nature research

Peer Review Information

Journal: Nature Genetics
Manuscript Title: Distinction between the effects of parental and fetal genomes on fetal growth
Corresponding author name(s): Dr. Valgerdur Steinthorsdottir, Dr. Kari Stefansson

Reviewer Comments & Decisions:

Decision Letter, initial version:

3rd February 2021

Dear Dr. Steinthorsdottir,

Your Article "Distinction between the effects of parental and fetal genomes on fetal growth" has been seen by two referees. You will see from their comments below that, while they find your work of interest, they have raised several relevant points. We are very interested in the possibility of publishing your study in Nature Genetics, but we would like to consider your response to these points in the form of a revised manuscript before we make a final decision on publication.

To guide the scope of the revisions, the editors discuss the referee reports in detail within the team, including with the chief editor, with a view to identifying key priorities that should be addressed in revision, and sometimes overruling referee requests that are deemed beyond the scope of the current study. In this case, we ask that you address all technical queries with clarifications and revisions to the text and display items where appropriate and that you make the full genome-wide association summary statistics publicly available. We hope you will find this prioritized set of referee points to be useful when revising your study. Please do not hesitate to get in touch if you would like to discuss these issues further.

We therefore invite you to revise your manuscript taking into account all reviewer and editor comments. Please highlight all changes in the manuscript text file. At this stage we will need you to upload a copy of the manuscript in MS Word .docx or similar editable format.

We are committed to providing a fair and constructive peer-review process. Do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

When revising your manuscript:

*1) Include a "Response to referees" document detailing, point-by-point, how you addressed each referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the referees along with the revised manuscript.

*2) If you have not done so already please begin to revise your manuscript so that it conforms to our Article format instructions, available
here.
Refer also to any guidelines provided in this letter.

*3) Include a revised version of any required Reporting Summary:
https://www.nature.com/documents/nr-reporting-summary.pdf
It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review.
A revised checklist is essential for re-review of the paper.

Please be aware of our guidelines on digital image standards.

Please use the link below to submit your revised manuscript and related files:

[REDACTED]

Note: This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

We hope to receive your revised manuscript within 4-8 weeks. If you cannot send it within this time, please let us know.

Please do not hesitate to contact me if you have any questions or would like to discuss these revisions further.

Nature Genetics is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit please visit www.springernature.com/orcid.

We look forward to seeing the revised manuscript and thank you for the opportunity to review your work.

Sincerely,
Kyle

Kyle Vogan, PhD
Senior Editor
Nature Genetics
https://orcid.org/0000-0001-9565-9665

Referee expertise:

Referee #1: Genetics, complex traits, statistical methods

Referee #2: Genetics, complex traits, fetal development

Reviewers' Comments:

Reviewer #1:
Remarks to the Author:

The authors' report a large scale GWAS meta-analysis of birthweight and several associated perinatal traits (birth length, ponderal index, gestational age) in the DECODE cohort and UK Biobank/EGG consortium. This is a great paper and a major step forward in elucidating the genetics of prenatal traits like birthweight. This manuscript differs from previous efforts in this area (Warrington et al. 2019 Nat Genet; Horikoshi et al. 2016 Nature) in that the authors are able to use genealogical information within the DECODE cohort to phase their sample and partition genetic effects on birthweight into maternal untransmitted, maternal transmitted and paternal transmitted haplotype effects. The authors subsequently cluster markers into categories based on each SNP's estimated effects on the phenotypes and find evidence for parent of origin effects at several loci. The authors also examine the genetic correlation between several traits and birthweight and amongst other things provide evidence that blood pressure and birthweight correlate because of genetic pleiotropy in the offspring genome as opposed to the maternal genome.

The paper is very well written, the methods appropriate for the most part (although see comments below about claiming POEs and definitive comments regarding mode of transmission etc), and represents a major advance in understanding the genetics of complex perinatal traits. I only have relatively minor comments for improvement of an impressive paper.

Comments:

I may be incorrect, but as far as I can determine, the authors classify the mechanism of action at the different SNPs according to a parametric based clustering analysis, as opposed to say formal tests of statistical significance comparing different models of inheritance to each other. P values referring to the significance of individual allelic effects do not offer straight forward interpretations because of the confounding of maternal/paternal and offspring genetic effect. In other words I see these results as descriptive rather than definitive. Genetic effects are small in general, their standard errors large, and even with a sample the size of DECODE, the power to formally discriminate between the possible different models of inheritance is likely to be small for many loci. Assuming my interpretation is correct, I wonder whether this fact needs to be emphasized more in the discussion?

p. 3 "they used a newly developed structural modelling approach" – it would be good to cite the original paper Warrington et al (2018) Int J Epidemiol in which the approach was developed alongside the Warrington et al. 2019 Nature Genetics paper.

p. 3 "we performed an expanded GWAS meta-analyses" should be analysis.

