

Rare variant associations with birth weight identify genes involved in adipose tissue regulation, placental function and insulin-like growth factor signalling

Katherine A. Kentistou^{*1}, Brandon E. M. Lim^{*2}, Lena R. Kaisinger^{*1}, Valgerdur Steinthorsdottir³, Luke N Sharp², Kashyap A. Patel², Vinicius Tragante³, Gareth Hawkes², Eugene J. Gardner¹, Thorhildur Olafsdottir³, Andrew R. Wood², Yajie Zhao¹, Gudmar Thorleifsson³, Felix R. Day¹, Susan E. Ozanne⁶, Andrew T. Hattersley², Stephen O'Rahilly⁶, Kari Stefansson^{3,4}, Ken K. Ong^{*1,5}, Robin N. Beaumont^{*2}, John R. B. Perry^{*1,6}, Rachel M. Freathy^{*2}

1. MRC Epidemiology Unit, Box 285 Institute of Metabolic Science, University of Cambridge School of Clinical Medicine, Cambridge CB2 0QQ, UK

2. Department of Clinical and Biomedical Sciences, Faculty of Health and Life Sciences, University of Exeter, Exeter, UK

3. deCODE genetics/Amgen, Inc., 102 Reykjavik, Iceland

4. Faculty of Medicine, University of Iceland, 101 Reykjavik, Iceland

5. Department of Paediatrics, University of Cambridge, Cambridge CB2 0QQ, UK

6. MRC Metabolic Diseases Unit, Institute of Metabolic Science, University of Cambridge School of Clinical Medicine, University of Cambridge, Cambridge CB2 0QQ, UK

* These authors contributed equally to this work.

Abstract

Investigating the genetic factors influencing human birth weight may lead to biological insights into fetal growth and long-term health. Genome-wide association studies of birth weight have highlighted associated variants in more than 200 regions of the genome, but the causal genes are mostly unknown. Rare genetic variants with robust evidence of association are more likely to point to causal genes, but to date, only a few rare variants are known to influence birth weight. We aimed to identify genes that harbour rare variants that impact birth weight when carried by either the fetus or the mother, by analysing whole exome sequence data in UK Biobank participants.

We annotated rare (minor allele frequency <0.1%) protein-truncating or high impact missense variants on whole exome sequence data in up to 234,675 participants with data on their own birth weight (fetal variants), and up to 181,883 mothers who reported the birth weight of their first child (maternal variants). Variants within each gene were collapsed to perform gene burden tests and for each associated gene, we compared the observed fetal and maternal effects. We identified 8 genes with evidence of rare fetal variant effects on birth weight, of which 2 also showed maternal effects. One additional gene showed evidence of maternal effects only. We observed 10/11 directionally concordant associations in an independent sample of up to 45,622 individuals (sign test $P=0.01$).

Of the genes identified, *IGF1R* and *PAPPA2* (fetal and maternal-acting) have known roles in insulin-like growth factor bioavailability and signalling. *PPARG*, *INHBE* and *ACVR1C* (all fetal-acting) have known roles in adipose tissue regulation and rare variants in the latter two also showed associations with favourable adiposity patterns in adults. We highlight the dual role of *PPARG* in both adipocyte differentiation and placental angiogenesis. *NOS3*, *NRK*, and *ADAMTS8* (fetal and maternal-acting) have been implicated in both placental function and hypertension.

Analysis of rare coding variants has identified regulators of fetal adipose tissue and fetoplacental angiogenesis as determinants of birth weight, as well as further evidence for the role of insulin-like growth factors.

Introduction

An improved understanding of the genetic contribution to birth weight may highlight mechanisms relevant to fetal growth restriction or overgrowth, as well as links between fetal growth and later disease¹. Studies of rare, monogenic forms of diabetes have highlighted the key role of fetal insulin, since rare single gene mutations in the fetus that reduce insulin secretion are generally associated with a reduced birth weight. These include mutations that cause neonatal diabetes, where the severity of the insulin secretory defect correlates with the degree of reduction in birth weight². In babies who produce no insulin at all due to pancreatic agenesis, or to complete loss-of-function (LoF) mutations in the insulin gene, birth weights are 50% of those of a healthy term baby³.

Family studies of mutations in the *GCK* gene have demonstrated the role of maternal genetic variants, which may influence birth weight independently of variants inherited by the fetus. Maternal heterozygous mutations in *GCK*, which reduce glucose sensing and cause fasting glucose to be regulated at a higher set-point, result in greater maternal glucose availability to the fetus, as glucose crosses the placenta. This causes the fetus, if it does not carry the mutation, to produce more fetal insulin and grow bigger in response⁴. Conversely, if the fetus carries a *GCK* mutation, its ability to sense the increase in glucose is impaired, causing reduced fetal insulin secretion and concomitant reduced fetal growth.

Genome-wide association studies of birth weight have identified >200 regions of the genome where common variants (minor allele frequency (MAF) >1%) are associated with birth weight^{5,6}. Variants at the identified loci influence birth weight by direct effects of the fetal genotype, indirect effects of the maternal genotype acting on the intrauterine environment, or a combination of the two. Most of these variants likely influence growth via mechanisms that are independent of fetal insulin secretion⁷. Parent-of-origin effects in the fetus have been observed at several birth weight loci⁶. However, the causal genes at the vast majority of identified loci are unknown.

Whole exome sequence (WES) data in biobank-scale studies offer the opportunity to identify genes that are causally related to birth weight, and the potential to uncover new mechanisms of importance for the regulation of fetal growth. We performed exome-wide association studies (ExWAS) of rare variant (MAF <0.01) gene burden with birth weight in up to 234,675 UK Biobank participants who reported their own birth weight (fetal variants), and up to 181,883 female UK Biobank participants who reported the birth weight of their first child (maternal variants). The identified genes highlight key roles for adipose tissue regulation and fetoplacental angiogenesis, in addition to the role of insulin-like growth factor bioavailability and signalling, in regulating human birth weight.

Results

We identified 8 genes in which rare (MAF<0.1%) deleterious LoF fetal variants, defined as either high-confidence protein truncating variants, or missense variants with a CADD score

≥ 25 , were associated with birth weight at exome-wide significance (BOLT-LMM $P < 1.64 \times 10^{-6}$): *ACVR1C*, *IGF1R*, *INHBE*, *NOS3*, *NRK*, *NYNRIN*, *PAPPA2*, *PPARG* (Figure 1, Supplementary Figures 1-3, Supplementary Table 1). In the analysis of maternal variants and offspring birth weight, 3 genes showed associations at exome-wide significance (BOLT-LMM $P < 1.58 \times 10^{-6}$). These included 2 genes also identified in the fetal analysis (*IGF1R*, *NOS3*) and additionally *ADAMTS8* (Figure 1, Supplementary Figures 4-5, Supplementary Table 1). We also confirmed all of these significant associations by conducting independent analyses using REGENIE (see Methods) and we did not observe any method-related biases (Supplementary Figures 1, Supplementary Table 1).

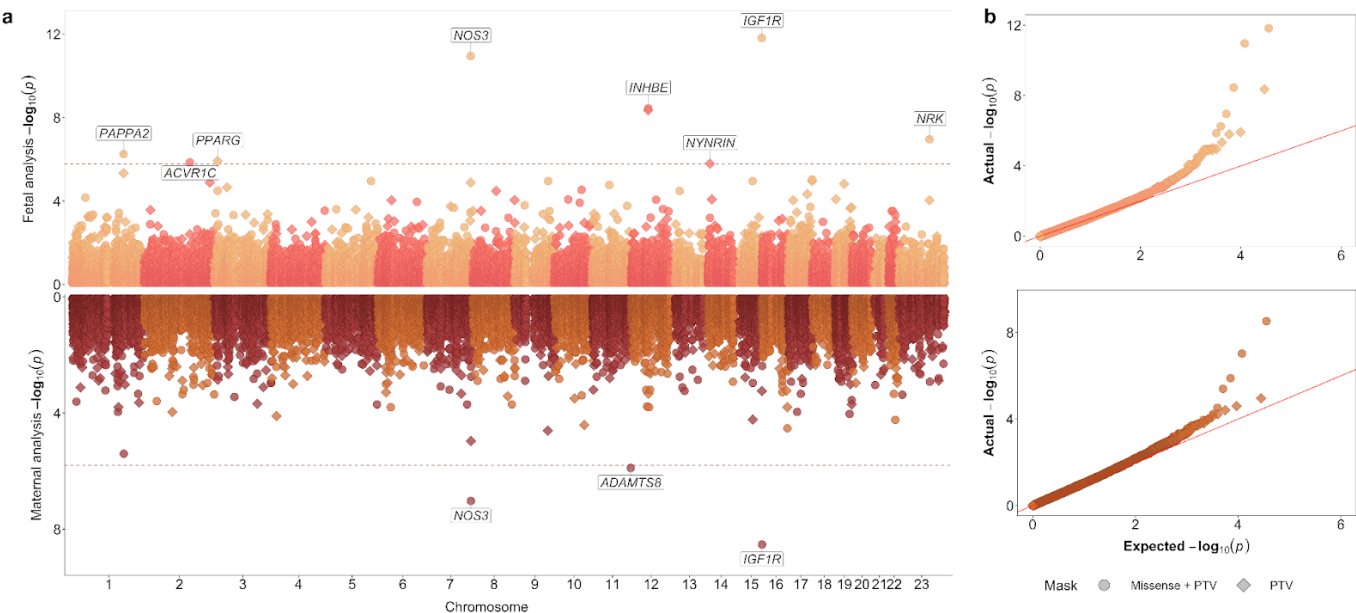


Figure 1 | Exome-wide rare (MAF<0.1%) fetal or maternal variant gene burden associations with birth weight. (a) Miami plot showing gene burden test results from BOLT-LMM for birth weight, with the fetal exome-wide analysis (up to n=234,675) on the top panel and the maternal exome-wide analysis (up to n=181,883) on the bottom panel. Gene associations passing exome-wide significance, at $P < 1.64 \times 10^{-6}$ in the fetal analysis and $P < 1.58 \times 10^{-6}$ in the maternal analysis, are labelled. The two variant collapsing masks are indicated by point shapes. (b) QQ plots for exome-wide gene burden associations. Results for the exome-wide significant genes are included in Supplementary Table 1.

