Genome-wide association study of offspring birth weight in 86,577 women highlights maternal genetic effects that are independent of fetal genetics

Robin N Beaumont[1]*, Nicole M Warrington[2]*, Alana Cavadino[3], Jessica Tyrrell[1,4], Michael Nodzenski[5], Momoko Horikoshi[6,7], Frank Geller[8], Ronny Myhre[9], Rebecca C. Richmond[10-12], Lavinia Paternoster[10], Jonathan P. Bradfield[13], Eskil Kreiner-Møller[14], Ville Huikari [15], Sarah Metrustry[16], Kathryn L. Lunetta[17,18], Jodie N. Painter[19], Jouke-Jan Hottenga[20,21], Catherine Allard[22], Sheila J. Barton[23], Ana Espinosa[24-26], Julie A. Marsh[27], Catherine Potter[28], Ge Zhang[29-31], Wei Ang[27], Diane J. Berry[32], Luigi Bouchard[22,33,34], Shikta Das[35], Early Growth Genetics (EGG) Consortium§, Hakon Hakonarson[13,36,37], Jani Heikkinen[38], Øyvind Helgeland[39-41], Berthold Hocher[42,43], Albert Hofman[44], Hazel M. Inskip[23,45], Samuel E Jones[1], Manolis Kogevinas[24-26], Penelope A. Lind[19], Letizia Marullo[46], Sarah E. Medland[19], Anna Murray[1], Jeffrey C. Murray[47], Pål R. Njølstad[39,40], Ellen A. Nohr[48], Christoph Reichetzeder[43,49], Susan M. Ring[10,11], Katherine S Ruth[1], Loreto Santa-Marina[25,50,51], Denise M. Scholtens[5], Sylvain Sebert[15,52], Verena Sengpiel[53], Marcus A Tuke[1], Marc Vaudel[39], Michael N Weedon[1], Gonneke Willemsen[20,21], Andrew R Wood[1], Hanieh Yaghootkar[1], Louis J. Muglia[30,31], Meike Bartels[20,21], Caroline L. Relton[10,11,28], Craig E. Pennell[27], Leda Chatzi[54], Xavier Estivill[24,25,55], John W. Holloway[56], Dorret I. Boomsma[20,21], Grant W. Montgomery[19], Joanne M. Murabito[18,57], Tim D. Spector[16], Christine Power[32], Marjo-Ritta Järvelin[15,52,58-60], Hans Bisgaard[14], Struan F.A. Grant[13,36,37], Thorkild I.A. Sørensen[10,61,62], Vincent W. Jaddoe[12,44,63], Bo Jacobsson[9,53], Mads Melbye [8,64], Mark I. McCarthy[6,7,65], Andrew T. Hattersley[1], M. Geoffrey Hayes[66], Timothy M. Frayling[1], Marie-France Hivert[67-69], Janine F. Felix[12,44,63], Elina Hyppönen[32,70,71], William L. Lowe, Jr.[66], David M Evans[2,10,11]†, Debbie A. Lawlor[10,11]†, Bjarke Feenstra[8]† and Rachel M. Freathy[1,10]†

*These authors contributed equally to this work

†These authors jointly directed this work

§A full list of consortium members is included in the **Supplementary Material**

Affiliations

1. Institute of Biomedical and Clinical Science, University of Exeter Medical School, University of Exeter, Royal Devon and Exeter Hospital, Barrack Road, Exeter, EX2 5DW, UK.
2. University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Queensland, Australia.
3. Centre for Environmental and Preventive Medicine, Wolfson Institute of Preventive Medicine, Barts and the London School of Medicine and Dentistry, Queen Mary University of London.
4. European Centre for Environment and Human Health, University of Exeter, The Knowledge Spa, Truro, TR1 3HD.
5. Department of Preventive Medicine, Northwestern University Feinberg School of Medicine.
6. Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, UK.
7. Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.
8. Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark.
9. Department of Genes and Environment, Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway.
10. Medical Research Council Integrative Epidemiology Unit at the University of Bristol, Oakfield House, Oakfield Grove Bristol, BS8 2BN, UK.

11. Population Health Science, Bristol Medical School, University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN, UK.
12. The Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, P.O.Box 2040, 3000 CA, Rotterdam, the Netherlands.
13. Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA.
14. Copenhagen Prospective Studies on Asthma in Childhood (COPSAC), Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark & Danish Pediatric Asthma Center, Copenhagen University Hospital, Gentofte, Denmark.
15. Institute of Health Sciences, University of Oulu, Oulu, Finland
16. Department of Twin Research, King's College London, St. Thomas' Hospital, London, UK.
17. Boston University School of Public Health, Department of Biostatistics, Boston, MA, USA.
18. Framingham Heart Study, Framingham, MA, USA.
19. QIMR Berghofer Medical Research Institute, Locked Bag 2000, Royal Brisbane Hospital, Herston, Qld 4029, Australia.
20. EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands.
21. Department of Biological Psychology, Vrije Universiteit Amsterdam, Van der Boechorststraat 1, 1081 BT Amsterdam, the Netherlands.
22. Centre de recherche du Centre Hospitalier Universitaire de Sherbrooke, Sherbrooke, QC, Canada.
23. Medical Research Council Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK
24. Pompeu Fabra University (UPF), Barcelona, Spain.
25. CIBER Epidemiología y Salud Pública (CIBERESP), Spain.
26. IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain.
27. Division of Obstetrics and Gynaecology, The University of Western Australia, Perth, Australia.
28. Institute of Genetic Medicine, Newcastle University, Central Parkway, Newcastle upon Tyne, NE1 3BZ, UK.
29. Human Genetics Division, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA.
30. Center for Prevention of Preterm Birth, Perinatal Institute, Cincinnati Children's Hospital Medical Center, OH, USA.
31. March of Dimes Prematurity Research Center Ohio Collaborative, Cincinnati, OH, USA.
32. Population, Policy and Practice, UCL Great Ormond Street Institute of Child Health, University College London, UK.
33. ECOGENE-21 and Lipid Clinic, Chicoutimi Hospital, Saguenay, QC, Canada.
34. Department of Biochemistry, Université de Sherbrooke, Sherbrooke, QC, Canada.
35. Department of Primary Care and Public Health, Imperial College London.
36. Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA.
37. Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA.
38. FIMM Institute for Molecular Medicine Finland, Helsinki University Helsinki, FI-00014, Finland.
39. KG Jebsen Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen, Norway.
40. Department of Pediatrics, Haukeland University Hospital, Bergen, 5021, Norway.
41. Department of Genetics and Bioinformatics, Domain of Health Data and Digitalisation, Institute of Public Health, Oslo, Norway.
42. The First Affiliated Hospital of Jinan University, Guangzhou, 510630, China.
43. Institute of Nutritional Science, University of Potsdam, Germany

44. Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, P.O.Box 2040, 3000 CA, Rotterdam, the Netherlands.
45. NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, Southampton, UK.
46. Department of Life Sciences and Biotechnology, Genetic Section, University of Ferrara, Ferrara, Italy.
47. Department of Pediatrics, University of Iowa, Iowa City, Iowa, USA.
48. Research Unit of Obstetrics & Gynecology, Institute of Clinical Research, University of Southern Denmark, Odense, Denmark.
49. Center for Cardiovascular Research / Charité, Berlin, Germany.
50. Subdirección de Salud Pública y Adicciones de Gipuzkoa, Donostia/San Sebastián, Spain.
51. Instituto de Investigación Sanitaria BIODONOSTIA, Donostia/San Sebastián, Spain
52. Department of Epidemiology and Biostatistics, School of Public Health, Medical Research Council-Health Protection Agency Centre for Environment and Health, Faculty of Medicine, Imperial College London, UK
53. Department of Obstetrics and Gynecology, Sahlgrenska Academy, Sahgrenska University Hospital, Gothenburg, Sweden.
54. Department of Social Medicine, University of Crete, Greece.
55. Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain.
56. Human Development & Health, Faculty of Medicine, University of Southampton.
57. Boston University School of Medicine Department of Medicine, Section of General Internal Medicine.
58. Biocenter Oulu, University of Oulu,Oulu, Finland.
59. Unit of Primary Care, Oulu University Hospital, Kajaanintie 50, P.O.Box 20, FI-90220 Oulu, 90029 OYS, Finland.
60. Department of Children and Young People and Families, National Institute for Health and Welfare, Aapistie 1, Box 310, FI-90101 Oulu, Finland.
61. Department of Public Health, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.
62. Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.
63. Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, P.O.Box 2040, 3000 CA, Rotterdam, the Netherlands.
64. Department of Medicine, Stanford University School of Medicine, Stanford, California, USA.
65. Oxford National Institute for Health Research (NIHR) Biomedical Research Centre, Churchill Hospital, Oxford, UK.
66. Division of Endocrinology, Metabolism and Molecular Medicine, Northwestern University Feinberg School of Medicine

.

67. Department of Population Medicine, Harvard Pilgrim Health Care Institute, Harvard Medical School, Boston, MA.
68. Diabetes Center, Massachussetts General Hospital, Boston, MA.
69. Department of Medicine, Universite de Sherbrooke, QC, Canada.
70. Centre for School of Population Health Research, School of Health Sciences, and Sansom Institute, University of South Australia, Adelaide, Australia.
71. South Australian Health and Medical Research Institute, Adelaide, Australia.

Corresponding authors:
Dr. Rachel M. Freathy
University of Exeter Medical School
Royal Devon and Exeter Hospital
Barrack Road
Exeter, UK
Tel: +44 (0)1392 408238
Email: r.freathy@ex.ac.uk

Prof. Debbie A. Lawlor
MRC Integrative Epidemiology Unit at the University of Bristol
Oakfield House,
Oakfield Road,
Bristol, UK
Tel: +44 (0)117 3310096
Email: d.a.lawlor@bristol.ac.uk

Dr. Bjarke Feenstra
Department of Epidemiology Research
Statens Serum Institut
Artillerivej 5
DK-2300 Copenhagen S
Denmark
Tel : +45 3268 3852
Email : fee@ssi.dk

Abstract

Genome-wide association studies (GWAS) of birth weight have focused on fetal genetics, while relatively little is known about how maternal genetic variation influences fetal growth. We aimed to identify maternal genetic variants associated with birth weight that could highlight potentially relevant maternal determinants of fetal growth.

We meta-analysed GWAS data on up to 8.7 million SNPs in up to 86,577 women of European descent from the Early Growth Genetics (EGG) Consortium and the UK Biobank. We used structural equation modelling (SEM) and analyses of mother-child pairs to quantify the separate maternal and fetal genetic effects.