P6. "One of the challenges in analysing genetic effect" should be an s on effects

-Could the authors put the standard errors of the effects in Supplementary Table 1?

-Gene names in Supplementary Tables should be in italics

-In ST7 birthweight is spelled wrong in the header

-Some of the labels in Figure 3 are a little confusing e.g. "maternal only effect" for variant rs231848 - I prefer the nomenclature in ST 8 that refer to these as "maternally transmitted only effect". Same comment for Paternal only effect.

p. 7 "Of the 243 fetal growth variants 22 show pattern of parent-of-origin effect (should be "show a parent of origin effect pattern) with the paternally and maternally transmitted alleles having different effect (should be effects)". Isn't this analysis confounded by maternal effects so a straight comparison between the two is not a valid test for POEs?

p. 8 "We have previously shown that this variant associates without parent of origin specific effects with reduced risk of type 2 diabetes and with greater height and higher BMI in adults". You can't say anything about direction of the effect without making clear the allele.

-In the header to ST9 I don't understand "(iv) a joint model including 3 PRS based on maternal transmitted and non-transmitted alleels and paternal transmitted and non-transmitted alleles" Isn't this 4 parameters?

-The results in ST9 and Figure 5 do not seem to match up (the p values different)?

-Why were GRS constructed using UKBB for things like T2D and CAD? I would have thought that you would have got more power by using GRS constructed using consortium data where cases are oversampled?

-In ST10 it would be nice to have a column classifying the BW SNPs

-p 11. "However, the maternally non-transmitted PRS associated positively with PI (although the signal was weak), indicating that an offspring that inherits height increasing alleles is born longer and leaner, whereas it is born heavier if maternally non-transmitted". This sentence could be written more clearly.

p.10 "A positive effect of the BMI PRS on BW was mostly seen through the maternal genome indicating that any effect of BMI PRS is through the intrauterine environment". But the PRS was not strongly associated with the untransmitted maternal score?

P.13. "that show a parent of origin specific effect" Again is it possible to say this definitively given no formal test for POE has been conducted? Perhaps a toning down of the language is appropriate.

p.14 "this does exclude…" Do you mean does not exclude? The Warrington et al (2019) GWAS utilized the UKBB and cohorts part of the EGG consortium. The authors did not have access to large numbers of parent-offspring pairs and trios to enable phasing of genotypes and consequently were unable to fit the sorts of models using transmitted and untransmitted alleles described in the present paper. Also being largely reliant on the UKBB, the authors were not able to easily condition on gestational age. I can see how a maternal effect of BP on birthweight through gestational age could explain the differences between Warrington et al and the present study. However, Tyrell et al (2016) reported an effect of maternal GRS for BP on BW after conditioning for gestational age. I'm therefore at a bit of a loss to explain the differences across all these studies. Perhaps the authors could also reference Tyrell

et al and speculate?

p.16 "we applied the method of LD score regression". What reference panel was used to index LD? I assume that this was an Icelandic panel to take into account the different LD structure in this population? Perhaps state this here?

p. 17 "a likelihood ratio test". Maybe this is obvious but I don't know how this is constructed given the data. I would like to see this in the author's response for my own knowledge. If it's obvious then I am fine with them not providing further details in the manuscript.

p.20 "the correlation between the PRS". Is this a correlation or an association?

Supplementary Note
-In the formula for the distribution of yi , g_i should be g_i transpose I think?

-In the third line of the formula for the log-likelihood are you missing a parenthesis to indicate that the summation is also over the last two terms (and also multiplied by one over sigma squared)?

-Has the matrix "G" been defined correctly? Should Q and Q transpose appear in the formula?

Finally, the manuscript uses publicly available data from the UK Biobank as well as summary results statistics provided by the EGG Consortium on their website. It would be nice if the authors could return the favour and upload the summary results from this study onto a website in the spirit of collegiality and collaboration.

David Evans.

Reviewer #2:
Remarks to the Author:

Juliusdottir, Steinthorsdottir and colleagues have carried out comprehensive GWAS and follow-up analyses of birth weight, length and ponderal index. This work, which I enjoyed reading, makes a valuable contribution to the field. The study has not only identified additional birth weight loci, but has analysed birth length and ponderal index, and crucially has used phased haplotype data in a large, well-powered sample to estimate effects of parental transmitted and non-transmitted alleles. The study has identified likely parent-of-origin effects and classified the likely origin of association at a substantial number of loci, and then used the phased data to investigate relationships with adult traits. My comments are mainly minor and are mostly concerned with clarifying the information presented.

1. The authors present a maternal, paternal and fetal GWAS of birth weight, but only fetal GWAS of birth length and ponderal index. Is there a reason for not running maternal / paternal GWAS analyses of birth length and ponderal index? It is possible that these would identify further loci.