Since maternal and fetal genotypes are correlated ($r=0.5$), we explored whether the observed associations represented fetal effects, maternal effects, or both. To do this, we used a weighted linear model (WLM) to approximately condition the fetal effect on the maternal genotype, and vice versa. Five out of nine associations (*ACVR1C*, *INHBE*, *NRK*, *NYNRIN*, *PPARG*) showed evidence of only fetal-genotype effects. *IGF1R*, *PAPPA2*, and *NOS3* were classified as both fetal- and maternal-acting, with rare variants in all three genes associating with a lower birth weight in both cases. *ADAMTS8* was the only gene classified as only maternal-acting (Figure 2a, Supplementary Table 2).

For fetal genes associated with birth weight, we explored whether they might act in a sex-specific manner. We found a nominally stronger effect for fetal *PPARG* variants on birth weight in females ($\beta_{\text{Female}} = -1.710$ [0.340] SDs) than in males ($\beta_{\text{Male}} = -0.177$ [0.592] SDs, $p_{\text{het}} = 2.46 \times 10^{-02}$), albeit only 3 male carriers were identified (Figure 2b, Supplementary Table 3).

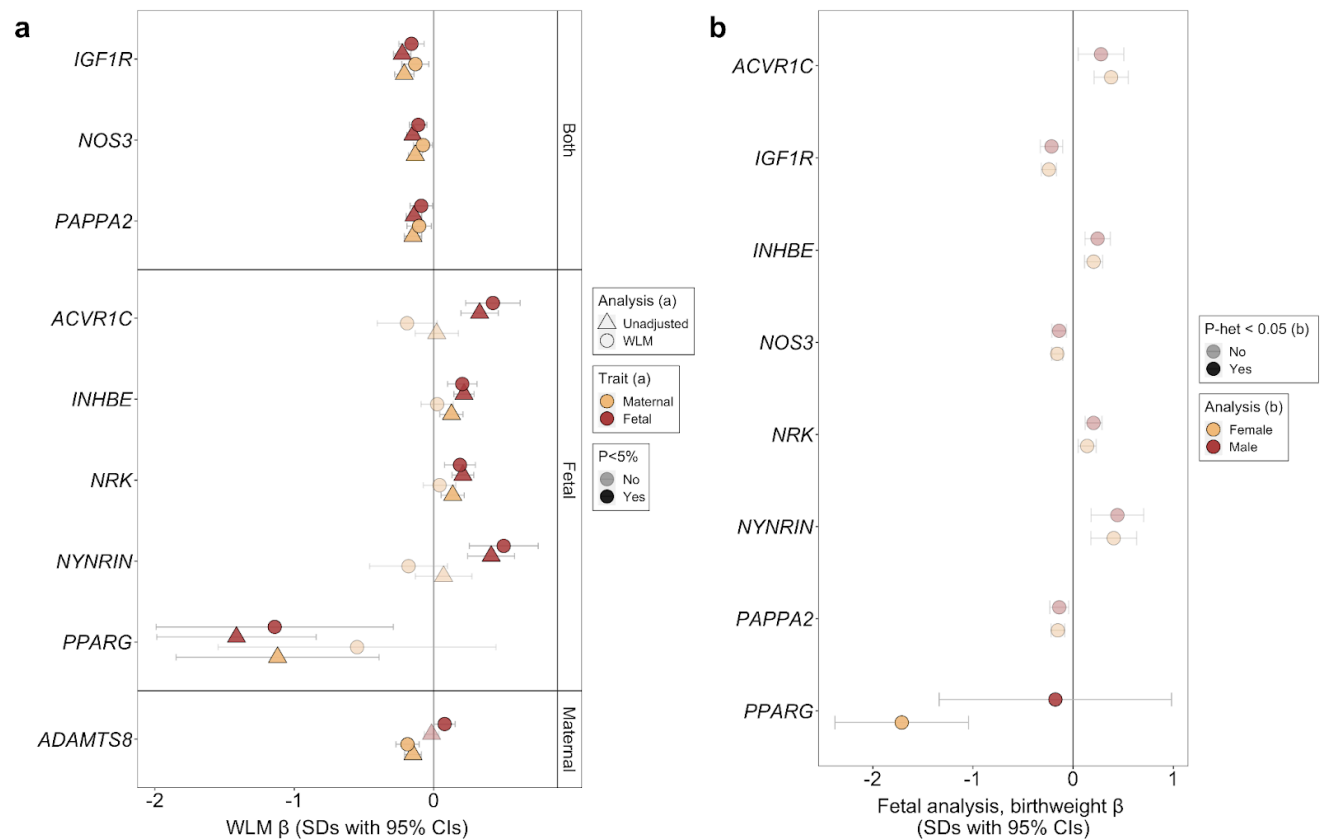


Figure 2 | Nine genes with rare variant associations with birth weight. (a) Weighted-linear models (WLMs) approximately conditioned the fetal effect on the maternal genotype, and vice versa. Marker opacity indicates gene burden associations at $P < 0.05$. (b) Sexually-dimorphic effects of the associated genes on birthweight. Markers are coloured yellow or red to indicate female- or male-only models. Marker opacity indicates associations with sexual dimorphism ($P < 0.05$). Units for effect estimates and accompanying 95% CIs are SDs in both panels. Relevant data are included in Supplementary Tables 2 and 3, accordingly.

Confirmation of exome associations

We aimed to replicate our associations in independent data on up to 45,622 Icelandic genomes. Despite the substantially smaller sample size than our discovery sample, we observed supportive evidence for the same type of rare LoF exome variants on birth weight

for the majority of our associations (10/11 showed directional consistency; Sign test $P=0.01$), with three genes associated at $P<0.05$; *INHBE*, *NRK* and *PAPPA2* (**Supplementary Figure 6, Supplementary Table 4**).

To identify whether any of the birth weight genes identified in our exome analyses are also supported by common variant associations, we examined the latest maternal and fetal GWAS summary statistics for birth weight⁶ and looked for GWAS signals within 500kb of our identified genes. We observed fetal GWAS signals proximal to *ACVR1C*, *IGF1R*, and *PAPPA2*, and there was also a maternal GWAS signal proximal to *NOS3* (**Supplementary Figure 7**). In all cases the lead associated variants were located within these genes (**Supplementary Table 5**) and proximity has been shown to be a good predictor of the causal gene at a GWAS locus^{8,9}. The lead SNPs at the *IGF1R* and *NOS3* loci were also eQTLs for these genes, with directions of effect concordant with the exome associations, i.e. the alleles associated with lower expression (LoF-like) were also associated with lower birth weight¹⁰. Of these four GWAS signals, *IGF1R*, *ACVR1C*, and *PAPPA2* were classified as having fetal only effects by Juliusdottir *et al.*⁶. The SNP near *NOS3* was not classified by Juliusdottir *et al.*⁶, but the *NOS3* lead SNP was classified by Warrington *et al.*⁵ as having only maternal effects. We did not observe attenuation of the rare variant associations with birth weight when conditioning on the proximal GWAS sentinel SNPs (**Supplementary Table 6**). Hence, these common variant associations provide independent support for these genes in the regulation of birth weight.

To assess evidence for additional candidate genes influencing birth weight at GWAS loci, which did not pass our stringent exome-wide significance thresholds, we looked at all genes within 300 kb of a reported GWAS locus and applied a Benjamini-Hochberg correction for multiple testing. The results are shown in **Supplementary Table 7**. In the fetal variant analysis we identified a further eight candidate genes (*CDK6*, *HGS*, *PAPPA*, *PHF19*, *PLAG1*, *PLCE1*, *PTEN* and *SKP2*), and in the maternal variant analysis we identified one additional candidate gene (*LMNA*). At six of these genes (*CDK6*, *PAPPA*, *PHF19*, *PLAG1*, *PLCE1* and *SKP2*) a GWAS lead SNP was located within the same gene, while the lead SNP within *PLCE1* was also a missense variant, further adding to the evidence that non-synonymous coding variation in these genes affects birth weight.

Finally, we explored whether the biology implicated by the exome-associated genes was consistent with that from GWAS studies. To do this, we performed a pathway enrichment analysis, using genes proximal to the GWAS signals from Juliusdottir *et al.*⁶. We saw enrichment for 185 GO pathways in the fetal GWAS of birth weight and 41 pathways in the maternal GWAS, with 21 pathways being enriched in both GWAS (fetal and maternal, **Supplementary Table 8**). In line with biology implicated by our exome associations, we observed significant enrichment for genes in multiple pathways implicated in insulin and growth factor response across both fetal and maternal GWAS genes and “circulatory system development” for fetal GWAS genes, among others.

Gene burden associations with related traits

To understand the wider phenotypic impacts of birth weight-associated rare genomic variation, we performed phenotypic association lookups in birthweight-related traits using data from the UK Biobank (see **Methods**).