Maternal SNPs at 10 loci (*MTNR1B, HMGA2, SH2B3, KCNAB1, L3MBTL3, GCK, EBF1, TCF7L2, ACTL9* and *CYP3A7*) showed evidence of association with offspring birth weight at $P<5x10^{-8}$. The SEM analyses showed at least 7 of the 10 associations were consistent with effects of the maternal genotype acting via the intrauterine environment, rather than via effects of shared alleles with the fetus. Variants, or correlated proxies, at many of the loci had been previously associated with adult traits, including fasting glucose (*MTNR1B, GCK* and *TCF7L2*) and sex hormone levels (*CYP3A7*), and one (*EBF1*) with gestational duration.

The identified associations indicate effects of maternal glucose, cytochrome P450 activity and gestational duration, and potential effects of maternal blood pressure and immune function on fetal growth. Further characterization of these associations, for example in mechanistic and causal analyses, will enhance understanding of the potentially modifiable maternal determinants of fetal growth, with the goal of reducing the morbidity and mortality associated with low and high birth weights.

Introduction

Individuals with birth weights approaching the lower and upper ends of the population distribution are more at risk of adverse neonatal and later-life health outcomes and mortality than those of average weight (1-5). The factors influencing birth weight involve both maternal and fetal genetic contributions in addition to the environment. Genome-wide association studies (GWAS) testing for common variant effects on *own* birth weight ("fetal" GWAS) have so far identified 60 robustly associated loci (6-8). The influence of common maternal genetic variation on offspring birth weight, beyond the effects of transmitted genetic variation, is poorly understood. Studies estimating the variance in birth weight explained by fetal or maternal genetic factors, using data on twins (9, 10), families (11) or mother-child pairs with genome-wide common variant data (8, 12), have consistently estimated a distinct maternal genetic contribution, which is smaller than the fetal genetic contribution, with estimates ranging from 3% to 22% of the variance explained (relative to 24% to 69% for fetal genetics).

Maternal genotypes may influence key maternal phenotypes, such as circulating levels of glucose and other metabolic factors, which could cross the placenta and affect the growth of the fetus. For example, women with hyperglycemia due to rare heterozygous mutations in the *GCK* gene have babies who are heavier at birth (provided the babies do not inherit the mutation) due to intrauterine exposure to high maternal glucose levels (13). Additionally, maternal genotypes may act upon other maternal attributes, such as vascular function or placental transfer of nutrients, which are also likely to influence fetal growth. Such maternal environmental effects could in turn influence fetal growth separately from the effects of any growth-related genetic variants that are inherited by the fetus directly from the mother (**Figure 1**). Supporting evidence for such effects from analyses of common genetic variants includes positive associations between maternal weighted allele scores for BMI or fasting glucose and offspring birth weight, and an inverse association between a maternal weighted allele score for systolic blood pressure and offspring birth weight (14).

The goal of the current study was to apply a GWAS approach to identify *maternal* genetic variants associated with offspring birth weight. This could potentially highlight novel pathways by which the maternal genotype influences offspring birth weight through the intra-uterine environment. We performed a meta-analysis of GWASs of offspring birth weight using maternal genotypes in up to 86,577 women of European descent from 25 studies, including 37,945 participants from studies collaborating in the Early Growth Genetics (EGG) Consortium and 48,632 participants from the UK Biobank (**Supplementary Figure 1**).

Results

The basic characteristics of study participants in the EGG Consortium discovery, EGG follow-up and UK Biobank GWAS analyses are presented in **Supplementary Tables 1, 2 and 3**, respectively.

Maternal SNPs at 10 loci were associated with offspring birth weight at $P < 5x10^{-8}$

We identified 10 autosomal loci that were associated with offspring birth weight at $P < 5x10^{-8}$ (**Figure 2; Table 1; Supplementary Figures 2 and 3; Supplementary Table 4**). The linkage disequilibrium (LD) score regression intercept (15) from the overall meta-analysis was 1.009, so there was little change in the test statistics after adjusting for this inflation. Three of these loci (*KCNAB1, EBF1* and *CYP3A7*) were identified in UK Biobank data only, and the index SNPs were unavailable in the EGG Consortium data. Consideration of results for proxy SNPs at these three loci from the EGG meta-analysis is below in the next section. For the index SNPs at the other 7 loci, we observed no strong evidence of heterogeneity in allelic effects between the EGG Consortium and UK Biobank components of the meta-analysis (**Supplementary Figure 4; Supplementary Table 4**). The majority of the index SNPs mapped to non-coding sequence and were not in strong LD with any coding variants (r^2<0.95), but the index SNP in *SH2B3*, rs3184504, is a non-synonymous coding variant (R262W). Approximate conditional analysis (see **Materials and Methods**) showed no evidence of secondary signals at any locus at $P < 5x10^{-8}$. In combination, the 10 loci explained 1.4 % (SE = 1.2 %) of variance in birth weight, while the variance in birth weight captured by all autosomal genotyped variants on the UK Biobank array was considerably greater: 11.1 % (SE = 0.6 %).

Birth weight-raising alleles at KCNAB1 and EBF1 were associated with longer gestational duration

The associations at *KCNAB1, EBF1* and *CYP3A7* resulted from analysis of UK Biobank data only, and index SNPs were unavailable in the EGG Consortium meta-analysis (imputed to HapMap Phase 2). To investigate further the evidence for association at these loci, we identified proxy SNPs ($r^2 = 1$) for *KCNAB1* (rs9872556) and *EBF1* (rs2964484) that were available in HapMap Phase 2 (no proxy SNP

was available at $r^2 > 0.5$ at the *CYP3A7* locus). Meta-analysis of the EGG Consortium and UK Biobank data showed weaker evidence of association overall, with some evidence of heterogeneity between the EGG meta-analysis and UK Biobank (*P* = 0.008 and 0.007, respectively; **Table 1**; **Supplementary Table 4**; **Supplementary Figure 4**). In the UK Biobank, women reported the birth weight of their first child, but not the duration of gestation. In contrast, analyses of birth weight in all but one EGG study (QIMR, n=892) were adjusted for the duration of gestation. It is therefore possible that the associations observed with birth weight at *KCNAB1* and *EBF1* in the UK Biobank reflect primary associations with gestational duration. Look-ups of the index SNPs and HapMap 2 proxy SNPs in a published dataset of the top 10000 associated SNPs from a GWAS of gestational duration and preterm birth in 43,568 women (16) showed evidence of association at both *EBF1* ($P < 10^{-12}$) and *KCNAB1* ($P < 10^{-3}$; **Supplementary Table 5**). The birth weight-raising alleles were associated with longer gestational duration.

Five associated SNPs were independent of those identified in previous fetal GWAS of birth weight

The index SNPs at four of the identified loci (*SH2B3, KCNAB1, TCF7L2* and *CYP3A7*), mapped > 2 Mb away from, and were statistically independent of any index SNPs previously associated with birth weight at $P < 5x10^{-8}$ in a fetal GWAS ($r^2 < 0.05$) (8). A summary of candidate genes at these 4 loci is shown in **Supplementary Table 6** (corresponding information for the other loci was reported in (8)). At *MTNR1B* and *HMGA2*, the same index SNP was associated with birth weight in the same direction in both the current study and the previous fetal GWAS. At the four remaining loci, the maternal GWAS index SNPs were within 0.5 to 15 kb of previously reported fetal GWAS index SNPs with very different strengths of pairwise LD between the maternal and fetal GWAS index SNPs. At the *EBF1* and *ACTL9* loci the maternal and fetal GWAS index SNPs were in strong LD (r^2 = 0.95 and 0.99, respectively), and the directions of association were consistent, suggesting that they were tagging the same causal variant. At the *L3MBTL3* locus the maternal and fetal directions of association were consistent, but the index SNPs were weakly correlated (r^2 = 0.13), and conditional analyses in UK

Biobank suggested weakening of the association when the *L3MBTL3* SNP identified in the fetal GWAS was accounted for (unconditional effect (SE) = 0.033 (0.007), P=1.4 x 10^{-6}; conditional effect (SE) =0.024 (0.008), P=1.2 x 10^{-3}; **Supplementary Table 7**). At the *GCK* locus the minor allele frequencies of the maternal and fetal GWAS index SNPs were very different (0.23 and 0.009, respectively), and in low pairwise LD (r^2 = 0.002). Analysis conditional on the fetal GWAS index SNP in UK Biobank did not alter the association at the maternal index SNP (**Supplementary Table 7**), suggesting that at *GCK,* the maternal association with birth weight was distinct from the previously-reported fetal association.

Structural equation modelling applied to UK Biobank data suggested most associations were driven by the maternal genotype.

The partial overlap between associations identified in the current study and those identified in the previous fetal GWAS of birth weight (**Figure 2; Supplementary Figure 2**) illustrates the expected correlation between maternal and fetal genotypes (r ≈ 0.5). The associations between maternal genotype and birth weight identified here may represent indirect effects of maternal genotype on birth weight acting via the maternal intrauterine environment, or primary effects of the fetal genotype on birth weight that are captured (due to correlation) when assaying the maternal genotype, or a mixture of maternal and fetal effects. Analysis of UK Biobank data using structural equation modelling (SEM; n=78,674 male and female unrelated participants, of whom 33,238 individuals only reported their own birth weight, 20,963 women only reported the birth weight of their first child and 24,473 women reported their own birth weight and that of their first child); see **Materials and Methods**) provided estimates of maternal effects adjusted for fetal genotype, and vice versa, and suggested that the associations at the majority of the loci were driven by the maternal genotype (**Figure 3**; **Supplementary Table 8**). In particular, the adjusted maternal effects estimated at 7 of the loci (*MTNR1B, KCNAB1, GCK, EBF1, TCF7L2, ACTL9* and *CYP3A7*) were separated from the adjusted fetal effect estimates by at least 2 standard errors. The only locus at which the

point estimate for the adjusted fetal effect was larger than that of adjusted maternal effect was *HMGA2*, suggesting this association was driven by the fetal genotype. Additional analyses (i) adjusting for fetal genotype in up to 8705 mother-child pairs, and (ii) comparing the unadjusted maternal effect estimates from the overall maternal GWAS (n=up to 86,577) with those from a published fetal GWAS (n=143,677), provided supporting evidence that the majority of the effects were maternally driven (**Supplementary Figure 5; Supplementary Tables 8 and 9**).

Known associations at the identified loci highlighted potentially relevant maternal traits including fasting glucose, blood pressure, immune function and sex hormone levels

Look-ups of index SNPs (N=7 loci), or SNPs in close LD (N=2 loci at $r^2 > 0.9$; N=1 at $r^2=0.4$), in available GWAS datasets for cardiometabolic and growth-related traits revealed several associations at $P < 5x10^{-8}$ (**Supplementary Table 10**), while further information on previously-reported associations was obtained from the NHGRI-EBI catalog of GWAS (see **Materials and Methods**).

The maternal birth weight-associated variants at *MTNR1B, GCK* and *TCF7L2* loci are known to be associated with fasting glucose and Type 2 diabetes susceptibility (17, 18), with the glucose-raising allele associated with higher offspring birth weight.