2. In this study, birth weight (in addition to length and ponderal index), has been used as a proxy for fetal growth, but these parameters can only give limited information about the intrauterine growth trajectory. I suggest the authors consider using a more specific term, such as "birth size", rather than "fetal growth", to refer to the phenotype under consideration. Similarly, "birth size traits" is preferable to "BW traits" in the results subheading.

3. In the methods, it states that prior to association analysis the measurements were adjusted for sex, age, year of birth and gestational age. In what sense are they adjusted for "age" here in addition to gestational age?

4. In the results it would be helpful to state whether birth weight was adjusted for sex and gestational age in the discovery meta-analysis. Although it is mentioned in the methods when describing the preparation of the Icelandic dataset, I think it is important to state in the results section to aid interpretation. The EGG&UKBiobank portion of the meta-analyses of birth weight would mostly have been unadjusted for gestational age due to birth weight in UK Biobank being self-reported, which might explain why some of the identified loci appeared to influence gestational duration.

5. Related to the above point, I have some queries about Supplementary Figure 1, which was cited to support the following sentence in the Results section: "We tested the effect of the associated variants on gestational age (GA) and found that only a fraction of the variants affected GA but adjusting BW for GA increased the significance of most variants (Supplementary Fig. 1)."

- In (a) and (b), what model was tested to generate the GA effects on the Y-axis (e.g. is it linear regression of gestational duration against SNP? – were there any covariates in the model)? And in what samples was this tested? – Icelandic only?

- Similarly, what samples were analysed to generate the effect estimates on the X-axis? – I initially assumed these were the effect estimates for BW from the overall meta-analysis, but since some of the SNPs do not have X error bars (meaning they do not have p<5e-04), that cannot be the case.

- Since the EGG&UKBiobank study was mostly not adjusted for gestational duration, while the current study was able to adjust for gestational duration, then it would make more sense for all plots to be generated from effects in Icelandic samples only, with full control over what is / is not adjusted for gestational duration. I realise that was the case for (c) and (d), but the legend suggests it is not the case for (a) and (b). It would be helpful to have this clarified.

I agree with the statement in the results about a small number of variants affecting GA, and that adjustment increased the significance of most associations, but I think there is a great opportunity to use these data to go a little further in commenting about the relationship between BW and GA at the identified variants. For example, how do the authors interpret the difference in slopes of the fitted lines between the maternal and fetal analyses (a) and (b) in supplementary fig 1? – how much can this tell us about the relative involvement of maternal and fetal genetics in growth vs. timing of delivery? And can the authors comment on the loci that seem to have a positive effect on birth weight and negative effect on gestational duration (YKT6, STK17A and others) vs. those that lie above the solid line (e.g. KCNAB1)? Can this inform about the likely causal relationships between BW>>GA or GA>>BW at these loci? (in the section on glycemic traits, the authors later allude to the fact that faster growing fetus would lead to shorter gestation, and it might be relevant to highlight that here too)

6. It is not clear why reference 2 (Zeng and Zhou, Front Genet) is cited in the introduction following, "It is not clear whether these relationships are causal, the consequence of confounding factors …or through shared genetics". That study used a Mendelian randomization approach and concluded support for causal relationships. However, the analyses in that study did not take into account the relationship between maternal and fetal genotypes. For careful consideration of the appropriate use of such methods, see https://wellcomeopenresearch.org/articles/2-11 or https://pubmed.ncbi.nlm.nih.gov/33272351/ . If the authors intend to cite evidence to support a causal relationship between lower birth weight and later life outcomes in humans, then studies of human populations at times of famine (e.g. Dutch Hungerwinter) are the most convincing.

7. Description of the EGG paper "230,069 with their birth weight also matched with their maternal genotype". I suggest to remove the word "also" as it implies that their own genotype was also available.

8. Figure 1 is a very nice overview of the GWAS results. I think it would be helpful to add the total N for each of the 5 GWAS analyses in brackets because it would more clearly indicate the power differences between them and help with interpretation of the relative numbers and sizes of peaks.

9. Supplementary tables: the previously published data are indicated as coming from the EGG Consortium, but this should be EGG + UK Biobank as the latter added substantial power in that study.

10. Figure 5: please clarify if BW and BL are each adjusted for gestational age, or if gestational age is adjusted for BW?

11. In the Discussion: "Our data indicate that variants that associate with adult T2D through low insulin secretion already affect insulin secretion in the developing fetus, leading to compromised insulin response and reduced fetal growth," is true, but there are clearly variants that only appear to have an effect through the maternal non-transmitted allele. There are other glycemic trait variants that were not identified with effects on birth weight in this study. In the light of the overall data, this should be rephrased to indicate that there is heterogeneity of effects on fetal insulin secretion: while some loci already affect insulin secretion in the fetus, this is not the case for all T2D loci.