We observed associations for rare LoF variants in *IGF1R* and *PAPPA2*, both of which are known regulators of insulin-like growth factor (IGF) regulation and bioactivity^{11–15}, with adult circulating IGF-1 levels, adult height and childhood height and body size (**Supplementary Table 9**). The *INHBE*, *NRK*, *NOS3* and *NYNRIN* genes were also associated with adult height. In all of these genes, apart from *INHBE*, carrying a rare LoF variant was associated with directionally concordant effects on birth weight and adult height, i.e. higher birth weight and higher adult height or *vice versa*.

Furthermore, we observed associations between variants in *ADAMTS8* and *NOS3* and adult blood pressure (**Supplementary Table 9**). For both of these genes, rare LoF variants were associated with lower birth weight and higher blood pressure.

Finally, several of the birth weight associated genes were also associated with measures of adiposity in adulthood, including waist-hip ratio (WHR) adjusted for BMI, and body fat percentage for rare LoF variants in *ACVR1C*, *INHBE* and *PPARG* (accordingly, **Supplementary Table 9**). These are further explored below.

ACVR1C, *INHBE* and favourable adiposity

ACVR1C (birth weight: $\beta_{\text{fetal}}=0.330$ SDs, $p=1.4\times 10^{-6}$, N=200 carriers, **Figure 3a**) encodes a type I receptor for the TGF-beta family of signalling molecules. It is predominantly expressed in adipose tissue¹⁰ and its primary function in metabolic regulation is to limit catabolic activities and preserve energy¹⁶. Low-frequency heterozygous mutations in humans have been linked to a favourable metabolic profile including lower WHR adjusted for BMI and protection against T2D^{17,18}. *INHBE* ($\beta_{\text{fetal}}=0.218$ SDs, $p=3.5\times 10^{-9}$, N=687 carriers, **Figure 3b**) encodes the inhibin subunit beta E of activin E, a growth factor belonging to the TGF-beta family. While this gene is predominantly expressed in the liver¹⁰, it does not appear to be necessary for normal liver function¹⁹, but rather acts as a negative regulator of energy storage in peripheral adipose tissue. Similarly to *ACVR1C*, a candidate receptor, heterozygous LoF carriers exhibit lower WHR adjusted for BMI and a favourable body fat distribution and metabolic profile^{20,21}. However, recent mouse studies^{22,23} showed that complete ablation of *INHBE* or *ACVR1C* leads to uncontrolled and pathological levels of lipolysis (see Discussion). We found that rare LoF variants in *ACVR1C* and *INHBE* are associated with increased birth weight and decreased WHR, indicative of a metabolically favourable adiposity pattern in adulthood (**Supplementary Table 9**).

Two genetics studies^{20,21} showed that the association between rare LoF variants in *INHBE* and WHR is primarily due to a splice acceptor variant (12:57456093:G:C) that substantially reduces the expression of the gene, attributable to either a change in secretion and/or protein

stability. They also reported that the favourable adiposity effect of this gene is almost entirely attenuated in the absence of this variant (gene-burden $p=2.01 \times 10^{-8}$ with splice variant, $p=0.34$ without²¹, which we also observe for WHR in our data (**Supplementary Table 10**). However, while this variant is also strongly associated with higher birth weight ($p=5.4 \times 10^{-6}$, **Figure 3b**), our observed birth weight association with *INHBE* does not fully attenuate in the absence of this variant (gene-burden $p=4.6 \times 10^{-9}$ with splice variant, $p=6.3 \times 10^{-5}$ without, **Supplementary Table 10**).

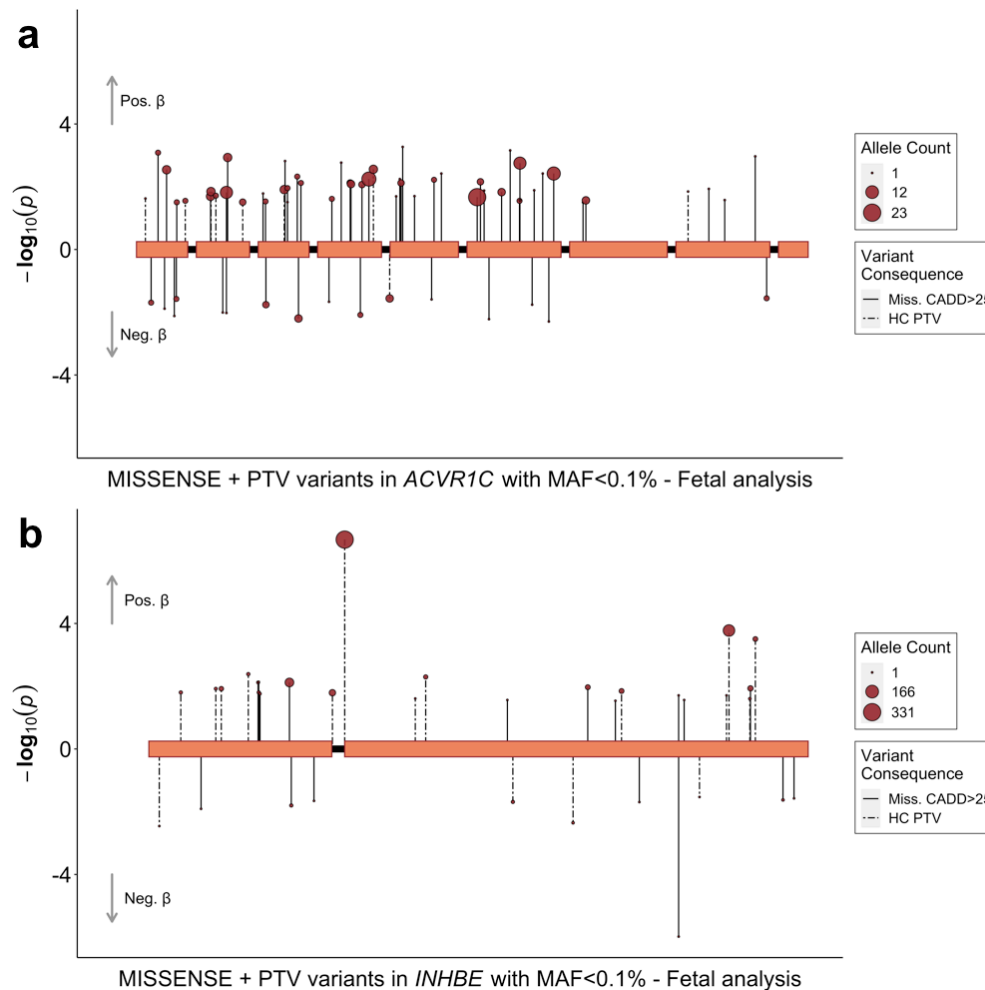


Figure 3 | Rare variant associations at *INHBE* and *ACVR1C* with fetal birth weight in the UK Biobank. Fetal variant-level associations between *ACVR1C* (a), *INHBE* (b) and birth weight. Included variants had a minor allele frequency (MAF) <0.1% and were annotated to be damaging variants defined as high-confidence protein truncating variants (PTV) or missense variants with a CADD score ≥ 25 . Each variant is presented as an individual line extending to its association p -value ($-\log_{10}$) in the direction indicating the direction of effect on birth weight in variant carriers. Dashed lines indicate PTVs and solid lines indicate missense variants. The point size indicates the number of carriers of each variant (i.e. the allele count).

PPARG in birth weight, adipogenesis, and placental formation

PPARG (birth weight: $\beta_{\text{fetal}} = -1.416$ SDs, $p = 1.2 \times 10^{-6}$, $N = 11$ carriers, **Figure 4b**) encodes the peroxisome proliferator-activated receptor PPAR-gamma, which acts as a master regulator of adipogenesis. Heterozygous LoF variants in humans have been linked to familial partial lipodystrophy (FPLD) characterised by loss of subcutaneous fat from the extremities and metabolic abnormalities such as insulin resistance and hypertriglyceridemia^{24–27}. Individuals with FPLD due to inherited heterozygous variants have also been shown to have low birthweight for their gestational age, when in the absence of maternal diabetes, hypertension or hypertriglyceridemia (e.g. when the *PPARG* variants are paternally inherited)²⁸. In addition to being highly expressed in adipocytes, *PPARG* is also highly expressed in trophoblasts in both rodent and human placentas²⁹ where it appears to play an important role in placental formation and angiogenesis by regulating the expression of proangiogenic factors such as proliferin (*Prl2c2*) and vascular endothelial growth factor (VEGF)³⁰. Its depletion results in embryonic lethality in mice due to placental dysfunction^{31,32}, while PPAR γ agonist activity was reduced in the blood and placentas of patients with severe preeclampsia, compared to those with normal pregnancy³³.

To distinguish if our observed birth weight association could be primarily due to *PPARG*'s function in fetal insulin resistance or placental function, we characterised rare variants in *PPARG* using Missense InTerpretation by Experimental Response (MITER) score data³⁴, as the gold-standard indicator of lipodystrophy-causing *PPARG* missense variants. Although MITER scores are not available for PTV variants, for which we observed the strongest birth weight association, we note that 8 out of our included 9 PTV variants are classified as pathogenic based on ACGS³⁵. Missense variants with a lipodystrophy-level MITER score (≤ -2) were collectively associated with a decrease in adult body fat percentage of similar magnitude to other prediction-based missense variant collapsing masks ($p = 9.70 \times 10^{-5}$, **Figure 4a, Supplementary Table 11**). We found only weak evidence of an association between the MITER score variants and birth weight ($p = 0.06$), although the effect estimate on birth weight was not significantly different from the other missense masks (**Figure 4a, Supplementary Table 11**). There was also no discernible correlation pattern between *PPARG* variant effect estimates on birth weight, body fat percentage and their corresponding MITER scores (Pearson correlations were, $R_{\text{BW-MITER}} = 0.026$, $R_{\text{BF\%-MITER}} = 0.187$, $R_{\text{BF\%-BW}} = 0.101$, **Figure 4c**). Taken together, our results suggest that the effects of rare *PPARG* PTVs on birth weight could be due to altered placentation, as well as the established role in adipogenesis.