The C-allele of the missense variant, rs3184504, in *SH2B3*, associated with higher birth weight in our study, has been associated with multiple cardiovascular traits (lower SBP and DBP (19, 20), altered lipid levels and lower risk of coronary artery disease, CAD (21, 22)), altered haematological traits (23-26), lower risk of autoimmune diseases and autoimmune disorders (27-32), lower kynurenine levels (33), higher risk of tonsillectomy (34) and higher risk of colorectal and endometrial cancer (35).

The maternal birth weight-raising allele at *ACTL9* was in LD ($r^2 = 1$) with alleles of nearby variants associated with lower risk of atopic dermatitis (36) and higher risk of tonsillectomy (34).

At the *CYP3A7* locus, a SNP in LD (rs34670419, r2=0.74) with our identified variant, rs45446698, has been associated with levels of the hormones, progesterone and dehydroepiandrosterone sulphate (DHEAS) (37). The maternal birth weight-raising allele was associated with lower hormone levels.

The index SNP, rs12520982, at *EBF1* is in LD ($r^2 = 1$) with rs2963463, recently associated with gestational duration and preterm birth (16). SNPs in weak LD ($r^2 < 0.1$) with *EBF1* rs12520982 were previously associated with blood pressure traits (19, 38, 39), while rs12520982 showed moderate evidence of association with SBP ($P = 1x10^{-5}$) and DBP ($P = 1x10^{-3}$) in the UK Biobank data (**Supplementary Table 10**). There were no known prior associations with the birth weight-associated variation at *KCNAB1*.

The variants at *HMGA2* and *L3MBTL3* have been associated with adult height (40). At *HMGA2,* and possibly also at *L3MBTL3,* the association with birth weight is through the fetal allele, not the maternal allele, so these associations with adult height are relevant for offspring, not mother. Associations at *HMGA2* were additionally observed with other growth and development phenotypes: infant length (41); infant head circumference (42), and primary or permanent tooth eruption (43, 44).

To identify biological pathways underlying maternal regulation of birth weight, we performed gene-set enrichment analysis using MAGENTA (45). Seven pathways reached FDR <0.05, including three involved in the metabolism of xenobiotics (**Supplementary Table 11**).

Discussion

In this study, we have identified variants in the maternal genome at 10 loci that are robustly associated with offspring birth weight. Five of the identified associations are independent of those reported in previous fetal GWAS of birth weight (8), bringing the total of known independent common variant associations with birth weight to 65. Since maternal and fetal genotype are correlated (r = 0.5), loci identified in GWAS of birth weight to date could either represent effects of the maternal genotype, acting via the intrauterine environment, or direct effects of the fetal genotype, or a mixture of the two (**Figure 1**). Our analyses, and those of 58 previously reported loci (46), suggest that while the majority of the 65 known associations indicate direct effects of the fetal genotype, at least 7 associations from the current study (those at *MTNR1B, EBF1, ACTL9, KCNAB1, GCK, TCF7L2* and *CYP3A7,* of which the first 3 were initially identified in fetal GWAS (8)) indicate maternal intrauterine effects.

The index SNP, rs45446698, at the *CYP3A7* locus, is an eQTL for *CYP3A7* in adrenal gland tissue (47). The *CYP3A7* gene is part of the cytochrome P450 family 3 subfamily A gene cluster, which encodes enzymes responsible for the metabolism of multiple and diverse endogenous and exogenous molecules (48), and SNP rs45446698 tags a haplotype of 7 highly correlated variants in the *CYP3A7* promoter, known as the *CYP3A7*1C* allele (49, 50). The *CYP3A7* gene is predominantly expressed in fetal development, but *CYP3A7*1C* results in expression in adult carriers (50, 51). The *CYP3A7*1C* allele, and correlated SNPs, have been associated with circulating levels of DHEAS, progesterone and 2-hydroxylation pathway estrogen metabolites (37, 52, 53). There were no associations between offspring birth weight and maternal SNPs at each of 9 loci (independent of *CYP3A7*) that are also known to influence levels of DHEAS or progesterone (37, 54) (data not shown), suggesting that neither DHEAS nor progesterone levels *per se* are likely to explain the association with birth weight. Since CYP3A enzymes metabolize a diverse range of substrates, there are many possible mechanisms by which maternal *CYP3A7*1C* might be associated with birth weight. In our conditional analysis, we

observed weak evidence of an independent association with the fetal allele at this locus in the opposite direction to that of the maternal allele. Further analyses in larger samples will be required to confirm this and to investigate possible mechanisms underlying this association. However, the association at this locus, together with the results of the gene-set enrichment analysis, which highlighted pathways involved in xenobiotic metabolism, suggests that it is a key avenue for future research into fetal outcomes.

The birth weight-raising maternal alleles at the identified loci (*MTNR1B, GCK* and *TCF7L2*) are strongly associated with higher fasting glucose and Type 2 diabetes in non-pregnant adults (17, 18), and with glycemic traits and gestational diabetes mellitus in pregnant women (55-57). The association between raised maternal glucose and higher offspring birth weight is the result of higher fetal insulin secretion in response to increased placental transfer of glucose (58). Our results confirm previous maternal candidate gene associations with birth weight at *TCF7L2* and *GCK* (55, 59, 60) and demonstrate the key role of maternal glucose levels in influencing offspring birth weight (14, 61). Notably, the Type 2 diabetes risk allele at each of these three loci was not associated with birth weight independently of the maternal allele when present in the fetus. This is contrary to what has been seen at other Type 2 diabetes loci such as *ADCY5* and *CDKAL1,* where risk alleles in the fetus were associated with lower birth weight (8). However, there is an additional low frequency fetal variant at the *GCK* locus, which is independent of the glucose-raising maternal variant associated with higher birth weight in the current study (8). Taken together with the known effects on birth weight of both maternal and fetal rare heterozygous *GCK* mutations (13), a complex picture of allelic variation relevant to fetal growth is emerging at this locus.

The association with birth weight at *HMGA2* was previously identified in a fetal GWAS of birth weight (same index SNP) (8), and our analyses showed that the maternal SNP in our study was probably capturing a direct effect of the SNP in the fetus on skeletal growth, given previous

associations with infant length, head circumference and adult height (40-42). The *L3MBTL3* locus identified in our study is also a known height locus, and the associated variant was correlated (r^2 = 0.13) with a SNP associated with birth weight in the previous fetal GWAS (8). It was less clear from our analyses whether the association at *L3MBTL3* originated from the maternal or fetal genotype. However, analyses of maternal height alleles transmitted to offspring vs. those not transmitted to offspring suggest that the majority of the association between maternal height and offspring birth weight is due to direct effects of fetal inherited alleles (62).

Our exploration of known associations at the remaining 4 loci indicated a number of potentially relevant maternal traits that could influence birth weight via the intrauterine environment, including higher blood pressure (associations at *SH2B3* and suggestive associations at *EBF1*, both between the blood pressure raising maternal allele and lower offspring birth weight), which has been causally associated with lower birth weight in Mendelian randomization analyses (14), and immune function (associations at *SH2B3* and *ACTL9*). However, further studies are needed to elucidate the mechanisms at these loci and at *KCNAB1*, which showed no previous associations with other traits.

We observed weak evidence of heterogeneity of effect sizes between the EGG Consortium and UK Biobank components of our meta-analysis at the *KCNAB1* and *EBF1* loci, which led us to investigate possible explanations. A key difference was that birth weight was adjusted for duration of gestation in the majority of EGG studies, while the duration of gestation was unavailable in the UK Biobank. This raised the possibility that birth weight associations at *KCNAB1* and *EBF1* might arise from a primary effect on gestational duration, i.e. these loci could be primarily influencing the timing of delivery, rather than fetal growth. It is of course possible that the heterogeneity indicated false positive associations in the UK Biobank dataset that were not replicated in the EGG dataset. However, directionally consistent evidence of association with gestational duration and preterm

birth in a recently published GWAS ($P < 5x10^{-8}$ at *EBF1;* $P < 10^{-3}$ at *KCNAB1*) suggests that this is unlikely.

There were some limitations to our study. First, the birth weight of first child was self-reported by mothers in the UK Biobank study, and so was likely subject to more error variation and potential bias than measured birth weight. However, maternal reports of offspring birth weight have been shown to be accurate (63, 64), and we showed that the birth weight of first child variable was associated with maternal smoking, height, BMI and socio-economic position in the expected directions. A second limitation of our study was that by performing a maternal GWAS of birth weight that does not account for the fetal genotype, the analysis was biased against identifying loci at which the fetal genotype exerts opposing effects. Proof-of-principle that such loci exist is demonstrated by the effects on birth weight of rare mutations in the *GCK* gene, which act in opposite directions when present in either mother or fetus, but result in normal birth weight if both mother and fetus inherit the mutation (13). Our analysis conditional on fetal genotype at the 10 loci using a novel method (46) had greatly increased power to resolve maternal vs. fetal effects compared with previous analyses in limited numbers of mother-child pairs (8). While it is not yet computationally feasible to run such an analysis genome-wide, future studies will benefit from considering maternal and fetal genotype simultaneously at the discovery stage and are thereby likely to uncover further loci.

In conclusion, we have identified 10 maternal genetic loci associated with offspring birth weight, 5 of which were not previously identified in fetal GWAS of birth weight, and at least 7 of which represent maternal intrauterine effects. Collectively, the identified associations highlight key roles for maternal glucose and cytochrome P450 activity and potential roles for maternal blood pressure and immune function. Future genetic, mechanistic and causal analyses will be required to characterize such intrauterine effects, leading to greater understanding of the maternal determinants of fetal growth, with the goal of reducing the morbidity and mortality associated with low and high birth weights.

Materials and Methods

Early Growth Genetics (EGG) Consortium discovery studies: genotyping, imputation and GWAS analysis

We studied 19,626 unrelated women of European ancestry from eleven studies with maternal genome-wide genotypes and offspring birth weight available. These included two sub-samples from the 1958 British birth cohort (1958BC-WTCCC2, n = 836; 1958BC-T1DGC, n = 858); the Avon Longitudinal Study of Parents and Children (ALSPAC, n = 7,304); a sub-sample of the Danish National Birth Cohort from the Genetics of Extreme Overweight in Young Adults study (DNBC-GOYA, n = 1,805); population-based controls from a case-control study of pre-term birth in the DNBC (DNBC-PTBCTRLS, n = 1,656); the Hyperglycemia and Adverse Pregnancy Outcome study (HAPO, n = 1,280); the Norwegian Mother and Child cohort study (MoBa, n = 650); the Northern Finland 1966 Birth Cohort study (NFBC1966, n = 2,035); the Netherlands Twin Register (NTR, n = 707); the Queensland Institute of Medical Research study of adult twins (QIMR, n = 892); the Twins UK study (TwinsUK, n = 1,603).