12. Second paragraph of page 14: "This does exclude maternal effect" – I think the authors meant to say "This does not exclude…"

13. In the discussion, the authors discuss the finding that a PRS for higher blood pressure was associated with lower BW through the fetal genome, but with less evidence for the maternal genome, and how this is consistent with a smaller study using phased haplotypes (Chen et al, PLoS Med), but inconsistent with findings from the large EGG&UKBiobank GWAS. Since the smaller study (as with the current study under review) found an effect of maternal blood pressure on shortening GA, and since the previous large GWAS was unadjusted for GA, could that explain some of the discrepancy? It might be a helpful point to add to this part of the discussion.

14. A table of basic characteristics (mean and SD birth weight, gestational duration, smoking etc) of study participants included in the new analyses would be helpful.

Author Rebuttal to Initial comments

Responses to comments from reviewers:

We thank the reviewers for their complements on our manuscript and for their excellent helpful comments that will allow us to further improve our manuscript.

Please find our responses to individual comments in bold below.

Reviewer #1:
Remarks to the Author:

The authors' report a large scale GWAS meta-analysis of birthweight and several associated perinatal traits (birth length, ponderal index, gestational age) in the DECODE cohort and UK Biobank/EGG consortium. This is a great paper and a major step forward in elucidating the genetics of prenatal traits like birthweight. This manuscript differs from previous efforts in this area (Warrington et al. 2019 Nat Genet; Horikoshi et al. 2016 Nature) in that the authors are able to use genealogical information within the DECODE cohort to phase their sample and partition genetic effects on birthweight into maternal untransmitted, maternal transmitted and paternal transmitted haplotype effects. The authors subsequently cluster markers into categories based on each SNP's estimated effects on the phenotypes and find evidence for parent of origin effects at several loci. The authors also examine the genetic correlation between several traits and birthweight and amongst other things provide evidence that blood pressure and birthweight correlate because of genetic pleiotropy in the offspring genome as opposed to the maternal genome.

The paper is very well written, the methods appropriate for the most part (although see comments below about claiming POEs and definitive comments regarding mode of transmission etc), and represents a major advance in understanding the genetics of complex perinatal traits. I only have relatively minor comments for improvement of an impressive paper.

Comments:

I may be incorrect, but as far as I can determine, the authors classify the mechanism of action at the different SNPs according to a parametric based clustering analysis, as opposed to say formal tests of statistical significance comparing different models of inheritance to each other. P values referring to the significance of individual allelic effects do not offer straight forward interpretations because of the confounding of maternal/paternal and offspring genetic effect. In other words I see these results as descriptive rather than definitive. Genetic effects are small in general, their standard errors large, and even with a sample the size of DECODE, the power to formally discriminate between the possible different models of inheritance is likely to be small for many loci. Assuming my interpretation is correct, I wonder whether this fact needs to be emphasized more in the discussion?

The reviewer is correct that our approach to classify the mode of association is a clustering method rather than a hypothesis testing method. Indeed, the four clusters of variants do not form a nested hierarchy of models that would easily lend themselves to hypothesis testing. Again, we agree with the reviewer that classification of variants with small effects is challenging, indeed we attempted classification of the 141 with the largest effects out of 243 variants. The strength of evidence of

classification varies substantially between markers depending on how strong the effects of the variants are and how far from other clusters the most likely cluster is. For 61 variants the probability of the most likely cluster was over 0.9, and it is therefore likely that most of these were accurately clustered. For the remaining variants the clustering is less accurate. To discuss these issues we have added the following sentence to the Discussion in the revised manuscript: "Our ability to accurately classify variants depends on their effect sizes and on how different their most likely cluster is from other clusters. For the variants with the greatest classification probabilities, the classification is most likely accurate, while for the variants with more ambiguous classification, more data is required for determination of which cluster they truly belong."

p. 3 "they used a newly developed structural modelling approach" – it would be good to cite the original paper Warrington et al (2018) Int J Epidemiol in which the approach was developed alongside the Warrington et al. 2019 Nature Genetics paper.

We have added the missing Warrington et al (2018) reference in the revised manuscript.

p. 3 "we performed an expanded GWAS meta-analyses" should be analysis.

This has now been corrected.

P6. "One of the challenges in analysing genetic effect" should be an s on effects

This has now been corrected in the revised manuscript.

-Could the authors put the standard errors of the effects in Supplementary Table 1?

We have updated our Supplementary tables to include standard errors in the revised manuscript.

-Gene names in Supplementary Tables should be in italics

This has now been amended in the revised manuscript.

-In ST7 birthweight is spelled wrong in the header

This has now been corrected.

-Some of the labels in Figure 3 are a little confusing e.g. "maternal only effect" for variant rs231848 - I prefer the nomenclature in ST 8 that refer to these as "maternally transmitted only effect". Same comment for Paternal only effect.

