis of the effect of the maternal disposition index on offspring birth weight (−0.17 s.d. per 1 s.d. higher maternal disposition index (95% CI: −0.26 to −0.08); Supplementary Table 15). In contrast, we estimated that birth weight was 0.10 s.d. (95% CI: 0.02 to 0.19) higher

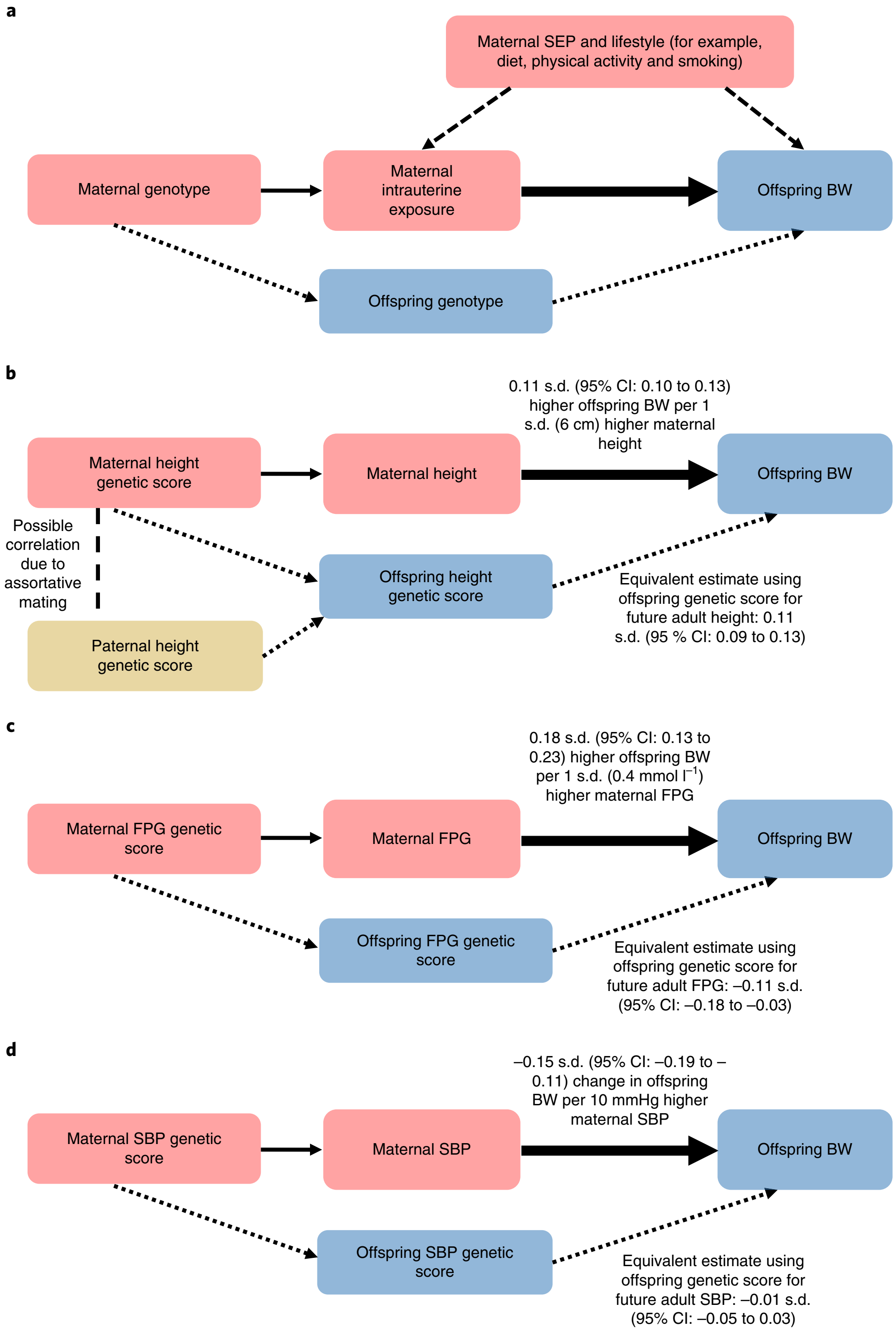

Fig. 3 | Mendelian randomization to assess the causal effect of maternal intrauterine exposures on offspring birth weight (adapted from Lawlor et al.[45]). **a**, Since maternal and fetal genotypes are correlated, it is essential to account for offspring genotype in this analysis. The continuous, thin arrow represents the relationship between the genetic instrument and intrauterine exposure. The dashed arrows represent potential confounding via maternal characteristics, which, under Mendelian randomization assumptions, are not associated with the genetic instrument. The dotted arrows represent potential violation of Mendelian randomization assumptions via offspring genotype. The thick arrow represents the causal effect of interest. **b**, Higher offspring birth weight is caused by direct fetal genetic effects of height-raising alleles and indirect effects of maternal height-raising alleles. Indirect maternal effects of height-raising alleles may increase offspring birth weight by increasing the space available for growth, but we cannot rule out alternative explanations; for example, assortative mating. **c**, Higher maternal fasting glucose levels increase offspring birth weight. Conversely, direct fetal genetic effects of glucose-raising alleles reduce birth weight. This is probably due to their effects on insulin: variants that lower maternal insulin levels increase maternal glucose, which crosses the placenta and stimulates fetal insulin-mediated growth. However, the same variants in the fetus cause lower fetal insulin levels, and consequently, reduced fetal insulin-mediated growth. **d**, Higher maternal SBP is causally associated with lower offspring birth weight. After adjusting for maternal effects, there was no evidence of an effect of offspring's own SBP genetic score on own birth weight. BW, birth weight; FPG, fasting plasma glucose; SEP, socio-economic position. 1 s.d. of BW = 484 g (refs. [9,45]).

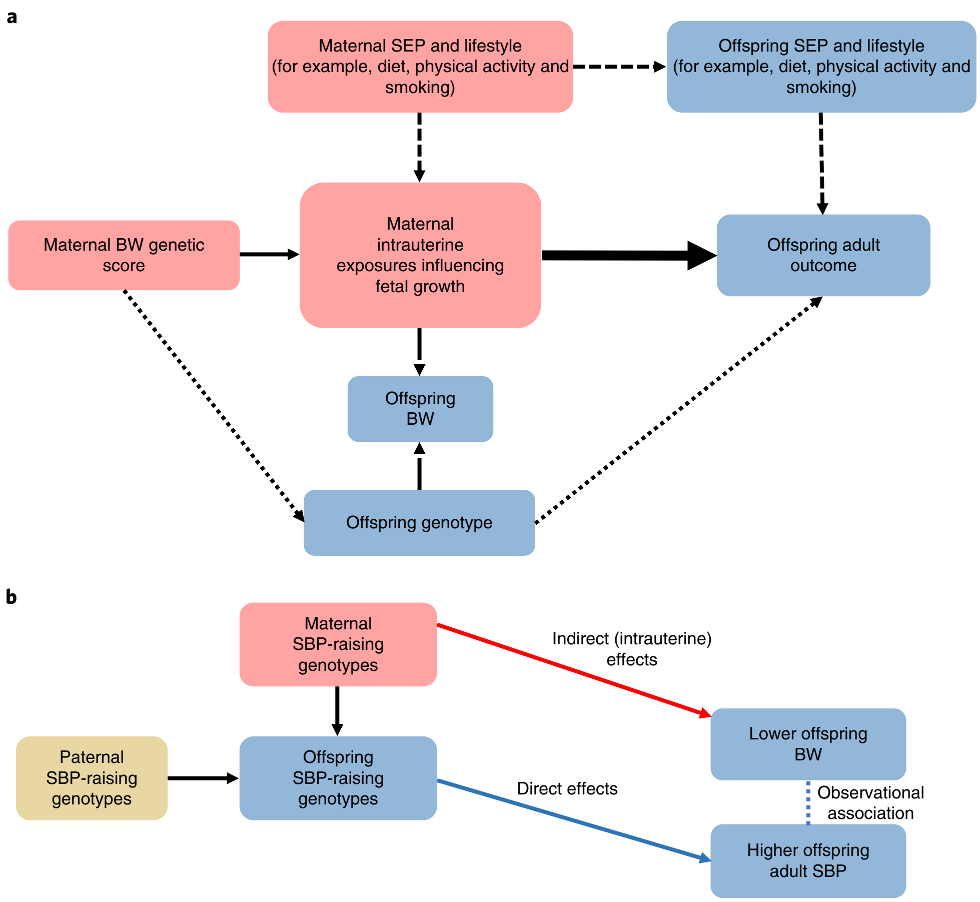

Fig. 4 | Mendelian randomization to assess the causal effect of intrauterine growth on offspring adult outcomes, using maternal intrauterine exposures that influence fetal growth. a, Maternal genotype should be associated with offspring birth weight independent of offspring genotype, so it is essential to adjust the analysis for offspring genotype. The continuous, thin arrow represents the relationship between the genetic instrument and intrauterine exposure. The long-dashed arrows denote the (maternal and possibly fetal) genotype associations with birth weight; these arrows highlight the assumption that genetic variation influences offspring adult outcome via intrauterine growth, not birth weight. The short-dashed arrows represent potential confounding via maternal and offspring characteristics. The dotted arrow represents potential violation of assumptions of the Mendelian randomization analysis via offspring genotype. The thick arrow represents the causal effect of interest. We have not estimated the size of the causal effect as we do not have effect estimates for the SNP-maternal intrauterine exposures influencing fetal growth. However, we have used the presence versus absence and direction of association in 3,886 mother–offspring pairs to indicate whether the intrauterine environment causes changes in adult offspring SBP (see Supplementary Table 18 for full results). **b**, Our results show that the observed negative correlation between birth weight and later SBP may be driven by the causal effect of higher maternal SBP on lower offspring birth weight (red arrow), in combination with the subsequent transmission of SBP-associated alleles to offspring (blue arrow), which then increase offspring SBP.

per 1 s.d. genetically higher fetal disposition index (Methods), highlighting that genetic variation underlying insulin secretion plays a key role in fetal growth, and suggesting that the genetic effects on the disposition index are similar in fetal and adult life.

Birth weight associations with previously reported GWAS SNPs for fasting glucose, T2D, insulin secretion and insulin sensitivity loci were directionally consistent with the overall genetic correlations, and supported the opposing contributions of fetal versus maternal glucose-raising alleles on birth weight (Supplementary Figs. 11–14). Taken together with the WLM-adjusted genetic correlations, the Mendelian randomization results underline the importance of fetal insulin in fetal growth, and show that fetal genetic effects link lower birth weight with reduced insulin secretion and higher T2D risk in later life[6]. However, further work is needed to investigate the role of indirect maternal genetic effects in the relationship between high birth weight and higher future risk of T2D.