Genotypes in each study were obtained through high-density SNP arrays and up to ~2.5million autosomal SNPs were imputed to HapMap Phase II. Study protocol was approved at each study centre by the local ethics committee and written informed consent had been obtained from all participants and/or their parent(s) or legal guardians. Study descriptions and basic characteristics of samples in the discovery phase are presented in **Supplementary Table 1**.

Within each study, we converted offspring birth weight (BW, grams) to a z-score ((BW value - mean(BW))/ standard deviation(BW)) to allow comparison of data across studies. We excluded multiple births, stillbirths, congenital anomalies (where known), and births before 37 weeks of gestation (where known). We assessed the association between each SNP and offspring birth weight using linear regression of the birth weight z-score against maternal genotype (additive genetic

model), with sex and gestational duration as covariables (gestational duration was unavailable in the QIMR study, which contributed 4.5 % of EGG participants). Ancestry principal components were included as covariables where necessary in the individual studies. Genome-wide association analyses were conducted using PLINK (65), SNPTEST (66), Mach2qtl (67) or Beagle (68) (see **Supplementary Table 1**).

Genome-wide meta-analysis of 11 EGG Consortium discovery studies

Prior to meta-analysis, SNPs with a minor allele frequency (MAF) <0.01 and poorly imputed SNPs (info<0.8 (PLINK), r2hat <0.3 (MACH or Beagle) or proper_info <0.4 (SNPTEST)) were excluded. To adjust for inflation in test statistics generated in each cohort, genomic control (69) was applied once to each individual study (see **Supplementary Table 1** for λ values in each study). Data annotation, exchange and storage were facilitated by the SIMBioMS platform (70). Quality control of individual study results and fixed-effects inverse variance meta-analyses were undertaken by two meta-analysts in parallel at different study centres using the software package METAL (2009-10-10 release) (71). We obtained association statistics for a total of 2,422,657 SNPs in the meta-analysis for which at least 7 of the 11 studies were included. The genomic control inflation factor, λ, in the overall meta-analysis was 1.007.

Follow-up of 18 SNPs in 13 additional EGG Consortium studies

We selected 15 SNPs that surpassed a *P*-value threshold of $P < 1 \times 10^{-5}$ for follow-up in additional, independent studies. Of these, one SNP (rs11020124) was in LD ($r^2 = 0.63$, 1000 Genomes Pilot 1 data) with SNP rs10830963 at the *MTNR1B* locus known to be associated with fasting glucose and Type 2 diabetes (72). We assumed that these represented the same association signal. Given its robust association with maternal glycemic traits likely to impact on offspring birth weight, we took only rs10830963 forward for follow-up at this locus. We then used the National Human Genome Research Institute (NHGRI) catalog of published GWAS (http://www.genome.gov/gwastudies,

accessed September 2012) to query the SNPs associated with birth weight between $P = 10^{-4}$ and $P = 10^{-5}$. We identified three further SNPs at loci with robust evidence ($P < 5 \times 10^{-8}$) of association with other phenotypes, and therefore higher prior odds of association with birth weight: rs2971669 near *GCK* ($r^2 = 0.73$ with rs4607517 associated with fasting glucose)(60); rs204928 in *LMO1* ($r^2 = 0.90$ with rs110419 associated with neuroblastoma)(73) and rs7972086 in *RAD51AP1* ($r^2 = 0.27$ with rs2970818 associated with serum phosphorus concentration)(74). We took forward SNPs rs4607517, rs204928 and rs7972086 for follow-up at these loci, giving a total of 18 SNPs to be examined in additional studies.

The descriptions, genotyping details and basic phenotypic characteristics of the follow-up studies are presented in **Supplementary Table 2**. Of a total of thirteen follow up studies (n = 18,319 individuals), 9 studies (n = 15,288) provided custom genotyping of between 4 and 18 SNPs, while 4 studies (n = 3,031 individuals) had *in silico* genome-wide or exome-wide SNP genotypes available. Where SNPs were imputed, we included only those with quality scores (r2hat or proper_info) >0.8. We excluded directly genotyped SNPs showing evidence of deviation from Hardy-Weinberg Equilibrium at P <0.0028 (Bonferroni corrected for 18 tests). Where genotypes were unavailable for the index SNP, we used the SNP Annotation and Proxy (SNAP) Search Tool to find proxy SNPs based on LD in the 1000 Genomes Pilot 1 dataset ($r^2 > 0.8$; https://www.broadinstitute.org/mpg/snap/ldsearch.php; accessed September 2012; see **Supplementary Table 12**).

Preparation, quality control and genetic analysis in UK Biobank samples

UK Biobank data were available for 502,655 participants, of whom 273,463 were women (75), and of these women, 216,811 reported the birth weight of their first child (in pounds) either at the baseline or follow-up assessment visit. We converted pounds to kg (multiplying by 0.45) for use in our analyses. No information was available on gestational duration or offspring sex. A total of N=64,072 women with offspring birth weight data available also had genotype data available in the May 2015

data release. Women identified as not of British descent (N=9,681) were excluded from the analysis along with those reporting offspring birth weights of < 2.5 kg or > 4.5kg (N=5,479). 'British-descent' was defined as individuals who both self-identified as white British and were confirmed as ancestrally Caucasian using principal components analyses (http://biobank.ctsu.ox.ac.uk). A total of 1,976 of the women were asked to repeat the questionnaire at a follow-up assessment and therefore had two reports of birth weight of first child. Those with differing values were excluded (N=280). This resulted in N=48,632 women with both genotype data and a valid offspring birth weight value, which was z-score transformed for analysis (**Supplementary Table 3**). UK Biobank carried out stringent quality control of the GWAS genotype scaffold prior to imputation up to a reference panel of a combined 1000 Genomes Project Consortium and UK10K Project Consortium. We tested for association with birth weight of first child using a linear mixed model implemented in BOLT-LMM (76) to account for cryptic population structure and relatedness. Genotyping array was included as a binary covariate in the regression model. Total chip heritability (i.e. the variance explained by all autosomal polymorphic genotyped SNPs passing quality control) was calculated using restricted maximum likelihood (REML) implemented in BOLT-LMM (76). We additionally analysed the association between birth weight of first child and directly genotyped SNPs on the X chromosome in 45,445 unrelated women identified by UK Biobank as white British. We excluded SNPs with evidence of deviation from Hardy-Weinberg equilibrium ($P < 1 \times 10^{-6}$), MAF < 0.01 or overall missing rate > 0.015, resulting in 17,352 SNPs for analysis in PLINK v.1.07, with the first 5 ancestry principal components as covariates.

In both the full UK Biobank sample and our refined sample, birth weight of first child was associated with mother's smoking status, maternal BMI and maternal height in the expected directions (**Supplementary Table 3**).

Overall meta-analysis of discovery and follow-up samples

A flowchart of the overall study design is presented in **Supplementary Figure 1**. We performed inverse variance, fixed-effects meta-analysis of the association between each SNP and birth weight z-score in up to 25 discovery and follow-up studies combined (maximum total n = 86,577 women; 8,723,755 SNPs with MAF ≥ 0.01 plus 17,352 X-chromosome SNPs in 45,445 women) using METAL (71). To check for population substructure or relatedness that was not adequately accounted for in the analysis, we examined the intercept value from univariate linkage-disequilibrium score regression (15).

Approximate conditional analysis

At each of the identified loci we looked for the presence of multiple distinct association signals in the region 1Mb up- and down-stream from the lead SNP through approximate conditional analysis. Conditional and joint analysis (COJO) in the analysis program, Genome-wide Complex Trait Analysis (GCTA) (77) was applied to identify secondary signals that attained genome-wide significance ($P < 5 \times 10^{-8}$) using a sample of 10,000 individuals selected at random from the UK Biobank to approximate patterns of LD between variants in these regions.

Candidate gene search

To search for candidate genes at the 4 loci not already covered by the previous fetal GWAS of birth weight (8), we identified the nearest gene, searched PubMed for relevant information on genes within 300kb of the index SNP, and queried the index SNP for eQTL or proxy SNPs ($r^2 > 0.8$) reported from GTEx v4, GEUVADIS, and 11 other studies using Haploreg v4.1 (http://archive.broadinstitute.org/mammals/haploreg/haploreg.php).

Estimating maternal and fetal genetic effects at the identified loci

The full details of the structural equation modelling (SEM) method for estimating the conditional fetal and maternal effects are described elsewhere (46). Briefly, we fitted a structural equation

model to three observed variables from the UK Biobank study; the participant's own self-reported birth weight, the birth weight of the first child reported by the women and the genotype of the participants. Our model included two latent variables; one for the individual's mother (i.e., grand-maternal genotype) and one for the genotype of the participant's offspring. We know these latent variables are correlated on average 50% with the individual's own genotype, hence the path coefficient between each of the latent variables and the observed genotype was set to 0.5. Our model also included residual error terms for the participant's own birth weight and the birth weight of their first child, a covariance parameter to quantify similarity between the error terms, and a variance parameter to model variation in the observed genotype. Using this model, we were able to simultaneously estimate the effect of maternal and fetal genotypes on offspring birthweight
To fit the SEM, we used OpenMx (78) in R (version 3.3.2) (79) with the raw UK Biobank data, and the P-value for the fetal and maternal paths was calculated using a Wald test. We fitted a second SEM without the child and maternal path to conduct a 2 degree of freedom test for the effect of the SNP on birth weight.

Genotype data from the UK Biobank May 2015 release was used for analysis. We included 57,711 participants who reported their own birth weight and 45,436 women who reported the birth weight of their first child, giving a total of a total of 78,674 unique individuals in the analysis (24,473 women had both their own and their offspring's birth weight). Individuals who were non-European or were related to others in the sample, or who were part of multiple births, were excluded. The birth weight of offspring phenotype was prepared as described above, while own birth weight was prepared as described previously (8). We adjusted the individuals' own birth weight for sex, and both birth weight measures for the 12 genetically determined principal components and genotyping batch before creating z-scores for analysis.

We analysed up to 8705 mother-child pairs from 4 studies with both maternal and fetal genotypes available (ALSPAC, EFSOCH, HAPO (non-GWAS) and DNBC-PTBCTRLS). We used linear regression to

test the association between birth weight z-score and maternal genotype conditional on fetal genotype and vice versa (also adjusting analyses for sex and gestational duration). We combined the results from the individual studies using inverse variance meta-analysis with fixed effects. We performed a further meta-analysis to combine the overall estimates with those from the SEM using UK Biobank data.