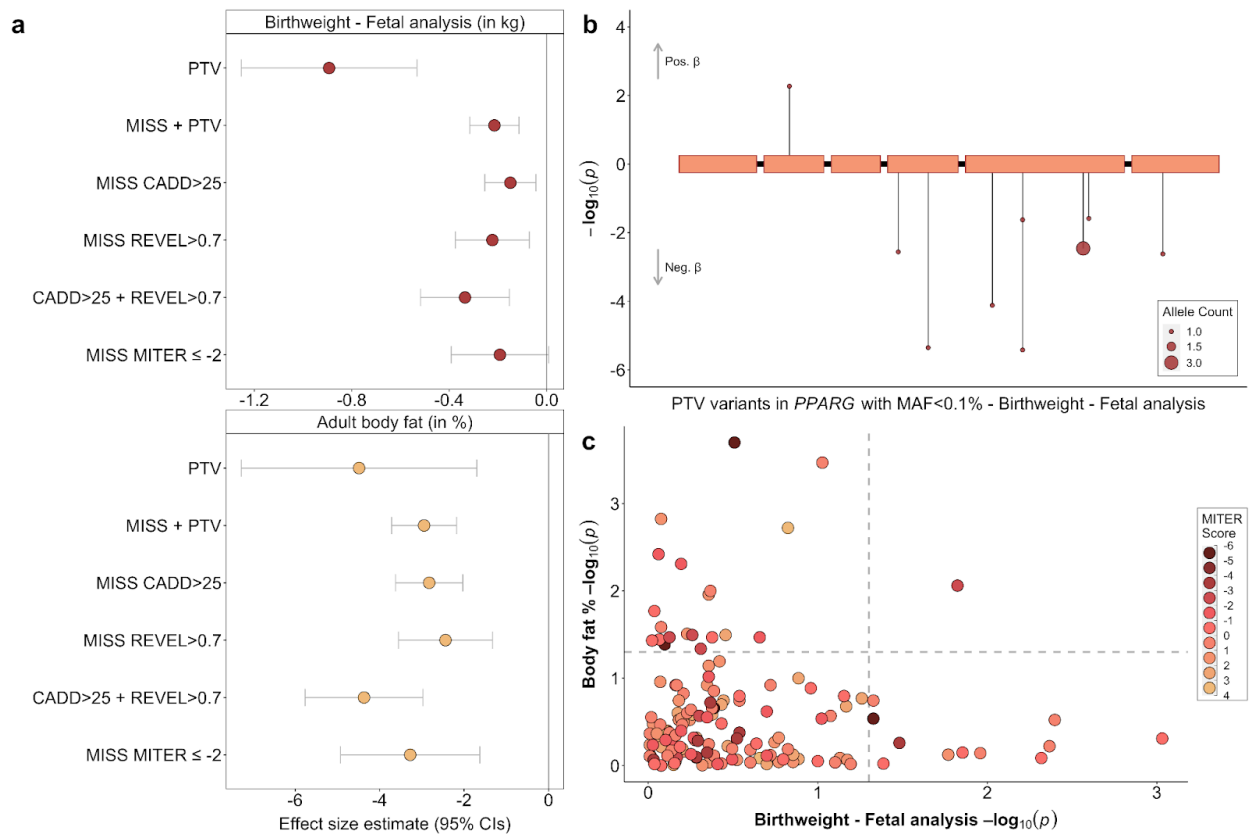


Figure 4 | Associations with rare LoF variants in *PPARG*. (a) Different (fetal) variant collapsing masks and their associations with birth weight (top) and adult body fat percentage (bottom). (b) The genomic location and associations with birth weight for all qualifying fetal high confidence PTVs within *PPARG*, in the discovery analysis. (c) Scatterplot of association P-values for birth weight (fetal variants) and body fat percentage, for all rare missense variants in *PPARG* annotated with a MITER score and found in the UK Biobank population. Marker colour indicates the variants MITER score, with scores ≤ -2 indicating lipodystrophy-level MITER scores. Relevant data is included in **Supplementary Table 11**.

As with *PPARG*, the *ADAMTS8*, *NRK* and *NOS3* genes have also been implicated in placental function. Specifically, *ADAMTS8* shows strong placental expression in the early stages of gestation and has been implicated in endometrium decidualization and trophoblast invasion^{36–39}. *NRK* encodes a serine/threonine kinase of the germinal centre kinase family and is expressed in skeletal muscle during development as well as the placenta, specifically the spongiotrophoblast layer, a fetus-derived region of the placenta^{40–42}. *Nrk*-deficient mouse fetuses display enlarged placentas, a likely result of enhanced trophoblast proliferation due to upregulation of AKT phosphorylation⁴³, as well as a higher birth weight and delayed delivery through yet unknown mechanisms⁴². Finally, common variants in *NOS3* have been previously associated with the risk of developing preeclampsia⁴⁴, although no association was found at this locus in a large-scale preeclampsia GWAS⁴⁵.

Comparison of burden test results with birth weight effects of pathogenic GCK variants

To investigate the sensitivity of our ExWAS approach to detect known rare variant effects on birth weight, we focused on the *GCK* gene. In the UK, pathogenic variants in *GCK* are the second most common cause of monogenic diabetes, and have established effects on fetal growth^{4,46}. Mothers with pathogenic *GCK* variants have a stably raised fasting glucose level, which results in greater glucose availability to the fetus, and increased birth weight due to fetal insulin secretion, which is raised in response to higher maternal and consequently fetal glucose. Conversely, fetal growth is reduced in babies who carry pathogenic *GCK* variants due to their reduced insulin secretion. These birth weight effects are masked if both mother and fetus carry the same mutation⁴. We annotated and combined pathogenic variants in *GCK* using clinical guidelines³⁵ and observed evidence of both fetal and maternal genetic associations with birth weight in the expected directions ($\beta_{\text{fetal}} = -0.311$ SDs, $p = 0.002$, $N = 86$ carriers; $\beta_{\text{maternal}} = 0.430$ SDs, $p = 0.001$, $N = 53$ carriers; **Supplementary Figure 8; Supplementary Table 12**). In comparison, our gene burden associations with rare predicted LoF variants in *GCK* were more modest but in the expected directions (**Supplementary Figure 8; Supplementary Table 12**). We found 63 of the known pathogenic *GCK* variants present in the UK Biobank WES dataset. Out of 16 variants in the PTV mask, 14 were annotated as pathogenic and of 55 variants in the Missense + PTV mask, 33 were annotated as pathogenic. The remaining pathogenic variants were annotated as non-deleterious and were not included in the masks.

Discussion

In this study, we used whole exome sequencing data from the UK Biobank cohort to understand the influence of rare genetic variants on human birth weight. We aimed to identify genes harbouring rare variants that impact birth weight when carried by either the fetus or the mother. We identified a total of 8 genes with rare fetal LoF variant effects on birth weight, two of which, *IGF1R* and *NOS3* showed some evidence of maternal effects. We identified one additional maternal effect gene, *ADAMTS8*. Findings were largely consistent in an independent cohort of Icelandic genomes, and four of the identified genes harboured independent GWAS lead SNPs in the largest and most recent birth weight GWAS meta-analysis, providing additional confirmation that the genes identified in this study are involved in birth weight regulation.

Among the genes associated with birth weight, several have well-established links with fetal growth. *IGF1R* and *PAPPA2* encode key components of IGF bioavailability and signalling. *IGF1R* is a transmembrane receptor tyrosine kinase activated by IGF1 (and also IGF2 and insulin), which mediates anabolic effects. Rare human homozygous, heterozygous or compound heterozygous mutations in *IGF1R* cause intrauterine growth restriction, reduced postnatal growth, short stature and microcephaly⁴⁷. *PAPPA2* encodes a metalloproteinase that regulates the bioavailability of IGFs. Rare homozygous mutations in *PAPPA2* are reported to cause severe short stature but with unclear effects on birth weight (reduced in 4

of 6 affected children)⁴⁸, while the protein originates primarily from trophoblasts in the early placenta^{49,50}. Our findings demonstrate that rare, damaging variants in *IGF1R* and *PAPPA2* contribute to population variation in birth weight, as well as in childhood and adult height.

Consistent with the fetal insulin hypothesis⁵¹, the association between the rare, damaging fetal variants in *IGF1R* and greater type 2 diabetes risk provides a potential mechanistic insight into the well-documented link between this later-life disease and lower birth weight. *IGF1R* is central to the development of several key tissues for glucose metabolism, including pancreatic islets, adipose tissue and skeletal muscle⁵², and in addition, loss of *IGF1R* function may lead to compensatory increases in growth hormone levels, which are associated with insulin resistance⁵³. The mechanism underlying the association we observed between maternal rare, damaging *IGF1R* variants and lower offspring birth weight after adjustment for fetal genotype is, however, less clear because we would expect such variants to increase birth weight if they result in higher maternal insulin resistance with concomitant higher glucose availability. However, there are some reported cases of heterozygous *IGF1R* mutations leading to hypoglycemia⁵⁴. We recommend that this result is confirmed in well-powered samples of mother-child pairs before potential mechanisms are investigated.