Blood pressure and birth weight. Observational studies of the relationship between birth weight and later-life blood pressure have produced mixed findings. Some studies indicate that lower birth weight is associated with higher later-life blood pressure and related comorbidities[39], whereas others have shown that this relationship could be driven by a statistical artifact due to adjusting for current weight[40,41]. We have previously shown that genetic factors account for a large proportion of an association between lower birth weight and higher blood pressure[5], but it was not clear whether this was due to direct fetal genotype effects or indirect maternal effects, or a combination of the two. We explored this association further using several complementary analyses. The estimate of the birth weight–SBP covariance explained was higher when using the maternal genotyped SNP associations with offspring birth weight (65% (95% CI: 57 to 74%)) than when using the fetal genotype associations with own birth weight (56% (95% CI: 48 to 64%); Supplementary Table 17). A similar pattern was seen with the birth weight–diastolic blood pressure (DBP) covariance (maternal: 72% (95% CI: 58 to 85%); fetal: 56% (95% CI: 46 to 67%); Supplementary Table 17). Together with the larger maternal than fetal genetic correlation for SBP (Fig. 2), these results point to the importance of indirect maternal effects of blood pressure on offspring birth weight (Supplementary Figs. 15 and 16). In line with this, Mendelian randomization analyses indicated that a 1 s.d. (10 mmHg) higher maternal SBP is causally associated with a 0.15 s.d. (95% CI: −0.19 to −0.11) lower offspring birth weight, independent of the direct fetal effects. In contrast, there was no fetal effect of SBP on own birth weight, after adjusting for the indirect maternal effect (−0.01 s.d. per 10 mmHg, 95% CI: −0.05 to 0.03; Fig. 3d and Supplementary Tables 14 and 15). Similar results were seen in the WLM-adjusted Mendelian randomization analyses of DBP on offspring and own birth weight.

Estimating the causal effect of birth weight-lowering intrauterine exposures on offspring SBP. A key question is whether maternal SNPs that reduce offspring birth weight through intrauterine effects are also associated with higher SBP in their adult offspring. Such an association would suggest that the maternal intrauterine effects cause the later SBP effect (that is, through developmental adaptations) (Fig. 4a and Supplementary Fig. 17). To investigate this possibility, we tested the conditional association between maternal and offspring genetic scores for birth weight and offspring SBP, as measured in 3,886 mother–offspring pairs in the UK Biobank, with sensitivity analyses in 1,749 father–offspring pairs. The fetal genetic score for lower birth weight was associated with higher offspring SBP, even after adjustment for maternal (or paternal) birth weight genotypes. However, when adjusted for fetal genotypes, the maternal genetic score for lower birth weight was associated with lower (not higher) offspring SBP (Supplementary Table 18). Taken together, our results show that the observed negative correlation between birth weight and later SBP is driven by (1) the causal effect of higher maternal SBP on lower offspring birth weight (Fig. 3d), in combination with (2) the subsequent transmission of SBP-associated alleles to offspring, which then increase offspring SBP (Fig. 4b), rather than by long-term developmental responses to adverse in utero conditions.

Discussion

In greatly expanded GWASs and follow-up analyses of own and offspring birth weight, we have identified 129 novel association signals and partitioned the genetic effects on birth weight into direct fetal and indirect maternal (intrauterine) effects. Using these partitioned effects, we identified fetal- and maternal-specific mechanisms and tissues involved in the regulation of birth weight, and mechanisms with directionally opposing effects in the fetus and mother (for example, insulin secretion and fasting glucose).

Mendelian randomization analyses using the WLM-adjusted estimates showed that (1) both direct fetal and indirect maternal effects of height-raising genotypes contribute to higher offspring birth weight; (2) fetal (and not maternal) genotype effects explain the negative genetic correlation between birth weight and later T2D; and (3) the negative genetic correlation between birth weight and adult SBP is the result of both indirect SBP-raising effects of maternal genotypes reducing offspring birth weight, and direct effects of fetal genotypes on higher adult SBP. The resolution of maternal versus fetal effects was higher in these Mendelian randomization analyses than has previously been achieved using analyses of available mother–child pairs[23], due to greater statistical power. Recently, a number of authors have attempted to use Mendelian randomization methodology to investigate causal links between birth weight and later T2D[42–44]. However, such naive Mendelian randomization analyses using two-sample approaches in unrelated sets of individuals, which do not properly account for the correlation between maternal and fetal genotype effects, may result in erroneous conclusions regarding causality. Future investigations into causal links between birth weight and later T2D or other disease outcomes will require larger samples than are currently available, with maternal and offspring genotypes and offspring later-life disease outcomes.

There are some limitations to this study (see Supplementary Note for a full discussion). First, the Mendelian randomization results concern birth weight variation within the normal range and do not necessarily reflect the effects of extreme environmental events (for example, famine), which may exert qualitatively different effects. Additionally, we have assumed a linear relationship between birth weight and later-life traits, which may be an oversimplification for some traits such as T2D. Second, birth weight is the end marker of a developmental process, with critical periods during the process that may make the fetus particularly sensitive to environmental influences. The Mendelian randomization analyses could therefore be masking effects at certain critical periods. Third, we have assumed that genetic variants identified in large GWASs of SBP and glycemic traits in males and non-pregnant females are similarly associated in pregnant women. Fourth, we have not investigated the potential gender difference in the associations between birth weight and later-life traits. Fifth, we have assumed that the critical period of exposure to indirect maternal genetic effects is pregnancy, and that the estimates do not reflect pre-pregnancy effects on primordial oocytes or postnatal effects[45]. Sixth, we have not considered paternal genotypes, and it is possible that this omission has biased the results of some of our analyses. Finally, although we were able to fit the full SEM at the 209 lead SNPs, we were unable to fit the SEM, including the 2 degrees of freedom test (that is, where maternal and fetal paths are constrained to 0) at all SNPs across the genome.

To conclude, the systematic separation of fetal from maternal genetic effects in a well-powered study has enhanced our understanding of the regulation of birth weight and of its links with later cardio-metabolic health. In particular, we show that the association between lower birth weight and higher adult blood pressure is attributable to genetic effects, and not to intrauterine programming. In successfully separating maternal from fetal genetic effects and using them in Mendelian randomization analyses, this work sets a precedent for future studies seeking to understand the causal role of the intrauterine environment in later-life health.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41588-019-0403-1.

Received: 29 June 2018; Accepted: 26 March 2019;

References

1. Barker, D. J. et al. Type 2 (non-insulin-dependent) diabetes mellitus, hypertension and hyperlipidaemia (syndrome X): relation to reduced fetal growth. *Diabetologia* **36**, 62–67 (1993).
2. Martin-Gronert, M. S. & Ozanne, S. E. Mechanisms underlying the developmental origins of disease. *Rev. Endocr. Metab. Disord.* **13**, 85–92 (2012).
3. Lumey, L. H., Stein, A. D. & Susser, E. Prenatal famine and adult health. *Annu. Rev. Public Health* **32**, 237–262 (2011).
4. Ben-Shlomo, Y. & Smith, G. D. Deprivation in infancy or in adult life: which is more important for mortality risk? *Lancet* **337**, 530–534 (1991).
5. Horikoshi, M. et al. Genome-wide associations for birth weight and correlations with adult disease. *Nature* **538**, 248–252 (2016).
6. Hattersley, A. T. & Tooke, J. E. The fetal insulin hypothesis: an alternative explanation of the association of low birthweight with diabetes and vascular disease. *Lancet* **353**, 1789–1792 (1999).
7. Beaumont, R. N. et al. Genome-wide association study of offspring birth weight in 86 577 women identifies five novel loci and highlights maternal genetic effects that are independent of fetal genetics. *Hum. Mol. Genet.* **27**, 742–756 (2018).
8. Horikoshi, M. et al. New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism. *Nat. Genet.* **45**, 76–82 (2013).
9. Freathy, R. M. et al. Variants in ADCY5 and near CCNL1 are associated with fetal growth and birth weight. *Nat. Genet.* **42**, 430–435 (2010).