Look-ups in published GWAS and NHGRI GWAS catalog

We looked up associations between the 10 identified loci and various anthropometric and cardiometabolic traits in available GWAS result sets. The traits and sources are shown in **Supplementary Table 10**. Where the index SNPs at *KCNAB1, EBF1* and *CYP3A7* were unavailable, we used proxies (r^2 = 0.99, 1.00 and 0.41, respectively). Since GWAS summary statistics for blood pressure were not publicly available, we used the UK Biobank May 2015 genetic data release and tested associations between the SNPs and systolic and diastolic blood pressure (SBP and DBP) in 127,968 and 127,776 British descent participants, respectively. Two blood pressure readings were taken approximately 5 minutes apart using an automated Omron blood pressure monitor. Two valid measurements were available for most participants, and the average was taken. Individuals were excluded if the two readings differed by more than 4.56SD, and blood pressure measurements more than 4.56SD away from the mean were excluded. We accounted for blood pressure medication use by adding 15 mmHg to the SBP measure and 10 mmHg to the DBP measure in those reporting regular use of any antihypertensive. Blood pressure was adjusted for age, sex and centre location and then inverse normalized before analysis.

We additionally queried the NHGRI-EBI catalog of published GWAS (http://www.ebi.ac.uk/gwas/home, last accessed 2 August 2017) for associations $P < 5x10^{-8}$ between any additional traits or diseases and SNPs within 500kb of, and in LD with, the index SNP at each locus.

Gene set enrichment analysis

We used Meta-Analysis Gene-set EnrichmeNT of variant Associations (MAGENTA) to test for pathway-based associations using summary statistics from the overall meta-analysis (45). The software mapped each gene to the SNP with the lowest P value within a 110kb upstream and 40kb downstream window. The P value (representing a gene score) was corrected for confounding factors such as gene size, SNP density and LD-related properties in a regression model. Genes within the HLA-region were excluded. Genes were then ranked by their adjusted gene scores. The observed number of gene scores in a given pathway with a ranked score above a given threshold (95th and 75^{th}percentiles) was calculated and this statistic was compared with 1,000,000 randomly permuted pathways of the same size. This generated an empirical *P* value for each pathway, and we considered pathways reaching FDR < 0.05 to be of interest. The 3230 biological pathways tested were from the BIOCARTA, Gene Ontology, Ingenuity, KEGG, PANTHER and REACTOME databases, with a small number of additional custom pathways.

Acknowledgements

We are extremely grateful to the participants and families who contributed to all of the studies and the teams of investigators involved in each one. These include interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. This research has been conducted using the UK Biobank Resource (Application numbers 7036 and 12703). For additional study-specific acknowledgements, please see Supplementary Material.

Funding

Researchers were funded by investment from the European Regional Development Fund (ERDF) and the European Social Fund (ESF) Convergence Programme for Cornwall and the Isles of Scilly [J.T.]; European Research Council (ERC) [grant: SZ-245 50371-GLUCOSEGENES-FP7-IDEAS-ERC to T.M.F., A.R.W.], [ERC Consolidator Grant, ERC-2014-CoG-648916 to V.W.V.J.], [P.R.N.]; University of Bergen, KG Jebsen and Helse Vest [P.R.N.]; Wellcome Trust Senior Investigator Awards [A.T.H. (WT098395), M.I.M. (WT098381)]; National Institute for Health Research (NIHR) Senior Investigator Award (NF-SI-0611-10219); Sir Henry Dale Fellowship (Wellcome Trust and Royal Society grant: WT104150) [R.M.F., R.N.B.]; 4-year studentship (Grant Code: WT083431MF) [R.C.R]; the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement (Grant number 669545; DevelopObese) [D.A.L.]; US National Institute of Health (grant: R01 DK10324) [D.A.L, C.L.R]; Wellcome Trust GWAS grant (WT088806) [D.A.L] and NIHR Senior Investigator Award (NF-SI-0611-10196) [D.A.L]; Wellcome Trust Institutional Strategic Support Award (WT097835MF) [M.A.T.]; The Diabetes Research and Wellness Foundation Non-Clinical Fellowship [J.T.]; Australian National Health and Medical Research Council Early Career Fellowship (APP1104818) [N.M.W.]; Daniel B. Burke Endowed Chair for Diabetes Research [S.F.A.G.]; UK Medical Research Council Unit grants MC_UU_12013_5 [R.C.R, L.P, S.R, C.L.R, D.M.E., D.A.L.] and MC_UU_12013_4 [D.M.E.]; Medical Research Council (grant: MR/M005070/1) [M.N.W., S.E.J.]; Australian Research Council Future Fellowship (FT130101709) [D.M.E] and (FT110100548) [S.E.M.];

NIHR Oxford Biomedical Research Centre (BRC); Oak Foundation Fellowship [B.F.]; FRQS research scholar and Clinical Scientist Award by the Canadian Diabetes Association and the Maud Menten Award from the Institute of Genetics–Canadian Institute of Health Research (CIHR) [MFH]; CIHR - Frederick Banting and Charles Best Canada Graduate Scholarships [C.A.]; FRQS [L.B.]; Netherlands Organization for Health Research and Development (ZonMw –VIDI 016.136.361) [V.W.V.J.]; National Institute on Aging (R01AG29451) [J.M.M.]; 2010-2011 PRIN funds of the University of Ferrara – Holder: Prof. Guido Barbujani, Supervisor: Prof. Chiara Scapoli – and in part sponsored by the European Foundation for the Study of Diabetes (EFSD) Albert Renold Travel Fellowships for Young Scientists, "5 per mille" contribution assigned to the University of Ferrara, income tax return year 2009 and the ENGAGE Exchange and Mobility Program for ENGAGE training funds, ENGAGE project, grant agreement HEALTH-F4-2007-201413 [L.M.]; ESRC (RES-060-23-0011) [C.L.R.]; National Institute of Health Research ([S.D., M.I.M.], Senior Investigator Award (NF-SI-0611-10196) [D.A.L]); Australian NHMRC Fellowships Scheme (619667) [G.W.M]. For study-specific funding, please see **Supplementary Material**.

The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Conflicts of interest: The authors declare no conflicts of interest. D.A.L has received support from Roche Diagnostics and Medtronic for biomarker research unrelated to the work presented here.

Author contributions

Study Design (of individual contributing EGG Consortium studies): D.A.L., H.B., M-F. H., J.F.F., D.I.B., J.P.B., S.F.A.G., H.H., W.L.L., M.G.H., H.M.I., T.I.A.S., X.E., L. S-M., M.K., L.C., S.S., B.J., J.M.M., M.M., J.C.M., C.E.P., T.D.S., B.H., C.L.R., S.E., E.H., V.W.V.J., C.Power, M-R.J., A.T.H.

Sample Collection: D.A.L., S.M.R., H.B., E.K-M., B.F., F.G., M.M., M-F.H., V.W.V.J., M.B., J.P.B., H.H., S.F.A.G., J.W.H., H.M.I., B.H., C.R., C.L.R., T.I.A.S, E.A.N., G.W., L.S-M., B.J., M.K., L.C., S.D., G.W.M., M-R.J., C.Power, E.H., T.D.S., C.E.P, A.T.H.

Genotyping: D.A.L., D.M.E., S.M.R., M-R.J., J.C.M., L.B., J.F.F., J-J.H., H.H., S.F.A.G., J.W.H., M.I.M., L.M., C.L.R., C.Potter, L.P., X.E., M.V., Ø.H., P.R.N., G.W.M., C.E.P., T.D.S., T.M.F, R.M.F.

Statistical Analysis: R.N.B., N.M.W., J.T., R.M.F., R.C.R., D.A.L., D.M.E., E.K-M., L.P., B.F., F.G., C.A., J.F.F., J-J.H., G.Z., L.J.M., M.G.H., D.M.S., M.N., K.L.L., S.E.J., K.S.R., H.Y., A.R.W., M.A.T., A.M., M.N.W., M.H., C.L.R., C.Potter, A.E., S.D., V.H., J.N.P., S.E.M., P.A.L., A.C., D.J.B., R.M., V.S., J.A.M., W.A., S.M., S.J.B.

Writing and overall study direction: R.N.B., N.M.W., A.C., J.T., T.M.F., M-F.H., J.F.F., E.H., W.L.L., D.M.E., D.A.L., B.F. and R.M.F.

All authors reviewed and edited the manuscript.

References

1 Group, H.S.C.R., Metzger, B.E., Lowe, L.P., Dyer, A.R., Trimble, E.R., Chaovarindr, U., Coustan, D.R., Hadden, D.R., McCance, D.R., Hod, M. *et al.* (2008) Hyperglycemia and adverse pregnancy outcomes. *N Engl J Med*, **358**, 1991-2002.
2 Hales, C.N., Barker, D.J., Clark, P.M., Cox, L.J., Fall, C., Osmond, C. and Winter, P.D. (1991) Fetal and infant growth and impaired glucose tolerance at age 64. *BMJ*, **303**, 1019-1022.
3 Harder, T., Rodekamp, E., Schellong, K., Dudenhausen, J.W. and Plagemann, A. (2007) Birth weight and subsequent risk of type 2 diabetes: a meta-analysis. *Am J Epidemiol*, **165**, 849-857.
4 Lawn, J.E., Blencowe, H., Oza, S., You, D., Lee, A.C., Waiswa, P., Lalli, M., Bhutta, Z., Barros, A.J., Christian, P. *et al.* (2014) Every Newborn: progress, priorities, and potential beyond survival. *Lancet*, **384**, 189-205.
5 Risnes, K.R., Vatten, L.J., Baker, J.L., Jameson, K., Sovio, U., Kajantie, E., Osler, M., Morley, R., Jokela, M., Painter, R.C. *et al.* (2011) Birthweight and mortality in adulthood: a systematic review and meta-analysis. *Int J Epidemiol*, **40**, 647-661.
6 Freathy, R.M., Mook-Kanamori, D.O., Sovio, U., Prokopenko, I., Timpson, N.J., Berry, D.J., Warrington, N.M., Widen, E., Hottenga, J.J., Kaakinen, M. *et al.* (2010) Variants in ADCY5 and near CCNL1 are associated with fetal growth and birth weight. *Nat. Genet.*, **42**, 430-435.
7 Horikoshi, M., Yaghootkar, H., Mook-Kanamori, D.O., Sovio, U., Taal, H.R., Hennig, B.J., Bradfield, J.P., St Pourcain, B., Evans, D.M., Charoen, P. *et al.* (2013) New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism. *Nat. Genet.*, **45**, 76-82.
8 Horikoshi, M., Beaumont, R.N., Day, F.R., Warrington, N.M., Kooijman, M.N., Fernandez-Tajes, J., Feenstra, B., van Zuydam, N.R., Gaulton, K.J., Grarup, N. *et al.* (2016) Genome-wide associations for birth weight and correlations with adult disease. *Nature*, **538**, 248-252.
9 Magnus, P. (1984) Further evidence for a significant effect of fetal genes on variation in birth weight. *Clin Genet*, **26**, 289-296.
10 Magnus, P. (1984) Causes of variation in birth weight: a study of offspring of twins. *Clin Genet*, **25**, 15-24.
11 Lunde, A., Melve, K.K., Gjessing, H.K., Skjaerven, R. and Irgens, L.M. (2007) Genetic and environmental influences on birth weight, birth length, head circumference, and gestational age by use of population-based parent-offspring data. *Am J Epidemiol*, **165**, 734-741.
12 Eaves, L.J., Pourcain, B.S., Smith, G.D., York, T.P. and Evans, D.M. (2014) Resolving the effects of maternal and offspring genotype on dyadic outcomes in genome wide complex trait analysis ("M-GCTA"). *Behav Genet*, **44**, 445-455.
13 Hattersley, A.T., Beards, F., Ballantyne, E., Appleton, M., Harvey, R. and Ellard, S. (1998) Mutations in the glucokinase gene of the fetus result in reduced birth weight. *Nat Genet*, **19**, 268-270.
14 Tyrrell, J., Richmond, R.C., Palmer, T.M., Feenstra, B., Rangarajan, J., Metrustry, S., Cavadino, A., Paternoster, L., Armstrong, L.L., De Silva, N.M. *et al.* (2016) Genetic Evidence for Causal Relationships Between Maternal Obesity-Related Traits and Birth Weight. *JAMA*, **315**, 1129-1140.
15 Bulik-Sullivan, B.K., Loh, P.R., Finucane, H.K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics, C., Patterson, N., Daly, M.J., Price, A.L. and Neale, B.M. (2015) LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet*, **47**, 291-295.
16 Zhang, G., Feenstra, B., Bacelis, J., Liu, X., Muglia, L.M., Juodakis, J., Miller, D.E., Litterman, N., Jiang, P.P., Russell, L. *et al.* (2017) Genetic Associations with Gestational Duration and Spontaneous Preterm Birth. *N Engl J Med*.
17 Dupuis, J., Langenberg, C., Prokopenko, I., Saxena, R., Soranzo, N., Jackson, A.U., Wheeler, E., Glazer, N.L., Bouatia-Naji, N., Gloyn, A.L. *et al.* (2010) New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.*, **42**, 105-116.