PPARG, *INHBE*, and *ACVR1C* displayed fetal effects on birthweight and are involved in adipose tissue differentiation and regulation. Recent genetics studies implicated rare LoF variants in *INHBE* in adult favourable adiposity traits, i.e. lower WHR adjusted for BMI^{20,21}. They also highlighted a splice variant (12:57456093:G:C) which affects either protein secretion or stability, as underpinning the favourable adiposity association with this gene. We found that the birth weight association with *INHBE* largely remains, even in the absence of this splice variant. This could indicate that some variants within *INHBE* may exert a role in early (*in utero*) adipogenesis, while others may do so throughout the life course, but further studies need to confirm this hypothesis.

Our identification of rare LoF variants in two genes, *ACVR1C* and *INHBE*, that have fetal associations with higher birth weight and also a more metabolically favourable body fat distribution in adulthood is in line with recent findings that common genetic variants in the fetus, which predispose to higher metabolically favourable adiposity in adulthood, are also associated with higher birth weight. Variants with greater effects on adiposity have greater effects on birth weight⁵⁵ and we therefore hypothesise that the birth weight effects of *ACVR1C* and *INHBE* are due primarily to fetal fat accretion rather than effects on lean mass. A fetal genetic predisposition to greater birth weight and later favourable body fat distribution may underlie epidemiological associations between greater skinfold thickness at birth and favourable metabolic outcomes in early childhood, especially when controlling for exposure to maternal glucose levels *in utero*⁵⁶. Further studies should assess the contributions of rare variation in *ACVR1C* and *INHBE* to the fat and lean mass components of birth weight.

More generally, the effects of loss of *INHBE* on *Alk7* (*ACVR1C*) signalling on metabolic health appear to be complex and dose-dependent. Humans haploinsufficient for either of these genes have a healthy metabolic profile and reduced risk of cardiometabolic disease

accompanied by a favourable distribution of body fat. However, it has been shown that the total absence of either INHBE or ACVR1C from conception results in uncontrolled adipocyte lipolysis, ectopic fat distribution and insulin resistance, at least in mice^{22,23}. This discrepancy may be explained by the fact that Alk7 signalling in adipose cells suppresses the expression of both PPAR γ and key molecular elements of the lipolytic machinery. We suggest that at modest levels of INHBE/Alk7 deficiency, such as occurs when only one functional allele is present, the effect on PPAR γ is dominant and carriers of such mutations develop more subcutaneous adipocytes as a result of the enhanced PPARG tone during development. Ultimately, this is metabolically beneficial as it will tend to protect the individual from any effects of any caloric overload in postnatal life. However, in the total absence of INHBE/Alk7 the effect on lipolysis becomes dominant with consequent build-up of ectopic fat and insulin resistance. While it is not entirely clear why subcutaneous, rather than visceral adipocytes would be favoured by PPAR γ activity, it is worthy of note that the administration of drugs activating PPAR γ preferentially promotes subcutaneous fat deposition⁵⁷.

Heterozygous LoF in *PPARG* has previously been linked to familial partial lipodystrophy, characterised by the loss of subcutaneous fat from the extremities and metabolic abnormalities such as insulin resistance and hypertriglyceridemia^{24–27}. However, *PPARG* has also been linked to placental development and angiogenesis³⁰. Here, we saw a poor correlation between the variants' lipodystrophy-causing potential and effect on birth weight, suggesting the birth weight effect could be driven by another mechanism. Given the role of *PPARG* in placental development, it is possible that *PPARG* LoF variants act via the placenta to exert an effect on birth weight. Our results are consistent with recent observations of a high rate of small-for-gestational-age birth weights in patients who had familial partial lipodystrophy due to inherited *PPARG* mutations and who were not exposed to maternal diabetes in utero²⁸. *PPARG* mutation carriers were also more likely to be born preterm, regardless of maternal diabetes status, which further supports the role of fetal *PPARG* in placental development. Data on gestational duration was not available for the UK Biobank participants, and it is possible that shorter gestation in rare *PPARG* variant carriers contributed to the associations we observed with birth weight. We note that the common fetal variant rs1801282 in *PPARG*, which is associated with type 2 diabetes risk in adulthood⁵⁸ via a fat distribution-mediated form of insulin resistance⁵⁹, is weakly associated with lower birth weight ($p=0.0259$)⁵ and was not associated with gestational duration in a fetal GWAS of gestational duration in 84,689 individuals ($p=0.487$)⁶⁰. Our observed associations between birthweight and *ADAMTS8*, *NRK* and *NOS3* are also likely driven by placental mechanisms, given the known roles of these genes^{36–44}. The potential mechanism underlying the association between birth weight and our other identified gene, *NYNRIN*, requires further study.

Our study has several strengths, including a large discovery sample, an independent replication sample, support from common variant associations proximal to five of the genes, and approximate adjustment for correlation between maternal and fetal genotype. However, there are limitations. The sample size available for replication was limited, and although we observed strong evidence for consistent effect estimates, we recommend confirmation in larger datasets as they become available, along with proper conditional analyses in

sufficiently-powered samples of mother-child pairs. The approximate adjustment for the correlation between maternal and fetal genotype used here is based on the strong assumption that all variants with a given annotation have equal effects on birth weight⁵, which may not be the case. We also acknowledge the limited diversity of our samples. Studies of rare variants in well-powered samples of diverse ancestries are a priority for future research. While our study was underpowered to detect very rare pathogenic variants that have known birth weight effects², our sensitivity analysis of *GCK* showed that the known maternal and fetal associations were detectable by our method, despite the inevitable lower sensitivity and specificity of the exome-wide approach, and despite the lack of adjustment for opposing effects of maternal and fetal genotypes.

Overall, this study advances our understanding of the fetal and maternal genetic underpinnings of birth weight, providing strong evidence for causal genes and insights into biological pathways which are important targets for future research aiming to understand fetal growth and its links with long-term health.

Acknowledgements

This research has been conducted using the UK Biobank Resource under Application Numbers 9905 and 7036. UK Biobank data has approval from the North West Multi-centre Research Ethics Committee (MREC) as a Research Tissue Bank (RTB). Use of the Icelandic data was approved by the National Bioethics Committee (VSN-15-169). All genotyped participants signed a written informed consent allowing the use of their samples and data in projects at deCODE genetics approved by the NBC. Data were anonymized and encrypted by a third-party system, approved and monitored by the Icelandic Data Protection Authority⁶¹. B.E.M.L., R.N.B. and R.M.F. were supported by a Wellcome Senior Research Fellowship (WT220390). This research was funded in part by the Wellcome Trust (WT220390). R.M.F. is also supported by a grant from the Eunice Kennedy Shriver National Institute Of Child Health & Human Development of the National Institutes of Health under Award Number R01HD101669. K.A.K., L.R.K., E.J.G., Y.Z., F.R.D., K.K.O. and J.R.B.P. are supported by the Medical Research Council (MC_UU_00006/2). This study was supported by the National Institute for Health and Care Research (NIHR) Exeter Biomedical Research Centre and the NIHR Cambridge Comprehensive Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. For the purpose of open access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Author contributions

K.A.K., B.E.M.L., L.R.K., K.K.O., R.N.B., J.R.B.P., and R.M.F. analysed and interpreted the original data and wrote the manuscript. V.S., V.T., T.O., G.T., and K.S. contributed the Icelandic replication data. L.S., K.A.P., G.H., E.J.G., A.R.W., Y.Z., F.R.D., S.E.O., A.T.H., and S.O. contributed to data analysis and/or interpretation. All authors revised, critically reviewed and approved the manuscript.

Competing interests

J.R.B.P. and E.J.G. are employees/shareholders of Insmed. J.R.B.P. also receives research funding from GSK. Y.Z. is a UK University worker at GSK. S.O. has undertaken remunerated consultancy work for Pfizer, Third Rock Ventures, AstraZeneca, NorthSea Therapeutics and Courage Therapeutics. V.S., V.T., T.O., G.T., and K.S. are employees of deCODE genetics, a subsidiary of Amgen.

Code availability

Code for WES data processing and association testing is available on GitHub (<https://github.com/mrcepid-rap/mrcepid-runassociationtesting>). No custom code was developed for this study.

Data availability

UK Biobank data are available on application via <https://www.ukbiobank.ac.uk/enable-your-research/register>.

Materials and methods

Exome wide association analyses in the UK Biobank

Analyses of the UK Biobank data were performed on the UK Biobank Research Analysis Platform (RAP; <https://ukbiobank.dnanexus.com/>). Birth weight was derived in the 453,505 UK Biobank participants with genetically defined European ancestry, as follows. Own birth weight (fetal genotype analysis) was derived using self-reported data (field 20022), while birth weight of the first child in UK Biobank women (maternal genotype analysis) was taken from hospital records and self-reports (fields 41284 and 2744). Data from each field were converted to kg (if in another unit) and if both were available, hospital records were preferentially retained for maximum accuracy. For both phenotypes, repeated responses across the assessment centre visits were combined by calculating average values. Individuals from multiple births (field 1777), who reported substantially different values between visits (>1kg difference), or extreme birth weight values (<1 or ≥7kg) were excluded from downstream analyses. After these exclusions, fetal birth weight data was available for 252,329 individuals (234,675 with exome and covariate data) and maternal birth weight data was available for the first child of 195,653 women (181,883 with exome and covariate data).