We have updated the labels in Figure 3 as suggested.

p. 7 "Of the 243 fetal growth variants 22 show pattern of parent-of-origin effect (should be "show a parent of origin effect pattern) with the paternally and maternally transmitted alleles having different effect (should be effects)". Isn't this analysis confounded by maternal effects so a straight comparison between the two is not a valid test for POEs?

We did not mean to imply that the statement "with the paternally and maternally transmitted alleles having different effects" was in itself evidence for parent-of-origin effect but rather the outcome of such an effect. The evidence for the parent-of-origin effect was presented in the previous section in the manuscript and based on the effects of the maternally and paternally transmitted alleles as well as the effect of the maternally non-transmitted allele. We have now modified the sentence to make this clearer: "Of the 243 fetal growth variants 22 show a parent-of-origin effect pattern, with the paternally and maternally transmitted alleles having different effects in the absence of an effect of the maternally non-transmitted allele".

p. 8 "We have previously shown that this variant associates without parent of origin specific effects with reduced risk of type 2 diabetes and with greater height and higher BMI in adults". You can't say anything about direction of the effect without making clear the allele.

We thank the reviewer for pointing out this error. We have now replaced "this variant" with "the BW raising allele rs76895963-G" in the revised manuscript.

-In the header to ST9 I don't understand "(iv) a joint model including 3 PRS based on maternal transmitted and non-transmitted alleels and paternal transmitted and non-transmitted alleles" Isn't this 4 parameters?

As the reviewer point out this should have read four PRS. This has now been corrected in ST9.

-The results in ST9 and Figure 5 do not seem to match up (the p values different)?

We thank the reviewer for pointing out this error. Unfortunately an older version of Figure 5 was included in the manuscript. A correct version of Figure 5 is now in the revised manuscript.

-Why were GRS constructed using UKBB for things like T2D and CAD? I would have thought that you would have got more power by using GRS constructed using consortium data where cases are oversampled?

We agree with the reviewer that we would have gained power by using consortium data. However, summary statistics generated by large consortia on common complex human traits often include Icelandic datasets and this is the case for both T2D (DIAGRAM consortium) and CAD (CARDIOGRAM).

So to avoid confounding we chose to generate risk scores based on UKBB data where no Icelandic samples are included.

-In ST10 it would be nice to have a column classifying the BW SNPs

In the revised manuscript we have added a column to Supplementary table 10 with information on clustering of variants.

-p 11. "However, the maternally non-transmitted PRS associated positively with PI (although the signal was weak), indicating that an offspring that inherits height increasing alleles is born longer and leaner, whereas it is born heavier if maternally non-transmitted". This sentence could be written more clearly.

We agree with the reviewer. We have rephrased this text in the revised manuscript. "There was a negative correlation between the height PRS and PI if maternally and paternally transmitted indicating that a child that inherits height increasing alleles is born longer and leaner. Conversely, the maternally non-transmitted height PRS associated positively with PI (although the signal was weak), indicating that height increasing alleles in the mother result in children that are slightly heavier relative to BL (Supplementary Fig. 6)".

p.10 "A positive effect of the BMI PRS on BW was mostly seen through the maternal genome indicating that any effect of BMI PRS is through the intrauterine environment". But the PRS was not strongly associated with the untransmitted maternal score?

The effect of the BMI PRS on BW is weak and therefore difficult to interpret. We agree with the reviewer that it is not strongly associated with the maternally non-transmitted alleles. However, we based our interpretation on two points: 1) there is no effect detected for the paternally transmitted allele while we see an effect of the maternally transmitted allele, 2) the data sets for mothers and fathers are of similar size and while we see a weak effect in the father the effect in the mother is both stronger and much more significant. We have now in the revised manuscript modified the text from: "A positive effect of the BMI PRS on BW was mostly seen through the maternal genome indicating that any effect of BMI PRS is through the intrauterine environment" to: "We see a weak positive effect of the BMI PRS on BW but it is unclear to what extent the effect is through the maternal or fetal genomes (Fig. 5)".

P.13. "that show a parent of origin specific effect" Again is it possible to say this definitively given no formal test for POE has been conducted? Perhaps a toning down of the language is appropriate.

Several variants have significant effects on BW when transmitted maternally, but do not show significant effects when not transmitted or when transmitted paternally. Similarly, several variants have significant effects on BW when transmitted paternally, but not otherwise. Absence of evidence is not evidence of absence, but it is still clear that the effects of these variants are far greater when

transmitted from the specific parent. No statistical test will help here because no statistical test can show that a variant does not associate through some mode of transmission – statistical tests quantify evidence for deviation from a null hypothesis, but not for the null hypothesis being true. We agree with the reviewer that the language should be toned down somewhat and have changed the statement in the Discussion from "show a parent-of-origin specific effect" to "have an effect consistent with being parent-of-origin specific".

p.14 "this does exclude…" Do you mean does not exclude? The Warrington et al (2019) GWAS utilized the UKBB and cohorts part of the EGG consortium. The authors did not have access to large numbers of parent-offspring pairs and trios to enable phasing of genotypes and consequently were unable to fit the sorts of models using transmitted and untransmitted alleles described in the present paper. Also being largely reliant on the UKBB, the authors were not able to easily condition on gestational age. I can see how a maternal effect of BP on birthweight through gestational age could explain the differences between Warrington et al and the present study. However, Tyrell et al (2016) reported an effect of maternal GRS for BP on BW after conditioning for gestational age. I'm therefore at a bit of a loss to explain the differences across all these studies. Perhaps the authors could also reference Tyrell et al and speculate?