18 Voight, B.F., Scott, L.J., Steinthorsdottir, V., Morris, A.P., Dina, C., Welch, R.P., Zeggini, E., Huth, C., Aulchenko, Y.S., Thorleifsson, G. *et al.* (2010) Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet*, **42**, 579-589.
19 International Consortium for Blood Pressure Genome-Wide Association, S., Ehret, G.B., Munroe, P.B., Rice, K.M., Bochud, M., Johnson, A.D., Chasman, D.I., Smith, A.V., Tobin, M.D., Verwoert, G.C. *et al.* (2011) Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature*, **478**, 103-109.
20 Levy, D., Ehret, G.B., Rice, K., Verwoert, G.C., Launer, L.J., Dehghan, A., Glazer, N.L., Morrison, A.C., Johnson, A.D., Aspelund, T. *et al.* (2009) Genome-wide association study of blood pressure and hypertension. *Nat Genet*, **41**, 677-687.
21 Dichgans, M., Malik, R., Konig, I.R., Rosand, J., Clarke, R., Gretarsdottir, S., Thorleifsson, G., Mitchell, B.D., Assimes, T.L., Levi, C. *et al.* (2014) Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. *Stroke*, **45**, 24-36.
22 Consortium, C.A.D., Deloukas, P., Kanoni, S., Willenborg, C., Farrall, M., Assimes, T.L., Thompson, J.R., Ingelsson, E., Saleheen, D., Erdmann, J. *et al.* (2013) Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet*, **45**, 25-33.
23 Gieger, C., Radhakrishnan, A., Cvejic, A., Tang, W., Porcu, E., Pistis, G., Serbanovic-Canic, J., Elling, U., Goodall, A.H., Labrune, Y. *et al.* (2011) New gene functions in megakaryopoiesis and platelet formation. *Nature*, **480**, 201-208.
24 Gudbjartsson, D.F., Bjornsdottir, U.S., Halapi, E., Helgadottir, A., Sulem, P., Jonsdottir, G.M., Thorleifsson, G., Helgadottir, H., Steinthorsdottir, V., Stefansson, H. *et al.* (2009) Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat Genet*, **41**, 342-347.
25 van der Harst, P., Zhang, W., Mateo Leach, I., Rendon, A., Verweij, N., Sehmi, J., Paul, D.S., Elling, U., Allayee, H., Li, X. *et al.* (2012) Seventy-five genetic loci influencing the human red blood cell. *Nature*, **492**, 369-375.
26 van Rooij, F.J., Qayyum, R., Smith, A.V., Zhou, Y., Trompet, S., Tanaka, T., Keller, M.F., Chang, L.C., Schmidt, H., Yang, M.L. *et al.* (2017) Genome-wide Trans-ethnic Meta-analysis Identifies Seven Genetic Loci Influencing Erythrocyte Traits and a Role for RBPMS in Erythropoiesis. *Am J Hum Genet*, **100**, 51-63.
27 Barrett, J.C., Clayton, D.G., Concannon, P., Akolkar, B., Cooper, J.D., Erlich, H.A., Julier, C., Morahan, G., Nerup, J., Nierras, C. *et al.* (2009) Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet*, **41**, 703-707.
28 de Boer, Y.S., van Gerven, N.M., Zwiers, A., Verwer, B.J., van Hoek, B., van Erpecum, K.J., Beuers, U., van Buuren, H.R., Drenth, J.P., den Ouden, J.W. *et al.* (2014) Genome-wide association study identifies variants associated with autoimmune hepatitis type 1. *Gastroenterology*, **147**, 443-452 e445.
29 Eriksson, N., Tung, J.Y., Kiefer, A.K., Hinds, D.A., Francke, U., Mountain, J.L. and Do, C.B. (2012) Novel associations for hypothyroidism include known autoimmune risk loci. *PLoS One*, **7**, e34442.
30 Hunt, K.A., Zhernakova, A., Turner, G., Heap, G.A., Franke, L., Bruinenberg, M., Romanos, J., Dinesen, L.C., Ryan, A.W., Panesar, D. *et al.* (2008) Newly identified genetic risk variants for celiac disease related to the immune response. *Nat Genet*, **40**, 395-402.
31 Liu, J.Z., van Sommeren, S., Huang, H., Ng, S.C., Alberts, R., Takahashi, A., Ripke, S., Lee, J.C., Jostins, L., Shah, T. *et al.* (2015) Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet*, **47**, 979-986.
32 Stahl, E.A., Raychaudhuri, S., Remmers, E.F., Xie, G., Eyre, S., Thomson, B.P., Li, Y., Kurreeman, F.A., Zhernakova, A., Hinks, A. *et al.* (2010) Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet*, **42**, 508-514.

33 Shin, S.Y., Fauman, E.B., Petersen, A.K., Krumsiek, J., Santos, R., Huang, J., Arnold, M., Erte, I., Forgetta, V., Yang, T.P. *et al.* (2014) An atlas of genetic influences on human blood metabolites. *Nat Genet*, **46**, 543-550.
34 Pickrell, J.K., Berisa, T., Liu, J.Z., Segurel, L., Tung, J.Y. and Hinds, D.A. (2016) Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet*, **48**, 709-717.
35 Cheng, T.H., Thompson, D., Painter, J., O'Mara, T., Gorman, M., Martin, L., Palles, C., Jones, A., Buchanan, D.D., Win, A.K. *et al.* (2015) Meta-analysis of genome-wide association studies identifies common susceptibility polymorphisms for colorectal and endometrial cancer near SH2B3 and TSHZ1. *Sci Rep*, **5**, 17369.
36 Paternoster, L., Standl, M., Waage, J., Baurecht, H., Hotze, M., Strachan, D.P., Curtin, J.A., Bonnelykke, K., Tian, C., Takahashi, A. *et al.* (2015) Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. *Nat Genet*, **47**, 1449-1456.
37 Ruth, K.S., Campbell, P.J., Chew, S., Lim, E.M., Hadlow, N., Stuckey, B.G., Brown, S.J., Feenstra, B., Joseph, J., Surdulescu, G.L. *et al.* (2016) Genome-wide association study with 1000 genomes imputation identifies signals for nine sex hormone-related phenotypes. *Eur J Hum Genet*, **24**, 284-290.
38 Wain, L.V., Verwoert, G.C., O'Reilly, P.F., Shi, G., Johnson, T., Johnson, A.D., Bochud, M., Rice, K.M., Henneman, P., Smith, A.V. *et al.* (2011) Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. *Nat Genet*, **43**, 1005-1011.
39 Warren, H.R., Evangelou, E., Cabrera, C.P., Gao, H., Ren, M., Mifsud, B., Ntalla, I., Surendran, P., Liu, C., Cook, J.P. *et al.* (2017) Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat Genet*, **49**, 403-415.
40 Wood, A.R., Esko, T., Yang, J., Vedantam, S., Pers, T.H., Gustafsson, S., Chu, A.Y., Estrada, K., Luan, J., Kutalik, Z. *et al.* (2014) Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet*, **46**, 1173-1186.
41 van der Valk, R.J., Kreiner-Moller, E., Kooijman, M.N., Guxens, M., Stergiakouli, E., Saaf, A., Bradfield, J.P., Geller, F., Hayes, M.G., Cousminer, D.L. *et al.* (2015) A novel common variant in DCST2 is associated with length in early life and height in adulthood. *Hum Mol Genet*, **24**, 1155-1168.
42 Taal, H.R., Pourcain, B.S., Thiering, E., Das, S., Mook-Kanamori, D.O., Warrington, N.M., Kaakinen, M., Kreiner-Moller, E., Bradfield, J.P., Freathy, R.M. *et al.* (2012) Common variants at 12q15 and 12q24 are associated with infant head circumference. *Nat Genet*, **44**, 532-538.
43 Fatemifar, G., Hoggart, C.J., Paternoster, L., Kemp, J.P., Prokopenko, I., Horikoshi, M., Wright, V.J., Tobias, J.H., Richmond, S., Zhurov, A.I. *et al.* (2013) Genome-wide association study of primary tooth eruption identifies pleiotropic loci associated with height and craniofacial distances. *Hum Mol Genet*, **22**, 3807-3817.
44 Geller, F., Feenstra, B., Zhang, H., Shaffer, J.R., Hansen, T., Esserlind, A.L., Boyd, H.A., Nohr, E.A., Timpson, N.J., Fatemifar, G. *et al.* (2011) Genome-wide association study identifies four loci associated with eruption of permanent teeth. *PLoS Genet*, **7**, e1002275.
45 Segre, A.V., Consortium, D., investigators, M., Groop, L., Mootha, V.K., Daly, M.J. and Altshuler, D. (2010) Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet*, **6**.
46 Warrington, N.M., Freathy, R.M., Neale, M.C. and Evans, D.M. Using structural equation modelling to jointly estimate maternal and foetal effects on birthweight in the UK Biobank. http://www.biorxiv.org/content/early/2017/07/06/160044.
47 Consortium, G.T. (2015) Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*, **348**, 648-660.
48 Rendic, S. and Di Carlo, F.J. (1997) Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers, and inhibitors. *Drug Metab Rev*, **29**, 413-580.
49 Johnson, N., De Ieso, P., Migliorini, G., Orr, N., Broderick, P., Catovsky, D., Matakidou, A., Eisen, T., Goldsmith, C., Dudbridge, F. *et al.* (2016) Cytochrome P450 Allele CYP3A7*1C Associates