10. Hattersley, A. T. et al. Mutations in the glucokinase gene of the fetus result in reduced birth weight. *Nat. Genet.* **19**, 268–270 (1998).
11. Eaves, L. J., Pourcain, B. S., Smith, G. D., York, T. P. & Evans, D. M. Resolving the effects of maternal and offspring genotype on dyadic outcomes in genome wide complex trait analysis ("M-GCTA"). *Behav. Genet.* **44**, 445–455 (2014).
12. Warrington, N. M., Freathy, R. M., Neale, M. C. & Evans, D. M. Using structural equation modelling to jointly estimate maternal and fetal effects on birthweight in the UK Biobank. *Int. J. Epidemiol.* **47**, 1229–1241 (2018).
13. Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 361–363 (2012).
14. GTEx Consortium. The genotype-tissue expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).
15. Peng, S. et al. Expression quantitative trait loci (eQTLs) in human placentas suggest developmental origins of complex diseases. *Hum. Mol. Genet.* **26**, 3432–3441 (2017).
16. Way, G. P., Youngstrom, D. W., Hankenson, K. D., Greene, C. S. & Grant, S. F. Implicating candidate genes at GWAS signals by leveraging topologically associating domains. *Eur. J. Hum. Genet.* **25**, 1286–1289 (2017).
17. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
18. Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
19. Finucane, H. K. et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* **50**, 621–629 (2018).
20. Zhang, G. et al. Genetic associations with gestational duration and spontaneous preterm birth. *N. Engl. J. Med.* **377**, 1156–1167 (2017).
21. Smith, G. D. & Ebrahim, S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* **32**, 1–22 (2003).
22. Smith, G. D. et al. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. *PLoS Med.* **4**, e352 (2007).
23. Tyrrell, J. et al. Genetic evidence for causal relationships between maternal obesity-related traits and birth weight. *J. Am. Med. Assoc.* **315**, 1129–1140 (2016).
24. Pierce, B. L. & Burgess, S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am. J. Epidemiol.* **178**, 1177–1184 (2013).
25. Walton, A. & Hammond, J. The maternal effects on growth and conformation in shire horse–Shetland pony crosses. *Proc. R. Soc. Lond. B* **125**, 311–335 (1938).
26. Smith, D. W. et al. Shifting linear growth during infancy: illustration of genetic factors in growth from fetal life through infancy. *J. Pediatr.* **89**, 225–230 (1976).
27. Sorensen, T. et al. Comparison of associations of maternal peri-pregnancy and paternal anthropometrics with child anthropometrics from birth through age 7 y assessed in the Danish National Birth Cohort. *Am. J. Clin. Nutr.* **104**, 389–396 (2016).
28. Hypponen, E., Power, C. & Smith, G. D. Parental growth at different life stages and offspring birthweight: an intergenerational cohort study. *Paediatr. Perinat. Epidemiol.* **18**, 168–177 (2004).
29. Knight, B. et al. Evidence of genetic regulation of fetal longitudinal growth. *Early Hum. Dev.* **81**, 823–831 (2005).
30. Nahum, G. G. & Stanislaw, H. Relationship of paternal factors to birth weight. *J. Reprod. Med.* **48**, 963–968 (2003).
31. Griffiths, L. J., Dezateux, C. & Cole, T. J. Differential parental weight and height contributions to offspring birthweight and weight gain in infancy. *Int. J. Epidemiol.* **36**, 104–107 (2007).
32. Wood, A. R. et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46**, 1173–1186 (2014).
33. Zhang, G. et al. Assessing the causal relationship of maternal height on birth size and gestational age at birth: a Mendelian randomization analysis. *PLoS Med.* **12**, e1001865 (2015).
34. Tyrrell, J. et al. Height, body mass index, and socioeconomic status: Mendelian randomisation study in UK Biobank. *Br. Med. J.* **352**, i582 (2016).
35. Li, X., Redline, S., Zhang, X., Williams, S. & Zhu, X. Height associated variants demonstrate assortative mating in human populations. *Sci. Rep.* **7**, 15689 (2017).
36. Pedersen J. *Diabetes and Pregnancy: Blood Sugar of Newborn Infants.* PhD thesis (Danish Science Press, 1952).
37. Metzger, B. E. et al. Hyperglycemia and adverse pregnancy outcomes. *N. Eng. J. Med.* **358**, 1991–2002 (2008).
38. Crowther, C. A. et al. Effect of treatment of gestational diabetes mellitus on pregnancy outcomes. *N. Eng. J. Med.* **352**, 2477–2486 (2005).
39. Jarvelin, M. R. et al. Early life factors and blood pressure at age 31 years in the 1966 northern Finland birth cohort. *Hypertension* **44**, 838–846 (2004).
40. Tu, Y. K., West, R., Ellison, G. T. & Gilthorpe, M. S. Why evidence for the fetal origins of adult disease might be a statistical artifact: the "reversal paradox" for the relation between birth weight and blood pressure in later life. *Am. J. Epidemiol.* **161**, 27–32 (2005).
41. Huxley, R., Neil, A. & Collins, R. Unravelling the fetal origins hypothesis: is there really an inverse association between birthweight and subsequent blood pressure? *Lancet* **360**, 659–665 (2002).
42. Wang, T. et al. Low birthweight and risk of type 2 diabetes: a Mendelian randomisation study. *Diabetologia* **59**, 1920–1927 (2016).
43. Freathy, R. M. Can genetic evidence help us to understand the fetal origins of type 2 diabetes? *Diabetologia* **59**, 1850–1854 (2016).
44. Zanetti, D. et al. Birthweight, type 2 diabetes mellitus, and cardiovascular disease: addressing the Barker hypothesis with Mendelian randomization. *Circ. Genom. Precis. Med.* **11**, e002054 (2018).
45. Lawlor, D. et al. Using Mendelian randomization to determine causal effects of maternal pregnancy (intrauterine) exposures on offspring outcomes: sources of bias and methods for assessing them. *Wellcome Open Res.* **2**, 11 (2017).

Acknowledgements

Full acknowledgements and supporting grant details can be found in the Supplementary Information.

Author contributions

The central analysis and writing team comprised N.M.W., R.N.B., M.H., F.R.D., K.K.O., M.I.M., J.R.B.P., D.M.E. and R.M.F. Statistical analysis was performed by N.M.W., R.N.B., M.H., F.R.D., Ø.H., C.Lau., J.B., S.P., K.H., B.F., A.R.W., A.Mah., J.T., N.R.R., N.W.R., Z.Q., G-H.M., M.Vau., M.N., T.M.S., M.H.Z., J.P.B., N.G., M.N.K., R.L.-G., F.G., T.S.A., L.P., R.R., V.H., J.-J.H., L.-P.L., A.C., S.M., D.L.C., Y.W., E.T., C.A.W., C.T.H., N.V.-T., P.K.J., J.N.P., I.N., R.M., N.P., E.M.v.L., R.J., V.L., R.C.R., A.E., S.J.B., W.A., J.A.M., K.L.L., C.A., G.Z., L.J.M., J.Heik., A.H.C.v.K., B.D.C.v.S., K.J.G., N.R.v.Z., C.M.-G., Z.K., S.D., H.M., E.V.R.A., M.Mur., S.B.-G., D.M.H., J.M.Mer., K.E.S., P.A.L., S.E.M., B.M.S., J.-F.C., K.Pan., F.S., D.T., I.P., M.A.T., H.Y., K.S.R., S.E.J., P.-R.L., A.Mur., M.N.W., E.Z., G.V.D., Y.-Y.T., M.G.H., K.L.M., J.F.F., D.M.S., N.J.T., A.P.M., D.A.L., J.R.B.P., D.M.E. and R.M.F. Genotyping was performed by F.R.D., Ø.H., T.M.S., M.H.Z., N.G., R.L.-G., L.P., J.-J.H., L.-P.L., J.W.H., X.E., L.M., L.B., C.S.M., C.Lan., J.L., R.A.S., J.H.Z., G.H., S.M.R., A.J.B., J.F.-T., C.M.-G., H.G.d.H., F.R.R., Z.K., P.M.-V., H.M., E.V.R.A., M.Bus., M.A., P.K., M.Stu., T.A.L., C.M.v.D., A.K., E.Z., S.-M.S., G.W.M., H.C., J.F.W., T.G.M.V., C.E.P., E.E.W., T.D.S., T.L., P.V., H.B., K.B., J.C.M., F.R., J.F.F., T.H., O.P., A.G.U., M.-R.J., W.L.L., G.D.S., N.J.T., N.J.W., H.H., S.F.A.G., T.M.F., D.A.L., P.R.N., K.K.O., M.I.M., J.R.B.P., D.M.E. and R.M.F. Sample collection and phenotyping were performed by F.R.D., B.F., C.J.M., J.C., J.P.B., M.N.K., R.L.-G., F.G., R.R., I.N., H.M.I., J.W.H., L.S.-M., C.R., B.H., C.L.R., M.Kog., L.C., M.-F.H., C.S.M., F.D.M., C.Lan, J.L., R.A.S., J.H.Z., S.M.R., C.M.-G., H.G.d.H., Z.K., P.M.-V., S.D., G.W., M.M.-N., M.Sta., C.E.F., C.T., C.E.M.v.B., M.Bus., D.M.H., A.L., B.A.K., M.Bar., J.S., R.K.V., S.M.W., B.L.C., A.T., K.F.M., A.-M.E., T.A.L., A.K., H.N., K.Pah., O.T.R., B.J., G.V.D., S.-M.S., G.W.M., J.F.W., T.G.M.V., M.Vri., J.-C.H., L.J.B., C.E.P., L.S.A., J.B.B., J.G.E., E.E.W., A.T.H., T.D.S., M.Käh., J.S.V., T.L., P.V., H.B., K.B., M.Mel., E.A.N., D.O.M.-K., J.F.F., V.W.V.J., C.Pis., A.A.V., M.-R.J., C.Pow., E.H., W.L.L., G.D.S., N.J.W., H.H., S.F.A.G., D.A.L., K.K.O., M.I.M. and J.R.B.P. The study designers and principal investigators included J.P.B., I.N., H.M.I., L.S.-M., X.E., B.H., J.M.Mur., M.Kog., L.C., M.-F.H., F.D.M., M.A., A.T., M.Stu., K.F.M., A.-M.E., T.A.L., C.M.v.D., W.K., A.K., H.N., K.Pah., O.T.R., B.J., E.Z., G.V.D., Y.-Y.T., S.-M.S., G.W.M., H.C., J.F.W., T.G.M.V., M.Vri., E.J.C.N.d.G., H.N.K., J.-C.H., L.J.B., C.E.P., J.Hein., L.S.A., J.B.B., K.L.M., J.G.E., E.E.W., A.T.H., T.D.S., M.Käh., J.S.V., T.L., D.I.B., S.S., P.V., T.I.A.S., H.B., K.B., J.C.M., M.Mel., E.A.N., D.O.M.-K., F.R., A.H., J.F.F., V.W.V.J., T.H., C.Pis., A.A.V., O.P., A.G.U., M.-R.J., C.Pow., E.H., W.L.L., N.J.T., A.P.M., N.J.W., H.H., S.F.A.G., T.M.F., D.A.L., P.R.N., S.J., K.K.O., M.I.M., J.R.B.P. and R.M.F.

Competing interests

A.A.V. is an employee of AstraZeneca. S.F.A.G. has received support from GlaxoSmithKline for research that is not related to the study presented in this paper. D.A.L. has received support from Medtronic and Roche Diagnostics for biomarker research that is not related to the study presented in this paper. M.I.M. serves on advisory panels for Pfizer, Novo Nordisk and Zoe Global, has received honoraria from Merck, Pfizer, Novo Nordisk and Eli Lilly, has stock options in Zoe Global, and has received research funding from AbbVie, AstraZeneca, Boehringer Ingelheim, Eli Lilly, Janssen, Merck, Novo Nordisk, Pfizer, Roche, Sanofi–Aventis, Servier and Takeda.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41588-019-0403-1.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to D.M.E. or R.M.F.