We thank the reviewer for pointing out this unfortunate error. "not" has now been added to the text. We appreciate the luxury, compared to previous studies, of having access to data with complete information on gestational age and also with parent of origin assigned to each haplotype. It seems clear that the effect of BP risk scores on gestational age explains at least some of the discrepancy between our study and that of Warrington et al. where gestational age was not accounted for. However, other factors may also contribute. Looking at the Tyrrell study it seems that even though they find association between the BP risk score in the maternal genotype and BW adjusted for gestational age this effect is reduced and barely significant after accounting for the fetal genotype (see Figure 2 and Table 3). Unfortunately, they did not have complete information on the fetal genomes so the adjusted analysis was less powered. Nevertheless the results they present are not convincing in terms of the effect of the maternal BP risk score on BW adjusted for gestational age, after accounting for the fetal genome.

We have added a reference to the Tyrrell paper to our text as well as speculations on the discrepancies between studies which may include other aspects of the study design in addition to adjustment for gestational age. This may include differences in the risk scores used (ranging from using 68 SBP associated variants to a full polygenic risk score) as well as differences in the way the fetal genotype is accounted for.

"The association of blood pressure raising alleles in the maternal genome with GA may at least partly explain the observed effect on BW when not adjusting for GA[12]. While Tyrrell et al. reported an effect of a maternal blood pressure risk score on BW adjusted for GA this was at least partly dependent on

the fetal genotype[6]. There are several differences in the design of the studies that may contribute to different outcomes while addressing the effect of SBP PRS on BW. This includes difference in the design of the SBP PRS and how GA and the fetal genotypes are accounted for. Importantly, here we were able to adjust for GA and test directly the effect of transmitted and non-transmitted alleles".

p.16 "we applied the method of LD score regression". What reference panel was used to index LD? I assume that this was an Icelandic panel to take into account the different LD structure in this population? Perhaps state this here?

When performing the LD score regression we used an LD score estimated from the European ancestry (EUR) samples in the 1000 Genomes Project. It has been shown that EUR reference panel is adequate for studies of northern European ancestry, and the intercept might be slightly biased upwards: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4495769/. We have now stated this in the revised manuscript on p 17: "using an LD score estimated from the European ancestry samples in the 1000 Genomes Project".

p. 17 "a likelihood ratio test". Maybe this is obvious but I don't know how this is constructed given the data. I would like to see this in the author's response for my own knowledge. If it's obvious then I am fine with them not providing further details in the manuscript.

We have added a section titled "Meta-Analysis" to the Supplementary Note with the details behind this statement and a reference to this section to the Methods in the revised manuscript.

p.20 "the correlation between the PRS". Is this a correlation or an association?

We agree with the reviewer that association is a more appropriate term in this case and this has been modified in the revised version of the manuscript.

Supplementary Note
-In the formula for the distribution of yi , g_i should be g_i transpose I think?

We prefer g_i to be a column vector and have changed the definition of g_i to be on its transpose for clarity in the revised manuscript. This should make the formulation clear and consistent.

-In the third line of the formula for the log-likelihood are you missing a parenthesis to indicate that the summation is also over the last two terms (and also multiplied by one over sigma squared)?

We have added the parenthesis to the third line as suggested.

-Has the matrix "G" been defined correctly? Should Q and Q transpose appear in the formula?

The definition of G is correct, but the Qs in the fourth line, where the expected log-likelihood is stated in terms of gamma and G, should have gone out. We have fixed this error in the revised manuscript.

Finally, the manuscript uses publicly available data from the UK Biobank as well as summary results statistics provided by the EGG Consortium on their website. It would be nice if the authors could return the favour and upload the summary results from this study onto a website in the spirit of collegiality and collaboration.

As indicated in the Data availability statement the GWAS summary statistics from this study will be made available on the deCODE website at publication.

Reviewer #2:
Remarks to the Author:

Juliusdottir, Steinthorsdottir and colleagues have carried out comprehensive GWAS and follow-up analyses of birth weight, length and ponderal index. This work, which I enjoyed reading, makes a valuable contribution to the field. The study has not only identified additional birth weight loci, but has analysed birth length and ponderal index, and crucially has used phased haplotype data in a large, well-powered sample to estimate effects of parental transmitted and non-transmitted alleles. The study has identified likely parent-of-origin effects and classified the likely origin of association at a substantial number of loci, and then used the phased data to investigate relationships with adult traits. My comments are mainly minor and are mostly concerned with clarifying the information presented.