with Adverse Outcomes in Chronic Lymphocytic Leukemia, Breast, and Lung Cancer. *Cancer Res*, **76**, 1485-1493.
50 Kuehl, P., Zhang, J., Lin, Y., Lamba, J., Assem, M., Schuetz, J., Watkins, P.B., Daly, A., Wrighton, S.A., Hall, S.D. *et al.* (2001) Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. *Nat Genet*, **27**, 383-391.
51 Burk, O., Tegude, H., Koch, I., Hustert, E., Wolbold, R., Glaeser, H., Klein, K., Fromm, M.F., Nuessler, A.K., Neuhaus, P. *et al.* (2002) Molecular mechanisms of polymorphic CYP3A7 expression in adult human liver and intestine. *J Biol Chem*, **277**, 24280-24288.
52 Smit, P., van Schaik, R.H., van der Werf, M., van den Beld, A.W., Koper, J.W., Lindemans, J., Pols, H.A., Brinkmann, A.O., de Jong, F.H. and Lamberts, S.W. (2005) A common polymorphism in the CYP3A7 gene is associated with a nearly 50% reduction in serum dehydroepiandrosterone sulfate levels. *J Clin Endocrinol Metab*, **90**, 5313-5316.
53 Sood, D., Johnson, N., Jain, P., Siskos, A.P., Bennett, M., Gilham, C., Busana, M.C., Peto, J., Dos-Santos-Silva, I., Keun, H.C. *et al.* (2017) CYP3A7*1C allele is associated with reduced levels of 2-hydroxylation pathway oestrogen metabolites. *Br J Cancer*, **116**, 382-388.
54 Zhai, G., Teumer, A., Stolk, L., Perry, J.R., Vandenput, L., Coviello, A.D., Koster, A., Bell, J.T., Bhasin, S., Eriksson, J. *et al.* (2011) Eight common genetic variants associated with serum DHEAS levels suggest a key role in ageing mechanisms. *PLoS Genet*, **7**, e1002025.
55 Freathy, R.M., Hayes, M.G., Urbanek, M., Lowe, L.P., Lee, H., Ackerman, C., Frayling, T.M., Cox, N.J., Dunger, D.B., Dyer, A.R. *et al.* (2010) Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study: common genetic variants in GCK and TCF7L2 are associated with fasting and postchallenge glucose levels in pregnancy and with the new consensus definition of gestational diabetes mellitus from the International Association of Diabetes and Pregnancy Study Groups. *Diabetes*, **59**, 2682-2689.
56 Hayes, M.G., Urbanek, M., Hivert, M.F., Armstrong, L.L., Morrison, J., Guo, C., Lowe, L.P., Scheftner, D.A., Pluzhnikov, A., Levine, D.M. *et al.* (2013) Identification of HKDC1 and BACE2 as genes influencing glycemic traits during pregnancy through genome-wide association studies. *Diabetes*, **62**, 3282-3291.
57 Kwak, S.H., Kim, S.H., Cho, Y.M., Go, M.J., Cho, Y.S., Choi, S.H., Moon, M.K., Jung, H.S., Shin, H.D., Kang, H.M. *et al.* (2012) A genome-wide association study of gestational diabetes mellitus in Korean women. *Diabetes*, **61**, 531-541.
58 Pedersen, J. (1952) Diabetes and pregnancy; blood sugar of newborn infants during fasting and glucose administration. *Nord Med*, **47**, 1049.
59 Freathy, R.M., Weedon, M.N., Bennett, A., Hypponen, E., Relton, C.L., Knight, B., Shields, B., Parnell, K.S., Groves, C.J., Ring, S.M. *et al.* (2007) Type 2 diabetes TCF7L2 risk genotypes alter birth weight: a study of 24,053 individuals. *Am. J. Hum. Genet.*, **80**, 1150-1161.
60 Weedon, M.N., Clark, V.J., Qian, Y., Ben-Shlomo, Y., Timpson, N., Ebrahim, S., Lawlor, D.A., Pembrey, M.E., Ring, S., Wilkin, T.J. *et al.* (2006) A common haplotype of the glucokinase gene alters fasting glucose and birth weight: association in six studies and population-genetics analyses. *Am. J. Hum. Genet.*, **79**, 991-1001.
61 Breschi, M.C., Seghieri, G., Bartolomei, G., Gironi, A., Baldi, S. and Ferrannini, E. (1993) Relation of birthweight to maternal plasma glucose and insulin concentrations during normal pregnancy. *Diabetologia*, **36**, 1315-1321.
62 Zhang, G., Bacelis, J., Lengyel, C., Teramo, K., Hallman, M., Helgeland, O., Johansson, S., Myhre, R., Sengpiel, V., Njolstad, P.R. *et al.* (2015) Assessing the Causal Relationship of Maternal Height on Birth Size and Gestational Age at Birth: A Mendelian Randomization Analysis. *PLoS Med*, **12**, e1001865.
63 Jensen, C.B., Gamborg, M., Heitmann, B., Sorensen, T.I. and Baker, J.L. (2015) Comparison of birth weight between school health records and medical birth records in Denmark: determinants of discrepancies. *BMJ Open*, **5**, e008628.

64 Tate, A.R., Dezateux, C., Cole, T.J., Davidson, L. and Millennium Cohort Study Child Health, G. (2005) Factors affecting a mother's recall of her baby's birth weight. *Int J Epidemiol*, **34**, 688-695.
65 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559-575.
66 Marchini, J., Howie, B., Myers, S., McVean, G. and Donnelly, P. (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.*, **39**, 906-913.
67 Li, Y., Willer, C.J., Ding, J., Scheet, P. and Abecasis, G.R. (2010) MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet. Epidemiol.*, **34**, 816-834.
68 Browning, B.L. and Browning, S.R. (2007) Efficient multilocus association testing for whole genome association studies using localized haplotype clustering. *Genet. Epidemiol.*, **31**, 365-375.
69 Devlin, B. and Roeder, K. (1999) Genomic control for association studies. *Biometrics*, **55**, 997-1004.
70 Krestyaninova, M., Zarins, A., Viksna, J., Kurbatova, N., Rucevskis, P., Neogi, S.G., Gostev, M., Perheentupa, T., Knuuttila, J., Barrett, A. *et al.* (2009) A System for Information Management in BioMedical Studies--SIMBioMS. *Bioinformatics*, **25**, 2768-2769.
71 Willer, C.J., Li, Y. and Abecasis, G.R. (2010) METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, **26**, 2190-2191.
72 Prokopenko, I., Langenberg, C., Florez, J.C., Saxena, R., Soranzo, N., Thorleifsson, G., Loos, R.J., Manning, A.K., Jackson, A.U., Aulchenko, Y. *et al.* (2009) Variants in MTNR1B influence fasting glucose levels. *Nat. Genet.*, **41**, 77-81.
73 Wang, K., Diskin, S.J., Zhang, H., Attiyeh, E.F., Winter, C., Hou, C., Schnepp, R.W., Diamond, M., Bosse, K., Mayes, P.A. *et al.* (2011) Integrative genomics identifies LMO1 as a neuroblastoma oncogene. *Nature*, **469**, 216-220.
74 Kestenbaum, B., Glazer, N.L., Kottgen, A., Felix, J.F., Hwang, S.J., Liu, Y., Lohman, K., Kritchevsky, S.B., Hausman, D.B., Petersen, A.K. *et al.* (2010) Common genetic variants associate with serum phosphorus concentration. *J. Am. Soc. Nephrol.*, **21**, 1223-1232.
75 Allen, N.E., Sudlow, C., Peakman, T., Collins, R. and Biobank, U.K. (2014) UK biobank data: come and get it. *Sci Transl Med*, **6**, 224ed224.
76 Loh, P.R., Tucker, G., Bulik-Sullivan, B.K., Vilhjalmsson, B.J., Finucane, H.K., Salem, R.M., Chasman, D.I., Ridker, P.M., Neale, B.M., Berger, B. *et al.* (2015) Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet*, **47**, 284-290.
77 Yang, J., Ferreira, T., Morris, A.P., Medland, S.E., Genetic Investigation of, A.T.C., Replication, D.I.G., Meta-analysis, C., Madden, P.A., Heath, A.C., Martin, N.G. *et al.* (2012) Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet*, **44**, 369-375, S361-363.
78 Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., Spies, J., Estabrook, R., Kenny, S., Bates, T. *et al.* (2011) OpenMx: An Open Source Extended Structural Equation Modeling Framework. *Psychometrika*, **76**, 306-317.
79 Ihaka, R. and Gentleman, R. (1996) R: a language for data analysis and graphics. *Journal of computational and graphical statistics*, **5**, 299-314.

Figure legends

Figure 1. A schematic diagram illustrating that maternal genetic factors may influence fetal growth indirectly through the intra-uterine environment, or directly through inheritance by the fetus.

Figure 2. Manhattan plot of associations between 8,723,755 maternal autosomal SNPs and 17,352 maternal X-chromosome SNPs (all MAF >1%) and offspring birth weight from the meta-analysis of up to 86,577 women. SNP position across the chromosomes (x-axis) and results of association tests between maternal genotype and offspring birth weight adjusted for sex and, where available, gestational duration ($-\log_{10}$ *P*-value; left y-axis) are shown. The index maternal SNP from the current study and all SNPs within 500kb of that SNP are highlighted either in red, or in purple. Those in purple indicate loci at which the index maternal SNP from the current study was within 500k of an index SNP associated previously with own birth weight (i.e. in a "fetal GWAS of birth weight") at $P < 5\times10^{-8}$ (8). SNPs within 500kb of the 54 other index SNPs previously identified in the fetal GWAS of birth weight are highlighted in green. The red, horizontal line indicates a *P* value of 5×10^{-8}.

Figure 3. Independent maternal and fetal effects on birth weight at the 10 identified loci, estimated from a meta-analysis of results using a structural equation model (SEM) in unrelated UK Biobank participants with results using conditional analysis in maternal-fetal pairs. All SNPs are aligned to the birth weight-raising allele reported in Table 1. The colour of each dot indicates the maternal genetic association P-value for birth weight, adjusted for the fetal genetic association: red, $P < 0.0001$; orange, $0.0001 \leq P < 0.001$; yellow, $0.001 \leq P < 0.05$; white $P \geq 0.05$.