Nicole M. Warrington[1,169], Robin N. Beaumont[2,169], Momoko Horikoshi[3,4,5,169], Felix R. Day[6,169], Øyvind Helgeland[7,8,9,169], Charles Laurin[10,11], Jonas Bacelis[12], Shouneng Peng[13,14], Ke Hao[13,14], Bjarke Feenstra[15], Andrew R. Wood[2], Anubha Mahajan[4,5], Jessica Tyrrell[2,16], Neil R. Robertson[4,5], N. William Rayner[4,5,17], Zhen Qiao[1], Gunn-Helen Moen[18,19], Marc Vaudel[7], Carmen J. Marsit[20], Jia Chen[21], Michael Nodzenski[22], Theresia M. Schnurr[23], Mohammad H. Zafarmand[24,25], Jonathan P. Bradfield[26,27], Niels Grarup[23], Marjolein N. Kooijman[28,29,30], Ruifang Li-Gao[31], Frank Geller[15], Tarunveer S. Ahluwalia[23,32,33], Lavinia Paternoster[10], Rico Rueedi[34,35], Ville Huikari[36], Jouke-Jan Hottenga[37,38,39], Leo-Pekka Lyytikäinen[40,41], Alana Cavadino[42,43], Sarah Metrustry[44], Diana L. Cousminer[45,46], Ying Wu[47], Elisabeth Thiering[48,49], Carol A. Wang[50], Christian T. Have[23], Natalia Vilor-Tejedor[51,52], Peter K. Joshi[53], Jodie N. Painter[54], Ioanna Ntalla[55], Ronny Myhre[56], Niina Pitkänen[57], Elisabeth M. van Leeuwen[29], Raimo Joro[58], Vasiliki Lagou[4,59,60], Rebecca C. Richmond[10,11], Ana Espinosa[61,62,63,64], Sheila J. Barton[65], Hazel M. Inskip[65,66], John W. Holloway[67], Loreto Santa-Marina[63,68,69], Xavier Estivill[70,71], Wei Ang[72], Julie A. Marsh[72], Christoph Reichetzeder[73], Letizia Marullo[74], Berthold Hocher[75,76], Kathryn L. Lunetta[77,78], Joanne M. Murabito[78,79], Caroline L. Relton[10,11], Manolis Kogevinas[61,62,63,64], Leda Chatzi[80], Catherine Allard[81], Luigi Bouchard[81,82,83], Marie-France Hivert[84,85,86], Ge Zhang[87,88,89], Louis J. Muglia[87,88,89], Jani Heikkinen[90], EGG Consortium[171], Camilla S. Morgen[92], Antoine H. C. van Kampen[93,94], Barbera D. C. van Schaik[93], Frank D. Mentch[26], Claudia Langenberg[6], Jian'an Luan[6], Robert A. Scott[6], Jing Hua Zhao[6], Gibran Hemani[10,11], Susan M. Ring[10,11], Amanda J. Bennett[5], Kyle J. Gaulton[4,95], Juan Fernandez-Tajes[4], Natalie R. van Zuydam[4,5], Carolina Medina-Gomez[28,29,96], Hugoline G. de Haan[31], Frits R. Rosendaal[31], Zoltán Kutalik[35,97], Pedro Marques-Vidal[98], Shikta Das[99], Gonneke Willemsen[37,38,39], Hamdi Mbarek[37,38,39,100], Martina Müller-Nurasyid[101,102,103], Marie Standl[48], Emil V. R. Appel[23], Cilius E. Fonvig[23,104,105], Caecilie Trier[23,104], Catharina E. M. van Beijsterveldt[38], Mario Murcia[63,106], Mariona Bustamante[61,62,63], Sílvia Bonas-Guarch[107], David M. Hougaard[108], Josep M. Mercader[107,109,110], Allan Linneberg[111,112], Katharina E. Schraut[53], Penelope A. Lind[54], Sarah E. Medland[54], Beverley M. Shields[113], Bridget A. Knight[113], Jin-Fang Chai[114], Kalliope Panoutsopoulou[17], Meike Bartels[37,38,39], Friman Sánchez[107,115], Jakob Stokholm[32], David Torrents[107,116], Rebecca K. Vinding[32], Sara M. Willems[29], Mustafa Atalay[58], Bo L. Chawes[32], Peter Kovacs[117], Inga Prokopenko[4,118], Marcus A. Tuke[2], Hanieh Yaghootkar[2], Katherine S. Ruth[2], Samuel E. Jones[2], Po-Ru Loh[119,120], Anna Murray[2], Michael N. Weedon[2], Anke Tönjes[121], Michael Stumvoll[117,121], Kim F. Michaelsen[122], Aino-Maija Eloranta[58], Timo A. Lakka[58,123,124], Cornelia M. van Duijn[29], Wieland Kiess[125], Antje Körner[117,125], Harri Niinikoski[126,127], Katja Pahkala[57,128], Olli T. Raitakari[57,129], Bo Jacobsson[9,12], Eleftheria Zeggini[17,130], George V. Dedoussis[131], Yik-Ying Teo[114,132,133], Seang-Mei Saw[114,134], Grant W. Montgomery[54], Harry Campbell[53], James F. Wilson[53,135], Tanja G. M. Vrijkotte[24], Martine Vrijheid[61,62,63], Eco J. C. N. de Geus[37,38,39], M. Geoffrey Hayes[136], Haja N. Kadarmideen[137], Jens-Christian Holm[23,104], Lawrence J. Beilin[138], Craig E. Pennell[50], Joachim Heinrich[48,139], Linda S. Adair[140], Judith B. Borja[141,142], Karen L. Mohlke[47], Johan G. Eriksson[143,144,145], Elisabeth E. Widén[90], Andrew T. Hattersley[2,113], Tim D. Spector[44], Mika Kähönen[146,147], Jorma S. Viikari[148,149], Terho Lehtimäki[40,41], Dorret I. Boomsma[37,38,39,100], Sylvain Sebert[36,150,151,152], Peter Vollenweider[98], Thorkild I. A. Sørensen[10,23,92], Hans Bisgaard[32], Klaus Bønnelykke[32], Jeffrey C. Murray[153], Mads Melbye[15,154], Ellen A. Nohr[155], Dennis O. Mook-Kanamori[31,156],

Fernando Rivadeneira[28,29,96], Albert Hofman[29], Janine F. Felix[28,29,30], Vincent W. V. Jaddoe[28,29,30], Torben Hansen[23], Charlotta Pisinger[157], Allan A. Vaag[111,158], Oluf Pedersen[23], André G. Uitterlinden[28,29,96], Marjo-Riitta Järvelin[36,150,151,159,160], Christine Power[43], Elina Hyppönen[43,161,162], Denise M. Scholtens[22], William L. Lowe Jr[136], George Davey Smith[10,11,163], Nicholas J. Timpson[10,11], Andrew P. Morris[4,164,165], Nicholas J. Wareham[6], Hakon Hakonarson[26,45,166], Struan F. A. Grant[26,45,46,166], Timothy M. Frayling[2], Debbie A. Lawlor[10,11,163], Pål R. Njølstad[7,8], Stefan Johansson[7,167], Ken K. Ong[6,168,170], Mark I. McCarthy[4,5,169,170], John R. B. Perry[6,170], David M. Evans[1,10,11,170]* and Rachel M. Freathy[2,10,170]*

---

[1]University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Queensland, Australia. [2]Institute of Biomedical and Clinical Science, College of Medicine and Health, University of Exeter, Royal Devon and Exeter Hospital, Exeter, UK. [3]RIKEN Centre for Integrative Medical Sciences, Laboratory for Endocrinology, Metabolism and Kidney Diseases, Yokohama, Japan. [4]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK. [5]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK. [6]MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, UK. [7]KG Jebsen Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen, Norway. [8]Department of Pediatrics, Haukeland University Hospital, Bergen, Norway. [9]Department of Genetics and Bioinformatics, Domain of Health Data and Digitalisation, Norwegian Institute of Public Health, Oslo, Norway. [10]Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK. [11]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. [12]Department of Obstetrics and Gynecology, Sahlgrenska University Hospital Östra, Gothenburg, Sweden. [13]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [14]Icahn Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [15]Department of Epidemiology Research, Statens Serum Institute, Copenhagen, Denmark. [16]European Centre for Environment and Human Health, University of Exeter, Truro, UK. [17]Wellcome Sanger Institute, Hinxton, UK. [18]Department of Endocrinology, Morbid Obesity and Preventive Medicine, Oslo University Hospital, Oslo, Norway. [19]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway. [20]Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA. [21]Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [22]Department of Preventive Medicine, Division of Biostatistics, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA. [23]Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [24]Department of Public Health, Amsterdam Public Health Research Institute, Amsterdam UMC, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands. [25]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam Public Health Research Institute, Amsterdam UMC, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands. [26]Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, USA. [27]Quantinuum Research, San Diego, CA, USA. [28]The Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands. [29]Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands. [30]Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands. [31]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands. [32]Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, Denmark. [33]Steno Diabetes Center Copenhagen, Gentofte, Denmark. [34]Department of Computational Biology, University of Lausanne, Lausanne, Switzerland. [35]Swiss Institute of Bioinformatics, Lausanne, Switzerland. [36]Center for Life Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Finland. [37]Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands. [38]Netherlands Twin Register, Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands. [39]Amsterdam Public Health, Amsterdam, the Netherlands. [40]Department of Clinical Chemistry, Fimlab Laboratories, Tampere, Finland. [41]Department of Clinical Chemistry, Finnish Cardiovascular Research Center–Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland. [42]Section of Epidemiology and Biostatistics, School of Population Health, University of Auckland, Auckland, New Zealand. [43]Population, Policy and Practice, UCL Great Ormond Street Institute of Child Health, University College London, London, UK. [44]Department of Twin Research, King's College London, St. Thomas' Hospital, London, UK. [45]Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA, USA. [46]Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [47]Department of Genetics, University of North Carolina, Chapel Hill, NC, USA. [48]Institute of Epidemiology, Helmholtz Zentrum München German Research Center for Environmental Health, Neuherberg, Germany. [49]Division of Metabolic and Nutritional Medicine, Dr. von Hauner Children's Hospital, University of Munich Medical Center, Munich, Germany. [50]School of Medicine and Public Health, Faculty of Medicine and Health, The University of Newcastle, Newcastle, New South Wales, Australia. [51]Center for Genomic Regulation, Barcelona Institute of Science and Technology, Barcelona, Spain. [52]Barcelonabeta Brain Research Center, Pasqual Maragall Foundation, Barcelona, Spain. [53]Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK. [54]QIMR Berghofer Medical Research Institute, Royal Brisbane Hospital, Brisbane, Queensland, Australia. [55]William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [56]Department of Genes and Environment, Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway. [57]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland. [58]Institute of Biomedicine, School of Medicine, University of Eastern Finland, Kuopio, Finland. [59]Department of Neuroscience, Katholieke Universiteit Leuven, Leuven, Belgium. [60]VIB Center for Brain and Disease Research, Leuven, Belgium. [61]ISGlobal, Barcelona Institute for Global Health, Barcelona, Spain. [62]Universitat Pompeu Fabra, Barcelona, Spain. [63]CIBER de Epidemiología y Salud Pública, Madrid, Spain. [64]Hospital del Mar Medical Research Institute, Barcelona, Spain. [65]Medical Research Council Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK. [66]NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, Southampton, UK. [67]Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK. [68]Subdirección de Salud Pública y Adicciones de Gipuzkoa, San Sebastián, Spain. [69]Instituto de Investigación Sanitaria Biodonostia, San Sebastián, Spain. [70]Sidra Medicine Research Department, Sidra Medicine, Doha, Qatar. [71]Genomics Unit, Dexeus Woman's Health, Barcelona, Spain. [72]Division of Obstetrics and Gynaecology, The University of Western Australia, Perth, Western Australia, Australia. [73]Institute of Nutritional Science, University of Potsdam, Nuthetal, Germany. [74]Department of Life Sciences and Biotechnology, Genetic Section, University of Ferrara, Ferrara, Italy. [75]Fifth Department of Medicine, University Medical Centre Mannheim, University of Heidelberg, Heidelberg, Germany. [76]Reproductive and Genetic Hospital of CITIC-Xiangya, Changsha, China. [77]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA. [78]Framingham Heart Study, Framingham, MA, USA. [79]Section of General Internal Medicine, Department of Medicine, Boston University School of