1. The authors present a maternal, paternal and fetal GWAS of birth weight, but only fetal GWAS of birth length and ponderal index. Is there a reason for not running maternal / paternal GWAS analyses of birth length and ponderal index? It is possible that these would identify further loci.

Although not mentioned in the manuscript we did run maternal and paternal GWAS analyses for birth length and ponderal index. However, these analyses did not yield any new signals. We therefore chose not to include them as they would only distract from the main focus of the paper that was on the birth weight analysis where we were able to combine our data with large public GWAS data.

2. In this study, birth weight (in addition to length and ponderal index), has been used as a proxy for fetal growth, but these parameters can only give limited information about the intrauterine growth trajectory. I suggest the authors consider using a more specific term, such as "birth size", rather than "fetal

growth", to refer to the phenotype under consideration. Similarly, "birth size traits" is preferable to "BW traits" in the results subheading.

The reviewer rightly points out that we are by no means covering all aspects of fetal growth in our analyses. In the absence of a better alternative we have used the term "fetal growth" to describe the traits that we are working on, i.e. birth weight, birth length, ponderal index and to some extent also gestational age. We feel that the term "birth size" does not completely cover birth weight and other traits that have been adjusted for gestational age and may be misleading as we are not working with the actual size of the newborn. We do, however, agree that the term "BW traits" is not very clear and have removed it from the text on p. 4 and p. 20.

3. In the methods, it states that prior to association analysis the measurements were adjusted for sex, age, year of birth and gestational age. In what sense are they adjusted for "age" here in addition to gestational age?

In the fetal analysis we adjusted for sex, year of birth and gestational age. In the maternal and paternal analysis we also adjusted for parental age. We have now clarified this in the Methods and added: "and in the case of the parents also age".

4. In the results it would be helpful to state whether birth weight was adjusted for sex and gestational age in the discovery meta-analysis. Although it is mentioned in the methods when describing the preparation of the Icelandic dataset, I think it is important to state in the results section to aid interpretation. The EGG&UKBiobank portion of the meta-analyses of birth weight would mostly have been unadjusted for gestational age due to birth weight in UK Biobank being self-reported, which might explain why some of the identified loci appeared to influence gestational duration.

We have now made this clear in the Results section p. 4 in the revised manuscript: "The Icelandic GWAS data was adjusted for gestational age (GA) but not the EGG/UKB consortium data where information on GA was only available for a small subset".

5. Related to the above point, I have some queries about Supplementary Figure 1, which was cited to support the following sentence in the Results section: "We tested the effect of the associated variants on gestational age (GA) and found that only a fraction of the variants affected GA but adjusting BW for GA increased the significance of most variants (Supplementary Fig. 1)."

- In (a) and (b), what model was tested to generate the GA effects on the Y-axis (e.g. is it linear regression of gestational duration against SNP? – were there any covariates in the model)? And in what samples was this tested? – Icelandic only?

Yes, it is linear regression of gestational duration against SNP in the Icelandic data only, adjusted for sex and YOB in the fetal analysis, and for sex, YOB and maternal age in the GA offspring mothers analysis. We have now added the following text to the Supplementary Fig. 1 legend "Effect estimates for GA were obtained by linear regression of gestational age against fetal growth variants in the Icelandic data".

- Similarly, what samples were analysed to generate the effect estimates on the X-axis? – I initially assumed these were the effect estimates for BW from the overall meta-analysis, but since some of the SNPs do not have X error bars (meaning they do not have p<5e-04), that cannot be the case.

The effect estimates on the X-axis were from the two BW meta-analyses (children and mothers) for all the 243 fetal growth variants, hence note that not all the variants have p-values below 5e-04 in each of the BW and offspring BW mothers meta-analyses. They were selected as fetal growth variants if they were significant in one of the five meta-analyses.

- Since the EGG&UKBiobank study was mostly not adjusted for gestational duration, while the current study was able to adjust for gestational duration, then it would make more sense for all plots to be generated from effects in Icelandic samples only, with full control over what is / is not adjusted for gestational duration. I realise that was the case for (c) and (d), but the legend suggests it is not the case for (a) and (b). It would be helpful to have this clarified.

We agree with the reviewer that limiting the BW association data to the Icelandic data that has been adjusted for GA gives us "cleaner" data. We have now updated Supplementary Figure 1 and all plots are now based on Icelandic data only. In particular, the BW effect estimates in plots a) and b) are now based on the Icelandic fetal and maternal BW GWAS data and hence adjusted for GA. We now see more clearly that the 3 variants that are associated with GA ($p<5x10^{-8}$) have little effect on BW that has been adjusted for GA. In fact for the *COL27A1* variant (p=0.22 in the Icelandic maternal data) the BW increasing allele in the Icelandic data is the opposite of that in the corresponding meta-analysis.