We conducted gene burden tests for both fetal and maternal birth weight, by combining effects of all rare (MAF <0.1%) variants with predicted deleterious functional consequences across all protein-coding genes, as described in detail in Gardner *et al.*⁵³. Briefly, we queried population-level VCF files with data for 454,787 individuals provided from the UK Biobank study. Using bcftools⁶² multi-allelic variants were split and left-normalised, and all variants filtered using a missingness based approach. SNV genotypes with depth <7 and genotype quality <20 or InDel genotypes with a depth <10 and genotype quality <20 were set to missing. We also tested for an expected reference and alternate allele balance of 50% for heterozygous SNVs using a binomial test; SNV genotypes with a binomial test p. value ≤1×10⁻³ were set to missing. Following genotype filtering, variants with >50% missing genotypes were excluded from further analysis. Variants were then annotated with the ENSEMBL Variant Effect Predictor (VEP⁶³, v104) with the ‘everything’ flag and the LOFTEE plugin⁶⁴. For each variant we prioritised a single MANE (v0.97) or VEP canonical ENSEMBL transcript and most damaging consequence as defined by VEP defaults. For the purposes of defining Protein Truncating Variants (PTVs), we grouped high-confidence (as defined by LOFTEE) stop gained, splice donor/acceptor, and frameshift consequences. To define ‘high-confidence’ PTVs, we used the LOFTEE algorithm⁶⁴. All variants were subsequently annotated using CADD (v1.6,⁶⁵). We then assessed the association between rare variant burden and birth weight, using BOLT-LMM (v2.3.5,⁶⁶) and a set of dummy genotypes representing per-gene carrier status, under two overlapping variant collapsing masks; i)

high-confidence protein truncating variants (denoted PTV) and ii) High-confidence PTVs plus missense variants with CADD scores ≥ 25 (denoted Missense + PTV). All analyses were controlled for age, age², the first ten genetic ancestry principal components as calculated in Bycroft *et al.*⁶⁷, WES batch, and sex when running sex-combined analyses. We excluded genes with <10 carriers of variants per mask.

This led to testing 16,735 genes with qualifying Missense + PTV variants and 13,684 with PTVs for fetal birth weight, and 17,745 with Missense + PTV variants and 13,968 with PTVs for maternal birth weight. The exome-wide significance thresholds were thus set to $p=1.64 \times 10^{-6}$ ($0.05/(16735+13684)$) for fetal birth weight and $p=1.58 \times 10^{-6}$ ($0.05/(17745+13968)$) for maternal birth weight.

Confirmatory gene burden analyses

To replicate the findings from the above analysis and account for potential bias arising from only using one discovery approach, a second team independently derived the birth weight phenotypes for the maternal and fetal genetic analyses and conducted additional burden association analyses, specifically for the significantly associated birth weight genes.

Own birth weight (fetal genotype analysis) was derived using self-reported data (field 20022), excluding individuals who were part of a multiple birth (field 1777). Offspring birth weight (maternal genotype analysis) was derived from self reported weight of first child (field 2744). Where individuals reported their birth weight at more than one assessment centre visit, reported birth weights were averaged unless the weights differed by ≥ 1 kg, in which case they were excluded. Own and offspring birth weight were available in 269,921 (232,876 with exome and covariate data) and 216,798 (152,585 with exome and covariate data) individuals respectively.

Multiallelic variants were similarly split and left-normalised. Variants flagged for exclusion by UK Biobank were removed, and remaining variants were annotated with their functional consequence in canonical transcripts using the Ensembl Variant Effect Predictor and the CADD and LOFTEE plugins (<https://github.com/konradjk/loftee>). Association testing was performed using REGENIE (v3.1.3,⁶⁸), and variants were grouped according to two collapsing masks; i) loss of function variants defined by LOFTEE as high confidence (denoted PTV) and ii) LoF plus missense SNVs or inframe insertions or deletions with CADD scores ≥ 25 (denoted Missense + PTV). All analyses were controlled for maternal age at first birth (field 3872) or year of birth, sex (for fetal birth weight), assessment centre, WES sequencing batch, and the first ten principal components.

Exome-wide significantly associated genes from the above BOLT-LMM analysis were queried in the REGENIE results.

Follow-up analyses of identified genes

To distinguish between fetal genotyping-acting ('fetal-acting') and maternal genotype-acting ('maternal-acting') effects at each birth weight-associated gene, we applied a weighted linear model (WLM) to the burden test results. WLM estimates associations of fetal genotype conditional on maternal genotype, and vice versa, without the need for data from mother-child pairs. The WLM approach has been described previously⁵.

To investigate potential sexual dimorphism, we performed sex-stratified burden tests using generalised linear models in the discovery sample and for the most strongly associated mask-gene combination. Sex-stratified effect estimates for each gene were compared using two-sample T-tests. Associations were deemed dimorphic if the p -value for the T-test was <0.05 .

Where described, further missense variant collapsing masks and leave-one out analyses were tested using BOLT-LMM or linear models in the same individuals and birth weight outcomes as the discovery analyses. Such masks were defined using a combination of CADD⁶⁵ and REVEL⁶⁹.

Conditional analyses where the effect of the identified rare exome variant collapsing masks was conditioned on the genotype at common birth weight GWAS signals were also tested using linear models in the same individuals and birth weight outcomes as the discovery analyses. Genotypes at GWAS signals (as identified in ⁶) were extracted in this discovery sample using plink (v1.90b6.18, ⁷⁰). Birthweight outcomes were regressed against the covariates used in discovery analyses and the concurrent effects of exome variant carriage and GWAS signal genotype were tested against the residual birth weight variance. A change in effect estimate of more than 10% was considered to show significant attenuation in the presence of the GWAS variants.

Related trait PheWAS

For all identified exome associations to the two birth weight traits, we performed small-scale PheWAS on a few predefined relevant phenotypes.

For BOLT-LMM analyses, phenotypic outcomes were defined as follows. BMI raw values were used from field 21001. Body fat % was extracted using data from field 23099. Data from subsequent visits were used, if missing for a given instance. Comparative height at age 10 was recoded to have individuals who in field 1697 reported being 'Shorter' as 0, 'About average' as 1, and 'Taller' as 2. Individuals who reported 'Do not know' or 'Prefer not to answer' were set to 'NA'. Following variable recoding, this phenotype was run as a continuous trait. Comparative size at age 10 was defined as above using data from field 1687. Systolic and diastolic blood pressure was derived using data from fields 4079 and 4080, accordingly. For each, the average blood pressure at each assessment visit was calculated over the two available measurements and data from subsequent visits were used, if missing for a given instance. Individuals indicating the use of blood pressure medications in

fields 6153 and 6177 were adjusted for in the association models. Height raw values were used from field 50. IGF-1 levels were derived using data from field 30770, with exclusion of individuals >5 standard deviations from the mean. T2D was derived as described in ⁵³ and using data from fields 4041, 10844, 2443, 6177, 6153, 20002, 20003, 41202, 41204, 40001 and 40002. WHR adjusted for BMI; WHR was calculated using data from field 48 and 49. >4 standard deviations from the mean excluded following initial calculation and prior to adjusting with BMI (field 21001). Paired BMI and WHR data from subsequent visits were used, if missing for a given instance.

For REGENIE analyses, phenotypic outcomes are as follows: BMI was taken from the baseline value for field 21001 and residualised on age at assessment (field 21003), genetic sex (field 22001), assessment centre (field 54) and PCs 1-5. Blood pressure measurements were taken from the automated readings for diastolic blood pressure (field 4079) and systolic blood pressure (field 4080). Where two blood pressure readings were available, the average between the two was taken, excluding individuals where the measurements differed by >4.56 SD. Individuals with a blood pressure >4.56 SD away from the mean were excluded. Blood pressure medication was accounted for by adding 10 and 15 mmHg to diastolic and systolic blood pressure, respectively. Height raw values were taken from baseline values for field 50, where people whose height was 100 cm and sitting height (field 20015) was 20 cm, were removed. Raw IGF-1 values were used from field 30770. Size at age 10 was based on a continuous simulation of childhood BMI. The simulation converted the categorical variables 'thinner', 'the same size' or 'plumper' in field 1687 to a continuous distribution based on the number of individuals selecting each category. The full methods for the simulation have previously been reported⁷¹. WHR was calculated using waist circumference (field 48) and hip circumference (field 49). If a follow-up visit was recorded, the WHR was taken from the follow-up data. Otherwise, it was taken from the first recorded WHR measurement.

The multiple testing threshold was set at $p < 0.005$, after correcting for the above 10 queried phenotypes.

GWAS associations at identified loci

GWAS signals for fetal and maternal birth weight were accessed from Juliusdottir *et al.*⁶. These signals were lifted back to GRCh37 annotated with all proximal genes, defined as those within 500kb up- or downstream of the genes start or end sites, using the NCBI RefSeq gene map for GRCh37 (via <http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database/>). For signals proximal to the exome-associated genes we assessed whether the exome-associated gene was the closest gene to the GWAS signal. These signals were also queried using eQTL data from the GTEx study (V8, ¹⁰) and plotted using LocusZoom⁷².

We performed gene-centric biological pathway enrichment analysis using g:Profiler (via the R client "gprofiler2"⁷³, v0.2.1, accessed on 23/11/2023). Pathway enrichment analyses were performed using the genes calculated to be the closest per signal to the GWAS signals from Juliusdottir *et al.*⁶. Pathways with a Bonferroni corrected p -value < 0.05 were considered significantly enriched in the birth weight GWAS.