Medicine, Boston, MA, USA. [80]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, USA. [81]Centre de recherche, Centre Hospitalier Universitaire de Sherbrooke, Sherbrooke, Quebec, Canada. [82]ECOGENE-21 and Lipid Clinic, Chicoutimi Hospital, Saguenay, Quebec, Canada. [83]Department of Biochemistry, Université de Sherbrooke, Sherbrooke, Quebec, Canada. [84]Department of Population Medicine, Harvard Pilgrim Health Care Institute, Harvard Medical School, Boston, MA, USA. [85]Diabetes Center, Massachusetts General Hospital, Boston, MA, USA. [86]Department of Medicine, Universite de Sherbrooke, Sherbooke, Quebec, Canada. [87]Human Genetics Division, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. [88]Center for Prevention of Preterm Birth, Perinatal Institute, Cincinnati Children's Hospital Medical Center, Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, USA. [89]March of Dimes Prematurity Research Center Ohio Collaborative, Cincinnati, OH, USA. [90]Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland. [92]Department of Public Health, Section of Epidemiology, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [93]Bioinformatics Laboratory, Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam Public Health Research Institute, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands. [94]Biosystems Data Analysis, Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, the Netherlands. [95]Department of Pediatrics, University of California San Diego, La Jolla, CA, USA. [96]Department of Internal Medicine, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands. [97]Institute of Social and Preventive Medicine, Lausanne University Hospital, Lausanne, Switzerland. [98]Department of Medicine, Internal Medicine, Lausanne University Hospital, Lausanne, Switzerland. [99]Medical Research Council Unit for Lifelong Health and Ageing at UCL, Institute of Cardiovascular Sciences, University College London, London, UK. [100]Amsterdam Reproduction and Development, Amsterdam, the Netherlands. [101]Institute of Genetic Epidemiology, Helmholtz Zentrum München German Research Center for Environmental Health, Neuherberg, Germany. [102]Department of Internal Medicine I (Cardiology), Hospital of the Ludwig-Maximilians-University Munich, Munich, Germany. [103]Division of Genetic Epidemiology, Institute for Medical Information Processing, Biometry and Epidemiology, Faculty of Medicine, Ludwig-Maximilians-University Munich, Munich, Germany. [104]Children's Obesity Clinic, Department of Pediatrics, Copenhagen University Hospital Holbæk, Holbæk, Denmark. [105]Hans Christian Andersen Children's Hospital, Odense University Hospital, Odense, Denmark. [106]FISABIO–Universitat Jaume I–Universitat de València, Joint Research Unit of Epidemiology and Environmental Health, Valencia, Spain. [107]Joint BSC–CGR–IRB Research Program in Computational Biology, Barcelona Supercomputing Center, Barcelona, Spain. [108]Danish Center for Neonatal Screening, Department for Congenital Disorders, Statens Serum Institut, Copenhagen, Denmark. [109]Programs in Metabolism and Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA. [110]Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA. [111]Department of Clinical Medicine, Faculty of Health and Medical Sciences, Copenhagen University, Copenhagen, Denmark. [112]Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, Frederiksberg, Denmark. [113]NIHR Exeter Clinical Research Facility, University of Exeter College of Medicine and Health and Royal Devon and Exeter NHS Foundation Trust, Exeter, UK. [114]Saw Swee Hock School of Public Health, National University of Singapore, National University Health System, Singapore, Singapore. [115]Computer Sciences Department, Barcelona Supercomputing Center, Barcelona, Spain. [116]Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain. [117]IFB Adiposity Diseases, University of Leipzig, Leipzig, Germany. [118]Section of Genomics of Common Disease, Department of Medicine, Imperial College London, London, UK. [119]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA. [120]Division of Genetics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. [121]Medical Department, University of Leipzig, Leipzig, Germany. [122]Department of Nutrition, Exercise and Sports, Faculty of Science, University of Copenhagen, Copenhagen, Denmark. [123]Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland. [124]Kuopio Research Institute of Exercise Medicine, Kuopio, Finland. [125]Pediatric Research Center, Department of Women's andChild Health, University of Leipzig, Leipzig, Germany. [126]Department of Pediatrics, Turku University Hospital, Turku, Finland. [127]Department of Physiology, University of Turku, Turku, Finland. [128]Paavo Nurmi Centre, Sports and Exercise Medicine Unit, Department of Physical Activity and Health, University of Turku, Turku, Finland. [129]Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland. [130]Institute of Translational Genomics, Helmholtz Zentrum München German Research Center for Environmental Health, Neuherberg, Germany. [131]Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University, Athens, Greece. [132]Department of Statistics and Applied Probability, National University of Singapore, Singapore, Singapore. [133]Life Sciences Institute, National University of Singapore, Singapore, Singapore. [134]Singapore Eye Research Institute, Singapore, Singapore. [135]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK. [136]Department of Medicine, Division of Endocrinology, Metabolism, and Molecular Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA. [137]Quantitative and Systems Genomics Group, Department of Bio and Health Informatics, Technical University of Denmark, Lyngby, Denmark. [138]School of Medicine, Royal Perth Hospital, University of Western Australia, Perth, Western Australia, Australia. [139]Institute and Outpatient Clinic for Occupational Social and Environmental Medicine, Inner City Clinic, University Hospital Munich, Ludwig-Maximilians-University Munich, Munich, Germany. [140]Department of Nutrition, University of North Carolina, Chapel Hill, NC, USA. [141]USC-Office of Population Studies Foundation, University of San Carlos, Cebu City, Philippines. [142]Department of Nutrition and Dietetics, University of San Carlos, Cebu City, Philippines. [143]National Institute for Health and Welfare, Helsinki, Finland. [144]Department of General Practice and Primary Health Care, University of Helsinki and Helsinki University Hospital, Helsinki, Finland. [145]Folkhälsan Research Center, Helsinki, Finland. [146]Department of Clinical Physiology, Tampere University Hospital, Tampere, Finland. [147]Department of Clinical Physiology, Finnish Cardiovascular Research Center–Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland. [148]Division of Medicine, Turku University Hospital, Turku, Finland. [149]Department of Medicine, University of Turku, Turku, Finland. [150]Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, UK. [151]Biocenter Oulu, University of Oulu, Oulu, Finland. [152]Department of Genomics of Complex Diseases, Imperial College London, London, UK. [153]Department of Pediatrics, University of Iowa, Iowa City, IA, USA. [154]Department of Medicine, Stanford School of Medicine, Stanford, CA, USA. [155]Research Unit for Gynaecology and Obstetrics, Institute of Clinical Research, University of Southern Denmark, Odense, Denmark. [156]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, the Netherlands. [157]Research Center for Prevention and Health, Center for Sundhed, Rigshospitalet Glostrup, Copenhagen University, Glostrup, Denmark. [158]Cardiovascular, Renal and Metabolism, Translational Medicine Unit, Early Clinical Development, IMED Biotech Unit, AstraZeneca, Gothenburg, Sweden. [159]Unit of Primary Health Care, Oulu University Hospital, Oulu, Finland. [160]Department of Life Sciences, College of Health and Life Sciences, Brunel University London, London, UK. [161]Australian Centre for Precision Health, University of South Australia Cancer Research Institute, Adelaide, South Australia, Australia. [162]South Australian Health and Medical Research Institute, Adelaide, South Australia, Australia. [163]NIHR Bristol Biomedical Research Centre, Bristol, UK. [164]Department of Biostatistics, University of Liverpool, Liverpool, UK. [165]Estonian Genome Center, University of Tartu, Tartu, Estonia. [166]Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [167]Department of Medical Genetics, Haukeland University Hospital, Bergen, Norway. [168]Department of Paediatrics, University of Cambridge, Cambridge, UK. [169]NIHR Oxford Biomedical Research Centre, Oxford University Hospitals NHS Foundation Trust, John Radcliffe Hospital, Oxford, UK. [169]These authors contributed equally: Nicole M. Warrington, Robin N. Beaumont, Momoko Horikoshi, Felix R. Day, Øyvind Helgeland. [170]These authors jointly supervised this work: Ken K. Ong, Mark I. McCarthy, John R. B. Perry, David M. Evans, Rachel M. Freathy. [171]A list of members and affiliations appears in the Supplementary Note. *e-mail: d.evans1@uq.edu.au; r.freathy@ex.ac.uk

Methods

Ethics statement. All human research was approved by the relevant institutional review boards and conducted according to the Declaration of Helsinki. Participants of all studies provided written informed consent. The UK Biobank has approval from the North West Multi-Centre Research Ethics Committee, which covers the United Kingdom. Ethical approval for the Avon Longitudinal Study of Parents and Children (ALSPAC) study was obtained by the ALSPAC Ethics and Law Committee and local research ethics committees. Ethical approval for the Exeter Family Study of Childhood Health (EFSOCH) study was given by the North and East Devon Local Ethics Committee. Approval for access to data and biological material relating to MoBa–HARVEST was granted by the Scientific Management Committee of MoBa and the Regional Committee for Medical and Health Research Ethics.

Statistical tests. Details of the statistical tests used in the various analyses are reported under the appropriate headings below. All tests were two sided, unless otherwise stated.