Figure 1

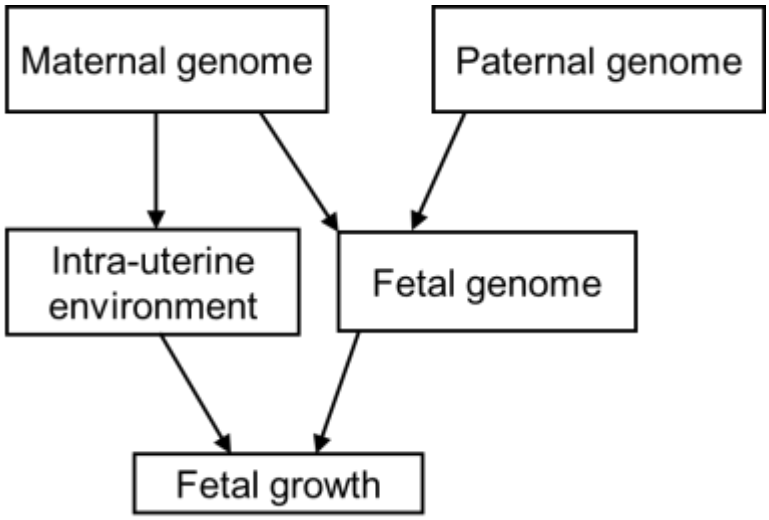
Maternal genome
Paternal genome
Intra-uterine environment
Fetal genome
Fetal growth

Figure 2

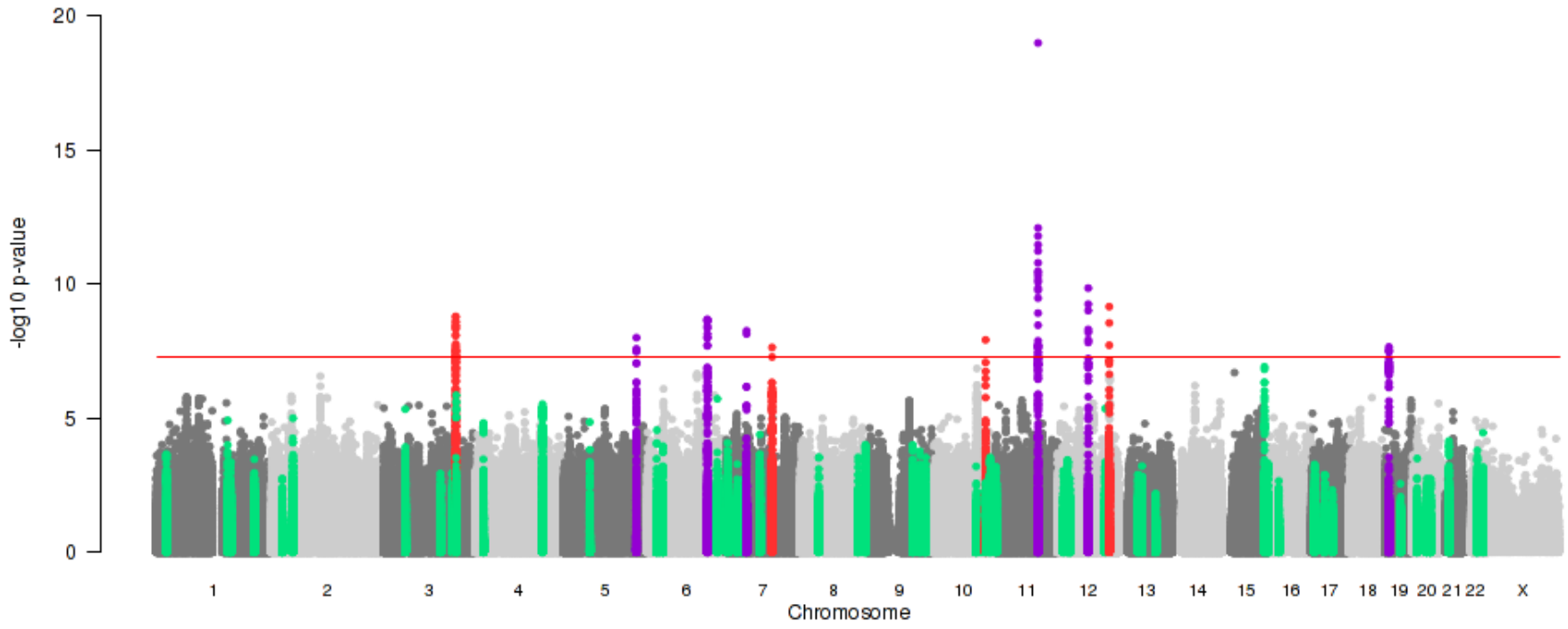
20
15
10
5
0
-log10 p-value
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Chromosome

Figure 3

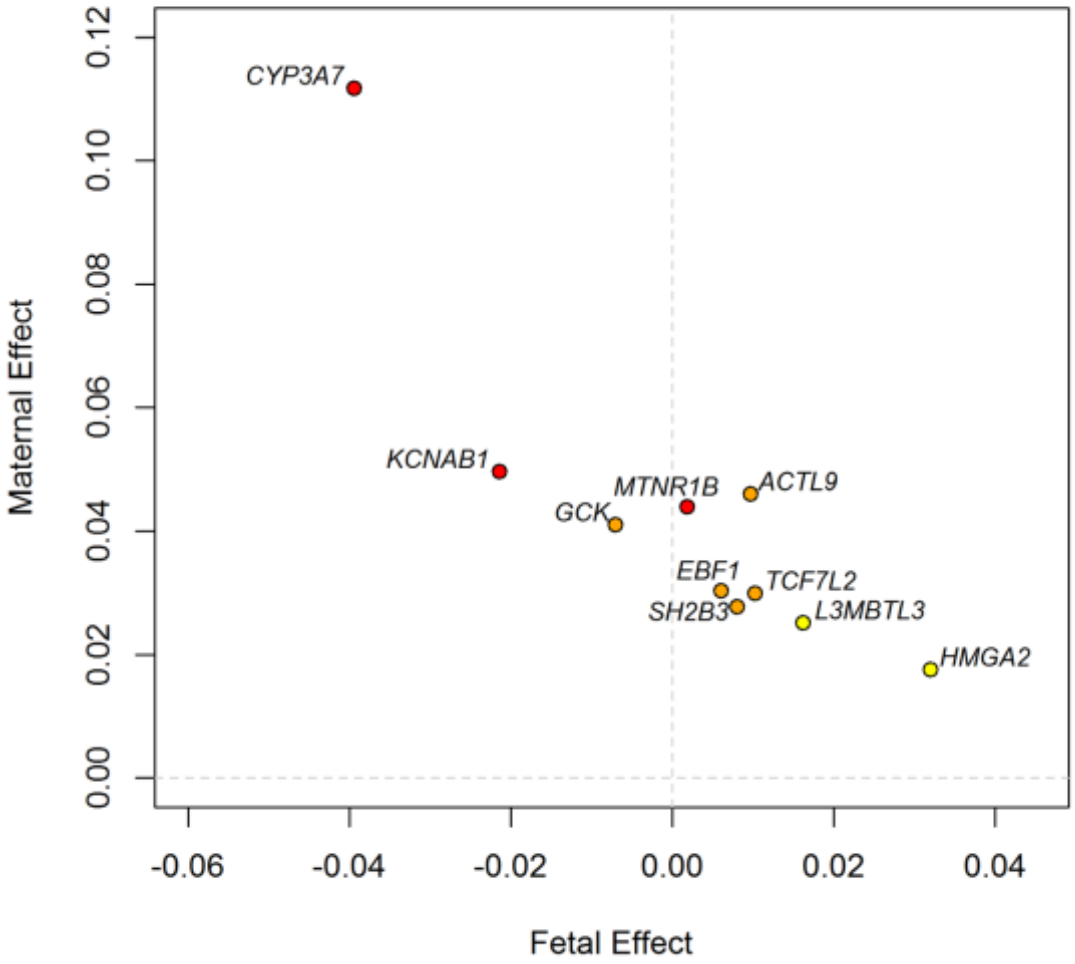
CYP3A7
KCNAB1
MTNR1B
ACTL9
GCK
EBF1
TCF7L2
SH2B3
L3MBTL3
HMGA2
Maternal Effect
0.00
0.02
0.04
0.06
0.08
0.10
0.12
-0.06
-0.04
-0.02
0.00
0.02
0.04
Fetal Effect

Table 1. Ten maternal genetic loci associated with offspring birth weight ($P < 5 \times 10^{-8}$) in a European ancestry meta-analysis of up to 86,577 women.

Locus	**Lead SNP**	**Chr**	**Position (bp, b37)**	**Alleles**	**EAF**	**EGG + UKBB meta-analysis**		
				Effect/Other		**Effect (SE)**	***P*-value**	***N***
MTNR1B	rs10830963	11	92708710	G/C	0.29	0.052 (0.006)	1.0×10^{-19}	71341
HMGA2	rs1351394	12	66351826	T/C	0.49	0.034 (0.005)	1.4×10^{-10}	68247
SH2B3	rs3184504	12	111884608	C/T	0.52	0.033 (0.005)	6.9×10^{-10}	68249
*KCNAB1**	rs7629460	3	155829938	C/A	0.59	0.039 (0.007)	1.6×10^{-9}	48632
KCNAB1†	rs9872556	3	155829855	T/C	0.57	0.029 (0.005)	8.1×10^{-8}	68253
L3MBTL3	rs9375694	6	130356608	G/A	0.31	0.035 (0.006)	2.1×10^{-9}	68223
GCK	rs2971669	7	44231778	T/C	0.23	0.038 (0.007)	5.5×10^{-9}	68162
*EBF1**	rs12520982	5	157894747	T/C	0.74	0.041 (0.007)	9.9×10^{-9}	48632
EBF1†	rs2964484	5	157897437	A/G	0.71	0.031 (0.006)	4.6×10^{-7}	67547
TCF7L2	rs7903146	10	114758349	T/C	0.30	0.034 (0.006)	1.2×10^{-8}	68253
ACTL9	rs2918299	19	8787273	C/T	0.83	0.041 (0.007)	2.2×10^{-8}	67603
*CYP3A7**	rs45446698	7	99332948	G/T	0.04	0.089 (0.016)	2.4×10^{-8}	48632

Effects were aligned to the birth weight-raising allele and are presented in SD units (1 SD of birth weight = 484g (7)). Chr, chromosome; bp, base pair; b37, build 37; EAF, effect allele frequency; SE, standard error.
*Index SNPs available in UK Biobank only.
†Results for proxy SNPs ($r^2 = 1$) at *KCNAB1* and *EBF1* available in both UK Biobank and EGG (imputed to HapMap Phase 2) are shown below the index SNPs, with full details in **Supplementary Table 4** (in both cases there was weak evidence of heterogeneity of effect size between EGG and UK Biobank, $P < 0.01$). No proxy for rs45446698 at *CYP3A7* with $r^2 > 0.5$ was available in HapMap Phase 2.