We also examined all genes within 300 kb of the previously published birth weight loci from Juliusdottir *et al.*⁶. For this restricted set of loci previously shown to have common genetic associations with birth weight, we applied a Benjamini-Hochberg procedure to identify genes with evidence of association.

Replication of exome associations in deCODE

Data preparation

The genome of the Icelandic population was characterised by whole-genome sequencing (WGS) of 63,460 Icelanders using Illumina standard TruSeq methodology to a mean depth of 35x (SD 8x) with subsequent long-range phasing⁷⁴, and imputing the information into 173,025 individuals chip-genotyped employing multiple Illumina platforms⁷⁵. Variant calling was performed using GraphTyper⁷⁶ (v2.6) and chip data was phased using SHAPEIT4⁷⁷.

We used Variant Effect Predictor (VEP)⁶³ to attribute predicted consequences to the variants sequenced in each dataset. We classified as high-impact variants those predicted as start-lost, stop-gain, stop-lost, splice donor, splice acceptor or frameshift, collectively called loss-of-function (LOF) variants. Variants of moderate impact are those classified as missense, splice-region and in-frame-indels.

Information on birthweight comes from the Icelandic birth register and Primary Health Care Clinics of the Capital area. Individual's fetal birth weight was available in up to 36,578 genotyped and 11,184 WGS participants, while maternal birth weight was available in 45,622 genotyped and 17,177 WGS participants.

Gene burden associations

We defined different models to group together various types of variant: (i) only LOF variants, filtered with LOFTEE⁶⁴; (ii) LOF and MIS (as described on Variant annotation), when predicted deleterious by CADD⁶⁵ score ≥ 25 . In all models, we used MAF $< 0.1\%$ to select variants for analyses.

Quantitative traits were analysed using a linear mixed model implemented in BOLT-LMM⁶⁶. To estimate the quality of the sequence variants across the entire set we regressed the alternative allele counts (AD) on the depth (DP) conditioned on the genotypes (GT) reported by GraphTyper⁷⁸. For a well behaving sequence variant, the mean alternative allele count for a homozygous reference genotype should be 0, for a heterozygous genotype it should be DP/2 and for homozygous alternative genotype it should be DP. Under the assumption of no sequencing or genotyping error, the expected value of AD should be DP conditioned on the genotype, in other words an identity line (slope 1 and intercept 0). Deviations from the identity line indicate that the sequence variant is spurious or somatic. We filter variants with slope less than 0.5. Additionally, GraphTyper employs a logistic regression model that assigns each variant a score (AAscore) predicting the probability that it is a true positive. We used only variants that have a AAscore > 0.8 .

Analysis of birth weight associations with pathogenic *GCK* mutations

We used a similar procedure to that used to define pathogenic variants in *GCK* in a recent paper⁷⁹. We reviewed all heterozygous missense/PTV variants in UK Biobank participants with whole exome sequencing that were observed at minor allele frequency (MAF) <0.001. We included variants in the analysis if missense/PTV variants were classified as pathogenic or likely pathogenic based on ACMG/AMP guidelines by clinical scientists at Exeter Molecular Genetic laboratory as part of routine clinical diagnostic care (i.e. previously seen in the MODY probands)³⁵ and were ultra-rare in the population (maximum allele count of 2 in gnomADv2.1, MAF <1.4x10⁻⁵). We manually reviewed sequence read data for all the pathogenic variants (missense and PTVs) in Integrative Genomics Viewer (IGV) to remove false-positive variants. We used linear regression to assess the difference in birth weight associated with being a carrier (separate maternal and fetal analyses) and adjusted for genotype, sex, mother's diabetes status, batch, year of birth, centre and principal components in the fetal analysis and genotype, batch, age at first birth, centre and principal components in the maternal birth weight analysis. We then compared the estimated effects and 95% confidence intervals for the pathogenic variants with those obtained from our PTV and PTV+missense masks in the REGENIE analyses.

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. Hughes, A. E., Hattersley, A. T., Flanagan, S. E. & Freathy, R. M. Two decades since the fetal insulin hypothesis: what have we learned from genetics? *Diabetologia* **64**, 717–726 (2021).
3. Hughes, A. E. *et al.* Monogenic disease analysis establishes that fetal insulin accounts for half of human fetal growth. *J. Clin. Invest.* **133**, e165402 (2023).
4. Hattersley, A. T. *et al.* Mutations in the glucokinase gene of the fetus result in reduced birth weight. *Nat. Genet.* **19**, 268–270 (1998).
5. Warrington, N. M. *et al.* Maternal and fetal genetic effects on birth weight and their relevance to cardio-metabolic risk factors. *Nat. Genet.* **51**, 804–814 (2019).
6. Juliusdottir, T. *et al.* Distinction between the effects of parental and fetal genomes on fetal growth. *Nat. Genet.* **53**, 1135–1142 (2021).
7. Hughes, A. E. *et al.* Fetal Genotype and Maternal Glucose Have Independent and Additive Effects on Birth Weight. *Diabetes* **67**, 1024–1029 (2018).
8. Aragam, K. G. *et al.* Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. *Nat. Genet.* **54**, 1803–1815 (2022).
9. Mountjoy, E. *et al.* An open approach to systematically prioritize causal variants and genes at all published human GWAS trait-associated loci. *Nat. Genet.* **53**, 1527–1533 (2021).
10. Lonsdale, J. *et al.* The genotype-tissue expression (GTEx) project. *Nat. Genet.* **45**, 580 (2013).
11. Sferruzzi-Perri, A. N., Sandovici, I., Constancia, M. & Fowden, A. L. Placental phenotype

- and the insulin-like growth factors: resource allocation to fetal growth. *J. Physiol.* **595**, 5057–5093 (2017).
12. Conover, C. A. *et al.* Metalloproteinase pregnancy-associated plasma protein A is a critical growth regulatory factor during fetal development. *Dev. Camb. Engl.* **131**, 1187–1194 (2004).
13. Bøtkjær, J. A., Noer, P. R., Oxvig, C. & Yding Andersen, C. A common variant of the pregnancy-associated plasma protein-A (PAPPA) gene encodes a protein with reduced proteolytic activity towards IGF-binding proteins. *Sci. Rep.* **9**, 13231 (2019).
14. Yan, X., Baxter, R. C. & Firth, S. M. Involvement of pregnancy-associated plasma protein-A2 in insulin-like growth factor (IGF) binding protein-5 proteolysis during pregnancy: a potential mechanism for increasing IGF bioavailability. *J. Clin. Endocrinol. Metab.* **95**, 1412–1420 (2010).
15. Dauber, A. *et al.* Mutations in pregnancy-associated plasma protein A2 cause short stature due to low IGF-I availability. *EMBO Mol. Med.* **8**, 363–374 (2016).
16. Ibáñez, C. F. Regulation of metabolic homeostasis by the TGF- β superfamily receptor ALK7. *FEBS J.* **289**, 5776–5797 (2022).
17. Justice, A. E. *et al.* Protein-coding variants implicate novel genes related to lipid homeostasis contributing to body-fat distribution. *Nat. Genet.* **1** (2019).
18. Emdin, C. A. *et al.* DNA Sequence Variation in ACVR1C Encoding the Activin Receptor-Like Kinase 7 Influences Body Fat Distribution and Protects Against Type 2 Diabetes. *Diabetes* **68**, 226–234 (2019).
19. Lau, A. L., Kumar, T. R., Nishimori, K., Bonadio, J. & Matzuk, M. M. Activin betaC and betaE genes are not essential for mouse liver growth, differentiation, and regeneration. *Mol. Cell. Biol.* **20**, 6127–6137 (2000).
20. Akbari, P. *et al.* Multiancestry exome sequencing reveals INHBE mutations associated with favorable fat distribution and protection from diabetes. *Nat. Commun.* **13**, 4844 (2022).

21. Deaton, A. M. *et al.* Rare loss of function variants in the hepatokine gene INHBE protect from abdominal obesity. *Nat. Commun.* **13**, 4319 (2022).
22. Griffin, J. D. *et al.* Hepatic Activin E mediates liver-adipose inter-organ communication, suppressing adipose lipolysis in response to elevated serum fatty acids. *Mol. Metab.* **78**, 101830 (2023).
23. Adam, R. C. *et al.* Activin E-ACVR1C cross talk controls energy storage via suppression of adipose lipolysis in mice. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2309967120 (2023).
24. Agarwal, A. K. & Garg, A. A novel heterozygous mutation in peroxisome proliferator-activated receptor-gamma gene in a patient with familial partial lipodystrophy. *J. Clin. Endocrinol. Metab.* **87**, 408–411 (2002).
25. Barroso, I. *et al.* Dominant negative mutations in human PPARgamma associated with severe insulin resistance, diabetes mellitus and hypertension. *Nature* **402**, 880–883 (1999).
26. Hegele, R. A., Cao, H., Frankowski, C., Mathews, S. T. & Leff, T. PPARG F388L, a transactivation-deficient mutant, in familial partial lipodystrophy. *Diabetes* **51**, 3586–3590 (2002).
27. Savage, D. B. *et al.* Human metabolic syndrome resulting from dominant-negative mutations in the nuclear receptor peroxisome proliferator-activated receptor-gamma. *Diabetes* **52**, 910–917 (2003).
28. Gosseume, C. *et al.* Perinatal, metabolic, and reproductive features in PPARG-related lipodystrophy. *Eur. J. Endocrinol.* **188**, lvad023 (2023).
29. Fournier, T., Tsatsaris, V., Handschuh, K. & Evain-Brion, D. PPARs and the placenta. *Placenta* **28**, 65–76 (2007).
30. Nadra, K. *et al.* PPARgamma in placental angiogenesis. *Endocrinology* **151**, 4969–4981 (2010).
31. Barak, Y. *et al.* PPAR gamma is required for placental, cardiac, and adipose tissue development. *Mol. Cell* **4**, 585–595 (1999).