GWAS of own birth weight. *European ancestry meta-analysis.* The European ancestry GWAS meta-analysis of own birth weight consisted of two components (Supplementary Fig. 1): (1) 80,745 individuals from 35 studies participating in the EGG Consortium from Europe, the United States and Australia; and (2) 217,397 individuals of white European origin from the UK Biobank (see Supplementary Note for details on phenotype preparation and GWAS analyses). We combined the summary statistics from the EGG meta-analysis with the UK Biobank summary statistics using a fixed-effects meta-analysis in GWAMA[46] (maximum $n = 297{,}142$). Variants failing GWAS quality-control filters, reported in <50% of the total sample size in the EGG component, or with MAF < 0.1% were excluded. We also performed a fixed-effects meta-analysis of the association summary statistics for 16,095 directly genotyped SNPs on the X chromosome from the UK Biobank and the EGG meta-analysis using GWAMA[46] (maximum $n = 270{,}929$). Genome-wide significance was defined as $P < 6.6 \times 10^{-9}$, as calculated by Kemp et al.[47], which was similar to the thresholds calculated using permutations by Jones et al.[48]. A locus was defined as one or more SNPs reaching genome-wide significance within a region of the genome; two genome-wide-significant SNPs were defined as belonging to two separate loci if the distance between them was ≥500 kb. The lead SNP within each locus was the SNP with the smallest *P* value.

Several sensitivity analyses were conducted to confirm the quality of our meta-analysis results (see Supplementary Note for details). Univariate linkage disequilibrium score regression[49] estimated the genomic inflation as 1.08, indicating that the majority of genome-wide inflation of the test statistics was due to polygenicity. To assess the impact of this inflation, we recalculated the association *P* values after adjusting the test statistics for the linkage disequilibrium score regression intercept (Supplementary Table 5).

COJO analysis to identify additional independent signals. Approximate COJO analysis[13] was performed in GCTA[50] using the European ancestry meta-analysis summary statistics to identify independent association signals attaining genome-wide significance ($P < 6.6 \times 10^{-9}$). The linkage disequilibrium reference panel was made up of 344,246 unrelated UK Biobank participants defined by the UK Biobank as having British ancestry, and SNPs were restricted to those present in the Haplotype Reference Consortium (HRC) reference panel. At each locus, only SNPs labeled by GCTA as 'independent' and not in linkage disequilibrium with the original lead SNP ($r^2 < 0.05$) were listed as secondary SNPs.

Trans-ethnic meta-analysis. To identify any further independent birth-weight-associated SNPs, we conducted a trans-ethnic meta-analysis combining three components (Supplementary Fig. 1): (1) 80,745 individuals from the European ancestry component within the EGG consortium; (2) 12,948 individuals from nine studies within the EGG consortium from diverse ancestry groups: African American, Afro-Caribbean, Mexican, Chinese, Thai, Filipino, Surinamese, Turkish and Moroccan; and (3) 227,530 individuals of all ancestries from the UK Biobank. The same strategy and variant filtering criteria were applied as in the European meta-analysis (Supplementary Fig. 1). Univariate linkage disequilibrium score regression[49] estimated the genomic inflation as 1.08. *P* values after adjustment of the test statistics for the linkage disequilibrium score regression intercept are presented in Supplementary Table 5.

GWAS of offspring birth weight. *European ancestry meta-analysis.* The European ancestry GWAS meta-analysis of offspring birth weight consisted of three components (Supplementary Fig. 2): (1) 12,319 individuals from 10 GWASs in the EGG consortium of European descent imputed up to the HapMap 2 reference panel; (2) 7,542 individuals from 2 GWASs in the EGG consortium of European descent imputed up to the HRC panel; and (3) 190,406 individuals of white European origin from the UK Biobank (see Supplementary Note for details on phenotype preparation and GWAS analyses). We conducted a European ancestry fixed-effects meta-analysis to combine the association summary statistics from the three components using GWAMA[46] (max $n = 210{,}267$). We also performed a fixed-effects meta-analysis of the association summary statistics for 18,137 directly genotyped SNPs on the X chromosome from the UK Biobank and EGG meta-analysis using GWAMA[46] (max $n = 197{,}093$). The same strategy and variant filtering criteria were applied as in the meta-analysis of own birth weight, and the same sensitivity analyses were conducted (Supplementary Note).

Univariate linkage disequilibrium score regression[49] estimated the genomic inflation as 1.05. We recalculated the *P* values after adjusting the test statistics for this linkage disequilibrium score intercept (Supplementary Table 5).

Approximate COJO analysis to identify additional independent signals. We performed approximate COJO analysis[13] using the European ancestry meta-analysis summary statistics of offspring birth weight. We used the same reference panel as in the own-birth-weight analysis and listed any secondary 'independent' SNPs associated with offspring birth weight.

Trans-ethnic meta-analysis. We conducted a trans-ethnic meta-analysis combining three components (Supplementary Fig. 2): (1) 12,319 individuals from 10 GWASs in the EGG consortium of European descent imputed up to the HapMap 2 reference panel; (2) 7,542 individuals from 2 GWASs in the EGG consortium of European descent imputed up to the HRC panel; and (3) 210,208 individuals of all ancestries from the UK Biobank. The same strategy and variant filtering criteria were applied as in the European meta-analysis (Supplementary Fig. 2), and the same sensitivity analyses were conducted (Supplementary Note). Univariate linkage disequilibrium score regression[49] estimated the genomic inflation as 1.04, and the recalculated *P* values after adjusting the test statistics for this linkage disequilibrium score intercept are presented in Supplementary Table 5.

Colocalization methods. For each signal where we identified different lead SNPs in the GWASs of own birth weight and offspring birth weight, we performed colocalization analysis using the 'coloc' R package[51]. For each signal, we input the regression coefficients, their variances, and SNP MAFs for all SNPs 500 kb up- and downstream of the lead SNP from the European meta-analysis. We used the coloc.abf() function, with its default parameters, to calculate posterior probabilities that the own birth weight and offspring birth weight lead SNPs were independent (H_3) or shared the same associated variant (H_4). We called variants the same signal if the H_4 posterior probability was greater than 0.50, and different signals if the H_3 posterior probability was greater than 0.50.

Estimation of genetic variance explained. First, we estimated the proportion of birth weight variance explained by fetal genotypes, maternal genotypes and the covariance between the two at the 190 genome-wide-significant signals in the Norwegian Mother and Child Cohort Study (MoBa–HARVEST; https://www.fhi.no/en/studies/moba/). This sample was independent of samples contributing to the discovery meta-analyses, apart from a small potential overlap with mothers from the MoBa 2008 sample that was included in the GWAS of offspring birth weight (<0.07% of the meta-analysis sample). For the 19 signals that had different maternal and fetal lead SNPs, the fetal SNP (and not the maternal SNP) was used in the analysis to avoid collinearity in the model. We excluded multiple births, babies of non-European descent, babies born before 37 weeks of gestation, and babies born with a congenital anomaly or still-born. Birth weight was *z*-score transformed and all models were adjusted for sex, gestational duration and the first four ancestry-informative principal components. We conducted a linear regression analysis in R[52] using 13,934 mother–offspring pairs where offspring birth weight was regressed on the maternal and fetal genotypes at all 190 SNPs simultaneously. The proportion of variance explained by fetal genotypes at the 190 genome-wide-significant signals was calculated as:

$$\sum_{i=1}^{190} \frac{2p_i(1-p_i)\hat{\beta}_{\mathrm{f}_i}^2}{\mathrm{var}(\mathrm{BW})}$$

Where p_i is the effect allele frequency of the *i*th SNP, $\hat{\beta}_{\mathrm{f}_i}$ is the regression coefficient for the effect of the offspring's genotype at the *i*th SNP on offspring birth weight, and var(BW) is the variance of offspring birth weight (which is approximately 1 as birth weight was *z*-score transformed). A similar formula was used to calculate the variance explained by maternal genotypes, using:

$$\sum_{i=1}^{190} \frac{2p_i(1-p_i)\hat{\beta}_{\mathrm{m}_i}^2}{\mathrm{var}(\mathrm{BW})}$$

Where $\hat{\beta}_{\mathrm{m}_i}$ is the regression coefficient for the effect of the maternal genotype at the *i*th SNP on offspring birth weight. The following formula was used to calculate twice the covariance:

$$\sum_{i=1}^{190} \frac{2p_i(1-p_i)\hat{\beta}_{\mathrm{f}_i}\hat{\beta}_{\mathrm{m}_i}}{\mathrm{var}(\mathrm{BW})}$$

Second, we used maternal GCTA[11] to estimate the proportion of variance in birth weight explained by genome-wide SNPs, or SNPs they tag, in the fetal genome or maternal genome, the covariance between the two and environmental factors in MoBa–HARVEST. The same phenotype was used as in the previous analysis and the model was adjusted for sex and gestational duration. Mothers or offspring were excluded if they were related to others in the sample, using a genetic relationship cut-off of 0.025, leaving $n = 7{,}910$ mother–offspring pairs available for analysis.

Identifying eQTL-linked genes. We used FUSION[53] with the v6p release of the GTEx data[14] to identify eQTL-linked genes. FUSION incorporates information from gene expression and GWAS data to translate evidence of association with a phenotype from the SNP level to the gene. Only gene-level results from the adjusted model were taken forward for consideration. Each of the genes implicated by this analysis survived multiple testing correction (Bonferroni corrected $P < 6 \times 10^{-7}$, after adjusting for 44 tissues) and were independent from other proximal genes tested in a joint model.

Placenta eQTL look ups. We annotated genome-wide-significant birth-weight-associated SNPs with gene expression data (200/209 SNPs available) from European ancestry placental samples in the Rhode Island Child Health Study[15] ($n = 123$ with a fetal genotype, including 71 with a birth weight appropriate for gestational age, 15 small for gestational age and 37 large for gestational age). SNPs were annotated if they had a genome-wide empirical false discovery rate < 0.01 for association with one or more transcripts, and $r^2 > 0.7$ with a lead eQTL SNP.

TAD pathways. TAD pathway analysis was performed using software described in Way et al.[16]. Briefly, the software uses publicly available TAD boundaries, identified in human embryonic stem cells and fibroblasts using a hidden Markov model[17], to prioritize candidate genes at GWAS SNPs. These TAD boundaries are stable across different cell types and can be used to identify genomic regions where non-coding causal variants are most likely to impact tissue-independent function.

Structural equation model for estimating adjusted maternal and fetal effects of the genome-wide-significant variants. The SEM approach used to estimate adjusted maternal and fetal effects has been described elsewhere[12] (for additional details, see Supplementary Note).

The SEM was fit to data from 146 genome-wide-significant lead fetal SNPs and 72 lead maternal SNPs from the GWAS meta-analysis (Supplementary Fig. 4). To identify a subset of unrelated individuals in the UK Biobank (as the SEM cannot easily account for relatedness), we generated a genetic relationship matrix using the GCTA software package[50] (version 1.90.2) and excluded one of every pair of related individuals with a genetic relationship greater than 9.375%. After the same exclusions were made as in the GWAS, 85,518 unrelated individuals of European ancestry with their own and their offspring's birth weight, 98,235 individuals with their own birth weight only, and 73,981 individuals with their offspring's birth weight only were available for analysis. We fit linear regression models to birth weight and offspring birth weight in this subset of unrelated, European ancestry individuals adjusting for sex (own birth weight only), assessment center and the top 40 ancestry-informative principal components provided by the UK Biobank, to account for any remaining population substructure. The residuals from these regression models were *z*-score transformed for analysis. Because we included the summary statistics from the meta-analysis of the EGG studies, rather than the individual-level data, we were unable to account for the small of subset individuals who contributed to both the own-birth-weight and offspring-birth-weight GWAS meta-analyses. Based on the results from simulations (not shown), we expect that this non-independence will result in very slightly smaller standard errors and an increased type 1 error rate, particularly for the fetal effect, which is estimated from a larger sample size than was available to estimate the maternal effect. Therefore, we conducted a sensitivity analysis that first excluded EGG studies from the meta-analysis of own birth weight that contributed to both GWAS meta-analyses of own and offspring birth weight (for example, ALSPAC), and then refitted the non-overlapping data in the SEM; these results are presented in Supplementary Table 19. For SNPs identified on the X chromosome, we fit a slightly different SEM due to males having double the expected genetic variance at X-linked loci compared with females. We did not incorporate summary statistics from the EGG Consortium as the GWAS results were not stratified according to sex (additional details on the X chromosome analysis are provided in the Supplementary Note and Supplementary Fig. 6b).

We used the estimates from the SEM to classify the lead SNPs into the following five categories. (1) Fetal only: the 95% CI surrounding the fetal effect estimate does not overlap zero and does not overlap the 95% CI around the maternal effect estimate. Additionally, the 95% CI surrounding the maternal effect estimate overlaps zero. (2) Maternal only: the 95% CI surrounding the maternal effect estimate does not overlap zero and does not overlap the 95% CI around the fetal effect estimate. Additionally, the 95% CI surrounding the fetal effect estimate overlaps zero. (3) Fetal and maternal, with effects going in the same direction: the 95% CIs around both the maternal and fetal effect estimates do not overlap zero, and their effect is in the same direction. (4) Fetal and maternal, with effects going in opposite directions: the 95% CIs around both the maternal and fetal effect estimates do not overlap zero, and their effects are in opposite directions. (5) Unclassified: SNPs do not fall into any of these categories; therefore, the 95% CIs around the maternal and fetal effect estimates overlap, and at least one overlaps zero.

Meta-analysis of maternal and fetal effects from a conditional regression analysis in mother–offspring pairs. We conducted conditional association analyses for all 209 lead SNPs in 18,873 mother–offspring pairs from 3 studies (MoBa–HARVEST, ALSPAC and EFSOCH), adjusting for both maternal and offspring genotype, and combined the summary statistics for each SNP in a fixed-effects meta-analysis using METAL[54]. We compared the estimates of the maternal and fetal effects of this meta-analysis with the SEM-adjusted maternal and fetal effects using a heterogeneity test (Supplementary Table 6).

Approximation of the SEM for genome-wide analyses. The SEM is computationally intensive to fit, making it difficult to run on all SNPs across the genome. Therefore, we developed an approximation of the SEM using a linear transformation and ordinary least squares linear regression, which we refer to as WLM-adjusted analyses. The full details of the derivation are provided in the Supplementary Note. Briefly, from ordinary least squares regression, we know that the estimated fetal effect size from the GWAS of own birth weight, $\hat{\beta}_{\mathrm{f_{unadj}}}$, is calculated by dividing the sample covariance between birth weight and SNP by the sample variance of the SNP. Similarly, the estimated maternal effect from the GWAS of offspring birth weight, $\hat{\beta}_{\mathrm{m_{unadj}}}$, is calculated by dividing the sample covariance between offspring birth weight and SNP by the sample variance of the SNP. It follows that an estimate of the fetal effect adjusted for the maternal genotype is (see Supplementary Note for full derivation):

$$\hat{\beta}_{\mathrm{f_{adj}}} = -\frac{2}{3}\hat{\beta}_{\mathrm{m_{unadj}}} + \frac{4}{3}\hat{\beta}_{\mathrm{f_{unadj}}}$$

and an estimate of the maternal effect adjusted for the fetal genotype is:

$$\hat{\beta}_{\mathrm{m_{adj}}} = \frac{4}{3}\hat{\beta}_{\mathrm{m_{unadj}}} - \frac{2}{3}\hat{\beta}_{\mathrm{f_{unadj}}}$$

If the model is truly linear, the same estimates can be obtained by transforming the reported birth weights rather than the regression coefficients[55]. See the Supplementary Note and Supplementary Fig. 18 for a flow diagram of the full analysis pipeline. A comparison of the results using this WLM method and the full SEM for the lead SNPs is presented in Supplementary Fig. 9.

Gene expression integration. To identify which tissue types were most relevant to genes involved in birth weight, we applied linkage disequilibrium score regression to specifically expressed genes ('LDSC-SEG')[19]. We used the summary statistics from the GWAS meta-analysis of own and offspring birth weight and the WLM-adjusted meta-analyses. Briefly, the method takes each tissue, ranking genes by a *t* statistic for differential expression, using sex and age as covariates, and excluding all samples in related tissues. It then takes the top 10% of ranked genes and makes a genome annotation including these genes (exons and introns) plus 100 kb on either side. Finally, it uses stratified linkage disequilibrium score regression to estimate the contribution of this annotation to per-SNP birth weight heritability, adjusting for all categories in the baseline model. We computed significance using a block jackknife over SNPs, and corrected for the number of tissues tested.

Gene-set enrichment analysis (MAGENTA). Pathway-based associations using summary statistics from the GWAS meta-analysis of own and offspring birth weight and WLM-adjusted meta-analysis were tested using MAGENTA[56]. Briefly, the software maps each gene to the SNP with the lowest *P* value within a 110-kb upstream and 40-kb downstream window. This *P* value is corrected for factors such as SNP density and gene size using a regression model. Genes within the HLA region were excluded. The observed number of gene scores within a given pathway with a ranked gene score above a given threshold (ninety-fifth or seventy-fifth percentile) was calculated. This statistic was compared with 1,000,000 randomly permuted pathways of the same size to calculate an empirical *P* value for each pathway. We considered pathways with a false discovery rate < 0.05 to be of interest. The 3,230 biological pathways tested were from the BIOCARTA, Gene Ontology, KEGG, PANTHER and READTOME databases, along with a small number of custom gene sets.

Gestational duration associations. We extracted the 209 lead birth-weight-associated SNPs from the summary statistics provided by 23andMe and published in a recent GWAS of gestational duration[20]. Any birth weight-associated SNP that was also associated with gestational duration ($P < 2.4 \times 10^{-4}$, corrected for 209 tests) was followed up in 13,206 mother–child pairs from the MoBa–HARVEST, ALSPAC and EFSOCH studies. Preterm births (gestational duration < 37 weeks) were removed before analysis, and gestational duration and birth weight were both *z*-score transformed. We conducted linear regression analyses to test the association between maternal or fetal genotype (both unadjusted and adjusted genotype effects) and gestational duration, birth weight or gestational duration adjusted for birth weight. The association analysis results were combined using inverse-variance-weighted meta-analysis. We also combined the unadjusted maternal SNP-gestational duration associations with the 23andMe summary statistics[20] using *P* value-based meta-analysis implemented in METAL[54].

Association between birth-weight-associated SNPs and a variety of traits. We performed GWASs on 78 traits in the UK Biobank using BOLT-LMM in an analogous way to the analysis of own birth weight. Phenotype definitions for the

78 traits are described by Frayling et al.[57]. Association statistics for the 209 lead birth-weight-associated SNPs were then extracted from the results (Supplementary Table 13). Additionally, we searched the NHGRI GWAS (https://www.ebi.ac.uk/gwas/; accessed 16 January 2018) for the 209 lead birth-weight-associated SNPs, or SNPs in high linkage disequilibrium with the 209 lead SNPs ($r^2 > 0.8$), and reported associations with other traits (Supplementary Table 13).

Linkage disequilibrium score regression. Linkage disequilibrium score regression, which has been described in detail elsewhere[18], was used to estimate the genetic correlation between birth weight and a range of traits and diseases. We used LD Hub[58] (ldsc.broadinstitute.org/) to perform the linkage disequilibrium score regression analyses. Due to the different linkage disequilibrium structure across ancestry groups, the summary statistics from the European-only birth weight analyses were uploaded to LD Hub, and genetic correlations were calculated with all available phenotypes. We conducted four separate analyses in LD Hub using (1) GWAS meta-analysis of own birth weight; (2) GWAS meta-analysis of offspring birth weight; (3) WLM-adjusted fetal effect; and (4) WLM-adjusted maternal effect.

To calculate the genetic correlation between the maternal and fetal effect estimates from the unadjusted and WLM-adjusted analyses, and also between gestational duration and the WLM-adjusted maternal and fetal effects, we used the scripts provided by the developer (https://github.com/bulik/ldsc).

Mendelian randomization analyses of maternal and fetal exposures on offspring birth weight. Two-sample Mendelian randomization analyses were performed with own or offspring birth weight as outcomes. The exposures included height, fasting glucose, disposition index of insulin secretion[59], insulin sensitivity, SBP and DBP. The SNP–exposure associations were taken from external studies (Supplementary Table 14). The SNP–outcome associations were taken from the current European GWAS meta-analyses of own birth weight, offspring birth weight, WLM-adjusted fetal effect and WLM-adjusted maternal effect. Two-sample Mendelian randomization regresses the effect sizes of SNP–outcome associations against the effect sizes of SNP–exposure associations, with an inverse-variance weighted analysis, giving similar results to the two-stage least-squares analysis in a single sample[60]. We performed several sensitivity analyses to assess the impact of genetic pleiotropy on the causal estimates including MR-Egger[61], Weighted Median[62] and Penalized Weighted Median[62] (see Supplementary Table 15 for results). Details of the R code for the Mendelian randomization analyses are provided elsewhere[61,62].

Due to the strong negative correlation between estimates of the maternal and fetal genetic effects on birth weight, we conducted simulations to confirm that this correlation did not bias the results of downstream Mendelian randomization analyses; these simulations are described in the Supplementary Note.