32. Kubota, N. *et al.* PPAR gamma mediates high-fat diet-induced adipocyte hypertrophy and insulin resistance. *Mol. Cell* **4**, 597–609 (1999).
33. Waite, L. L., Louie, R. E. & Taylor, R. N. Circulating Activators of Peroxisome Proliferator-Activated Receptors Are Reduced in Preeclamptic Pregnancy. *J. Clin. Endocrinol. Metab.* **90**, 620–626 (2005).
34. Majithia, A. R. *et al.* Prospective functional classification of all possible missense variants in PPARG. *Nat. Genet.* **48**, 1570–1575 (2016).
35. Ellard, S. *et al.* ACGS Best Practice Guidelines for Variant Classification in Rare Disease 2020. ACGS (2020).
36. Wen, J., Zhu, H. & Leung, P. C. K. Gonadal steroids regulate the expression of aggrecanases in human endometrial stromal cells in vitro. *J. Cell. Mol. Med.* **17**, 1325–1334 (2013).
37. Namli Kalem, M., Kalem, Z., Bakirarar, B. & Demircan, K. Adamts 1, 4, 5, 8, and 9 in Early Pregnancies. *Fetal Pediatr. Pathol.* **36**, 387–399 (2017).
38. Qu, H. & Khalil, R. A. Role of ADAM and ADAMTS disintegrin and metalloproteinases in normal pregnancy and preeclampsia. *Biochem. Pharmacol.* **206**, 115266 (2022).
39. Tatar, M., Uslu, S. & Öner, J. Expression of placental growth factor and a disintegrin and metalloprotease with a thrombospondin type motifs 1-4-8 during the three trimesters of rat pregnancy at the maternal-fetal interface. *Anat. Histol. Embryol.* **52**, 805–814 (2023).
40. Kanai-Azuma, M. *et al.* Nrsk: a murine X-linked NIK (Nck-interacting kinase)-related kinase gene expressed in skeletal muscle. *Mech. Dev.* **89**, 155–159 (1999).
41. Nakano, K., Yamauchi, J., Nakagawa, K., Itoh, H. & Kitamura, N. NESK, a member of the germinal center kinase family that activates the c-Jun N-terminal kinase pathway and is expressed during the late stages of embryogenesis. *J. Biol. Chem.* **275**, 20533–20539 (2000).
42. Denda, K. *et al.* Nrsk, an X-linked protein kinase in the germinal center kinase family, is

- 803 required for placental development and fetoplacental induction of labor. *J. Biol. Chem.* **286**,
804 28802–28810 (2011).
- 805 43. Morioka, Y., Nam, J.-M. & Ohashi, T. Nik-related kinase regulates trophoblast
806 proliferation and placental development by modulating AKT phosphorylation. *PloS One* **12**,
807 e0171503 (2017).
- 808 44. Zeng, F. *et al.* Associations between nitric oxide synthase 3 gene polymorphisms and
809 preeclampsia risk: a meta-analysis. *Sci. Rep.* **6**, 23407 (2016).
- 810 45. Honigberg, M. C. *et al.* Polygenic prediction of preeclampsia and gestational
811 hypertension. *Nat. Med.* **29**, 1540–1549 (2023).
- 812 46. Thanabalasingham, G. & Owen, K. R. Diagnosis and management of maturity onset
813 diabetes of the young (MODY). *BMJ* **343**, d6044 (2011).
- 814 47. Murray, P. G. & Clayton, P. E. Disorders of Growth Hormone in Childhood. in *Endotext*
815 (eds. Feingold, K. R. *et al.*) (MDText.com, Inc., South Dartmouth (MA), 2000).
- 816 48. Barrios, V. *et al.* Pregnancy-Associated Plasma Protein (PAPP)-A2 in Physiology and
817 Disease. *Cells* **10**, 3576 (2021).
- 818 49. Overgaard, M. T. *et al.* Pregnancy-associated Plasma Protein-A2 (PAPP-A2), a Novel
819 Insulin-like Growth Factor-binding Protein-5 Proteinase. *J. Biol. Chem.* **276**, 21849–21853
820 (2001).
- 821 50. Tóth, E. *et al.* Decreased Expression of Placental Proteins in Recurrent Pregnancy Loss:
822 Functional Relevance and Diagnostic Value. *Int. J. Mol. Sci.* **25**, 1865 (2024).
- 823 51. Hattersley, A. T. & Tooke, J. E. The fetal insulin hypothesis: an alternative explanation of
824 the association of low birthweight with diabetes and vascular disease. *Lancet Lond. Engl.*
825 **353**, 1789–1792 (1999).
- 826 52. Rother, K. I. & Accili, D. Role of insulin receptors and IGF receptors in growth and
827 development. *Pediatr. Nephrol.* **14**, 558–561 (2000).
- 828 53. Gardner, E. J. *et al.* Damaging missense variants in IGF1R implicate a role for IGF-1

829 resistance in the etiology of type 2 diabetes. *Cell Genomics* **2**, None (2022).

830 54. Solomon-Zemler, R. *et al.* A novel heterozygous IGF-1 receptor mutation associated with

831 hypoglycemia. *Endocr. Connect.* **6**, 395–403 (2017).

832 55. Thompson, W. D. *et al.* Fetal alleles predisposing to metabolically favorable adiposity are

833 associated with higher birth weight. *Hum. Mol. Genet.* **31**, 1762–1775 (2022).

834 56. Bianco, M. E. *et al.* Hyperglycemia and Adverse Pregnancy Outcome Follow-Up Study:

835 newborn anthropometrics and childhood glucose metabolism. *Diabetologia* **64**, 561–570

836 (2021).

837 57. Adams, M. *et al.* Activators of peroxisome proliferator-activated receptor gamma have

838 depot-specific effects on human preadipocyte differentiation. *J. Clin. Invest.* **100**, 3149–3153

839 (1997).

840 58. Altshuler, D. *et al.* The common PPARgamma Pro12Ala polymorphism is associated with

841 decreased risk of type 2 diabetes. *Nat. Genet.* **26**, 76–80 (2000).

842 59. Udler, M. S. *et al.* Type 2 diabetes genetic loci informed by multi-trait associations point

843 to disease mechanisms and subtypes: A soft clustering analysis. *PLOS Med.* **15**, e1002654

844 (2018).

845 60. Liu, X. *et al.* Variants in the fetal genome near pro-inflammatory cytokine genes on 2q13

846 associate with gestational duration. *Nat. Commun.* **10**, 3927 (2019).

847 61. Gulcher, J. R., Kristjánsson, K., Gudbjartsson, H. & Stefánsson, K. Protection of privacy

848 by third-party encryption in genetic research in Iceland. *Eur. J. Hum. Genet. EJHG* **8**,

849 739–742 (2000).

850 62. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping

851 and population genetical parameter estimation from sequencing data. *Bioinforma. Oxf. Engl.*

852 **27**, 2987–2993 (2011).

853 63. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).

854 64. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in

141,456 humans. *Nature* **581**, 434–443 (2020).

65. Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J. & Kircher, M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* **47**, D886–D894 (2019).

66. Loh, P.-R. *et al.* Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **47**, 284–290 (2015).

67. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).

68. Mbatchou, J. *et al.* Computationally efficient whole-genome regression for quantitative and binary traits. *Nat. Genet.* **53**, 1097–1103 (2021).

69. Ioannidis, N. M. *et al.* REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am. J. Hum. Genet.* **99**, 877–885 (2016).

70. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

71. Hawkes, G. *et al.* Genetic evidence that high BMI in childhood has a protective effect on intermediate diabetes traits, including measures of insulin sensitivity and secretion, after accounting for BMI in adulthood. *Diabetologia* **66**, 1472–1480 (2023).

72. Boughton, A. P. *et al.* LocusZoom.js: interactive and embeddable visualization of genetic association study results. *Bioinforma. Oxf. Engl.* **37**, 3017–3018 (2021).

73. Kolberg, L., Raudvere, U., Kuzmin, I., Vilo, J. & Peterson, H. gprofiler2 -- an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler. *F1000Research* **9**, ELIXIR-709 (2020).

74. Kong, A. *et al.* Detection of sharing by descent, long-range phasing and haplotype imputation. *Nat. Genet.* **40**, 1068–1075 (2008).

75. Gudbjartsson, D. F. *et al.* Large-scale whole-genome sequencing of the Icelandic population. *Nat. Genet.* **47**, 435–444 (2015).

- 881 76. Eggertsson, H. P. *et al.* GraphTyper2 enables population-scale genotyping of structural
882 variation using pangenome graphs. *Nat. Commun.* **10**, 5402 (2019).
- 883 77. Delaneau, O., Zagury, J.-F., Robinson, M. R., Marchini, J. L. & Dermitzakis, E. T.
884 Accurate, scalable and integrative haplotype estimation. *Nat. Commun.* **10**, 5436 (2019).
- 885 78. Halldorsson, B. V. *et al.* The sequences of 150,119 genomes in the UK Biobank. *Nature*
886 **607**, 732–740 (2022).
- 887 79. Mirshahi, U. L. *et al.* Reduced penetrance of MODY-associated HNF1A/HNF4A variants
888 but not GCK variants in clinically unselected cohorts. *Am. J. Hum. Genet.* **109**, 2018–2028
889 (2022).