Transmitted/non-transmitted allele scores in ALSPAC. Allelic transmission was determined for 4,962 mother–offspring pairs in ALSPAC. First, we converted maternal and fetal genotypes into best-guess genotypes where SNPs of interest had been imputed. Where one or both of the mother–offspring pair were homozygous, allelic transmission was trivial to determine. Where both mother and offspring were heterozygous for the SNP of interest, we used phase imputation generated using SHAPEIT2 (ref. [63]) to examine the haplotypes in the region of the SNP of interest to determine allelic transmission. Weighted allele scores were generated for maternal non-transmitted, shared (maternal transmitted) and paternally inherited fetal alleles for SBP, DBP, fasting glucose, insulin secretion and insulin sensitivity. Associations were tested between the weighted allele scores and birth weight.

Covariance between birth weight and adult traits explained by genotyped SNPs. The genetic and residual covariance between birth weight and several quantitative/disease phenotypes was calculated in the UK Biobank using REML in BOLT-LMM[64]. We included 215,444 individuals of European ancestry with data on own birth weight and 190,406 with data on offspring birth weight. SNPs with MAF $< 1\%$, evidence of deviation from Hardy–Weinberg equilibrium ($P \leq 1 \times 10^{-6}$) or an overall missing rate > 0.015 were excluded, resulting in 524,307 genotyped SNPs for analysis. We then calculated 96% CIs for the proportion of covariance explained by genotyped variants as $\text{gcov}/(\text{gcov} + \text{rcov}) \pm 1.96 \times \text{gcovSE}/\text{abs}(\text{gcov} + \text{rcov})$, where gcov is genetic covariance, rcov is the residual covariance, gcovSE is the standard error for gcov, and abs is the absolute value. Details of the phenotype preparation for the adult traits are provided in the Supplementary Note.

Association between maternal SNPs associated with offspring birth weight and later-life offspring SBP. Using the UK Biobank, we tested whether maternal SNPs associated with offspring birth weight were also associated with offspring SBP in later life. The UK Biobank released kinship information generated in KING[65], which included kinship coefficients and estimates of the proportion of SNPs with zero identical-by-state (IBS0). We defined parent–offspring pairs using the kinship coefficient and IBS0 cut-offs recommended by Manichaikul et al.[65]. There were 5,635 unique parent–offspring pairs of European ancestry with SBP data (for parents who had multiple offspring with SBP data, only the oldest offspring was included in the analysis); 3,886 mother–offspring pairs and 1,749 father–offspring pairs. We tested the relationship between unweighted allelic scores of birth-weight-associated SNPs in mothers/fathers and offspring SBP (see Supplementary Note for SBP phenotype preparation) before and after adjusting for offspring genotypes at the same loci. We examined unweighted allelic scores consisting of all autosomal lead birth-weight-associated SNPs available in the UK Biobank (205 SNPs), 72 autosomal SNPs that showed evidence of a maternal effect and 31 autosomal SNPs that showed evidence only of maternal effects on birth weight. We also looked at the SNPs previously associated with SBP (Supplementary Table 14) as a sensitivity analysis to rule out the possibility of postnatal pleiotropic effects contaminating our results. All analyses were adjusted for offspring age at SBP measurement, sex and assessment center.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The genotype and phenotype data are available on application from the UK Biobank (http://www.ukbiobank.ac.uk/). Individual cohorts participating in the EGG Consortium should be contacted directly as each cohort has different data access policies. GWAS summary statistics from this study are available via the EGG website (https://egg-consortium.org/).

Code availability

Custom-written code is available on request from N.M.W. (e-mail: n.warrington@uq.edu.au).

References

46. Magi, R. & Morris, A. P. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* **11**, 288 (2010).
47. Kemp, J. P. et al. Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. *Nat. Genet.* **49**, 1468–1475 (2017).
48. Jones, S. E. et al. Genome-wide association analyses of chronotype in 697,828 individuals provides insights into circadian rhythms. *Nat. Commun.* **10**, 343 (2019).
49. Bulik-Sullivan, B. K. et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
50. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
51. Giambartolomei, C. et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).
52. Ihaka, R. & Gentleman, R. R: a language for data analysis and graphics. *J. Comput. Graph. Stat.* **5**, 299–314 (1996).
53. Gusev, A. et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).
54. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
55. Mardia, K. V., Kent, J. T. & Bibby, J. M. *Multivariate Analysis* (Academic Press, 1979).
56. Segrè, A. V. et al. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet.* **6**, e1001058 (2010).
57. Frayling, T. M. et al. A common allele in FGF21 associated with sugar intake is associated with body shape, lower total body-fat percentage, and higher blood pressure. *Cell Rep.* **23**, 327–336 (2018).
58. Zheng, J. et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272–279 (2016).
59. Prokopenko, I. et al. A central role for GRB10 in regulation of islet function in man. *PLoS Genet.* **10**, e1004235 (2014).
60. Burgess, S., Scott, R. A., Timpson, N. J., Davey Smith, G. & Thompson, S. G. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur. J. Epidemiol.* **30**, 543–552 (2015).
61. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**, 512–525 (2015).
62. Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.* **40**, 304–314 (2016).
63. Delaneau, O., Zagury, J. F. & Marchini, J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat. Methods* **10**, 5–6 (2013).
64. Loh, P. R. et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **47**, 284–290 (2015).
65. Manichaikul, A. et al. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).

nature research

Corresponding author(s): Rachel Freathy

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Data collection was performed centrally by the UK Biobank and by the investigators of each contributing EGG Consortium study. We provide detailed supplementary tables with characteristics of each study and references to study descriptions.
Data analysis	We used BOLT-LMM (version 2.3.2), GWAMA (version 2.2.2), GCTA (version 1.91.1), FUSION (version released 1 Dec 2016), TAD pathway analysis (accessed on 3 Nov 2017), METAL (version released 25 March 2011), LDSC-SEG (using LDSC version 1.0.0), MAGENTA (version 2.4), LD Hub (version 1.9.0), Coloc (version 3.1), Stata (version 13.0 and 14.0), LDlink (version 3.0) and R to analyse the data and produce plots for this manuscript. A full description of software used in this paper is provided in the Methods section, along with references to the relevant journal articles.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The following Data Availability statement is included in the manuscript:
Data availability
The genotype and phenotype data are available upon application from the UK Biobank (http://www.ukbiobank.ac.uk/). Individual cohorts participating in the EGG consortium should be contacted directly as each cohort has different data access policies. GWAS summary statistics from this study are available via the EGG website (https://egg-consortium.org/).

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Full descriptions of how we defined the samples included in the GWAS of own or offspring birth weight are included in the methods section. Briefly, we aimed to bring together the largest possible sample size with birth weight and/or offspring birth weight and GWAS data to study the role of genetic variation in birth weight. We excluded some individuals, as described below, and the final sample size was then the number of individuals included in any given statistical test that had no missing data for all genotype and phenotype variables included in the model. These numbers are given clearly in the main body of the article and the Methods section. Our sample size was adequate to recover the 65 known birth weight associated regions and identify 129 novel independent association signals. Additionally, as mentioned in the results section, asymptotic power calculations showed that with our sample size we had 80% power to detect fetal (maternal) effects that explained 0.006% (0.008%) of the variance in birth weight (α=0.05).
Data exclusions	Established protocols were used to conduct rigorous data quality control for each study prior to the GWAS analyses. The following exclusions were made for the birth weight phenotype: i) twins and other multiple births, ii) individuals born before 37 weeks of gestation (where gestational duration was available), iii) any extreme outliers in the birth weight distribution, iv) babies born with congenital anomalies (where available). In the UK Biobank, where there was no gestational duration information, we included only individuals born between 2.5 and 4.5kg. Within each genotyped dataset, to guard against population stratification, principal components analysis was used to exclude ancestry outliers (for example, the European component of the UK Biobank was defined using k-means cluster analysis, described in the Methods section). In the EGG studies, SNPs were excluded if the INFO<0.4, minimac r2<0.3 or minor allele count was <3. In the UK Biobank, SNPs were excluded if the INFO<0.3, minor allele frequency was <0.1% or they were not HRC imputed SNPs. Additional data exclusions for individual studies are provided in Supplementary Tables 1-4.
Replication	To determine the overall evidence of association at each SNP, we performed a meta-analysis of available GWAS studies. We checked for evidence of heterogeneity between the European component of the UK Biobank and the EGG European component to verify that there was consistency between the major components of the meta-analysis. This is described in the manuscript.
Randomization	Individuals in the UK Biobank were not assigned to specific groups, but data and samples were collected at different study centres and were genotyped in batches and on different arrays. Strict quality control performed centrally by UK Biobank ensured that batch-effects were controlled for in the genotype data. We adjusted genetic association tests for both study centre and data release (see Methods section for detail). We performed Mendelian randomization analyses, based on the assumption that genotype is randomly assigned at conception and thus that the maternal genotype can be used as an unconfounded proxy for a maternal characteristic, such as glucose, height etc. Essentially, using this method (with appropriate control for fetal genotype), birth weight is assessed in individuals effectively randomized to different maternal glucose/height etc levels. The Methods are described in detail in the manuscript.
Blinding	Data collected by both UK Biobank and EGG studies were observational and had no specific interventions. As such, no blinding was required.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	The UK Biobank is a cohort of British residents between the ages of 37 and 73, recruited to 22 centres at baseline measurement. Of the European-ancestry individuals who reported their own birth weight, 61% of participants were women. The EGG Consortium studies are mainly birth cohort designs, in which parent participants were recruited in pregnancy. Analyses of offspring birth weight (maternal GWAS) contained all women. In the UK Biobank, sex of the baby was not known; in the EGG Consortium studies, the male/female ratio of offspring was approx. 50/50. More details are provided in the supplementary tables and references. Participants were not selected on the basis of disease status.
Recruitment	The EGG Consortium studies are mainly birth cohort designs, in which parent participants were recruited in pregnancy and/or offspring at birth. Details are given in the supplementary tables and references. The UK Biobank consists of participants recruited in middle-age, and birth weight data are recalled and self-reported. We have shown previously that the self-reported variable is associated as expected with factors like sex, maternal age and height (Tyrrell et al 2013 IJE) and that genetic associations with birth weight are similar in the UK Biobank and EGG Consortium birth cohorts with measured birth weight (Horikoshi et al Nature 2016, and current manuscript), so we do not consider the self-report data to be a problem, especially given the large sample size. We note that the UK Biobank had a low response rate (5%), so we cannot rule out potential bias from selection, for example, resulting in higher average socio-economic position.