I agree with the statement in the results about a small number of variants affecting GA, and that adjustment increased the significance of most associations, but I think there is a great opportunity to use these data to go a little further in commenting about the relationship between BW and GA at the identified variants. For example, how do the authors interpret the difference in slopes of the fitted lines between the maternal and fetal analyses (a) and (b) in supplementary fig 1? – how much can this tell us about the relative involvement of maternal and fetal genetics in growth vs. timing of delivery? And can the authors comment on the loci that seem to have a positive effect on birth weight and negative effect on gestational duration (YKT6, STK17A and others) vs. those that lie above the solid line (e.g. KCNAB1)? Can this inform about the likely causal relationships between BW>>GA or GA>>BW at these loci? (in the section on glycemic traits, the authors later allude to the fact

that faster growing fetus would lead to shorter gestation, and it might be relevant to highlight that here too)

It is important to keep in mind that most of the fetal growth variants are not associated with GA in the Icelandic data and in fact as shown in Supplementary Fig. 1, for most of them the association p-value is > 0.05. It is, therefore, difficult to read much into this data for individual variants.

6. It is not clear why reference 2 (Zeng and Zhou, Front Genet) is cited in the introduction following, "It is not clear whether these relationships are causal, the consequence of confounding factors …or through shared genetics". That study used a Mendelian randomization approach and concluded support for causal relationships. However, the analyses in that study did not take into account the relationship between maternal and fetal genotypes. For careful consideration of the appropriate use of such methods, see https://wellcomeopenresearch.org/articles/2-11 or https://pubmed.ncbi.nlm.nih.gov/33272351/ . If the authors intend to cite evidence to support a causal relationship between lower birth weight and later life outcomes in humans, then studies of human populations at times of famine (e.g. Dutch Hungerwinter) are the most convincing.

We thank the reviewer for pointing out those excellent papers. We have updated our references in the revised manuscript.

7. Description of the EGG paper "230,069 with their birth weight also matched with their maternal genotype". I suggest to remove the word "also" as it implies that their own genotype was also available.

For clarity we have removed the word "also" from the text on p. 3 in the revised manuscript.

8. Figure 1 is a very nice overview of the GWAS results. I think it would be helpful to add the total N for each of the 5 GWAS analyses in brackets because it would more clearly indicate the power differences between them and help with interpretation of the relative numbers and sizes of peaks.

We thank the reviewer for an excellent suggestion and have updated Figure 1 accordingly.

9. Supplementary tables: the previously published data are indicated as coming from the EGG Consortium, but this should be EGG + UK Biobank as the latter added substantial power in that study.

We have updated the Supplementary tables accordingly in the revised manuscript.

10. Figure 5: please clarify if BW and BL are each adjusted for gestational age, or if gestational age is adjusted for BW?

BW and BL are both adjusted for GA. Since our default is to adjust for GA in our Icelandic data, we only highlight it when we do not adjust for GA (eg $BW_{undajusted}$ as for example in Supplementary figure 6). We did not adjust for BW in our GA analyses. We have now added for clarity "BW and BL were also adjusted for GA" in the legend for Figure 5.

11. In the Discussion: "Our data indicate that variants that associate with adult T2D through low insulin secretion already affect insulin secretion in the developing fetus, leading to compromised insulin response and reduced fetal growth," is true, but there are clearly variants that only appear to have an effect through the maternal non-transmitted allele. There are other glycemic trait variants that were not identified with effects on birth weight in this study. In the light of the overall data, this should be rephrased to indicate that there is heterogeneity of effects on fetal insulin secretion: while some loci already affect insulin secretion in the fetus, this is not the case for all T2D loci.

We agree with the reviewer that the effect of variants that associate with T2D on BW is a story in itself and individual variants could have been discussed at great length. However, going into this in detail is beyond the scope of this paper. To be clear we would not expect all variants that associate with T2D to affect insulin secretion in the fetus. T2D risk variants are heterogeneous and many of them are unlikely to affect insulin secretion to any extent or in some cases only in older individuals. Furthermore, many of the reported T2D variants have very small effects so any effect on BW might be too small to detect given current sample sizes.

We are far from understanding how individual variants confer risk of T2D. Attempts have been made to cluster T2D risk variants based on their action. Most notably, Mahajan et al. 2018 clustered T2D risk variants into six clusters based on how they affect various metabolic and anthropometric traits. Two of their clusters are reported to have a main effect on insulin secretion. We note some similarities between those two clusters, Insulin secretion 1 and Insulin secretion 2 (Supplementary Figure 6, Mahajan et al. 2018) and variants that associate with BW in our study. In particular, the Insulin secretion 1 cluster includes variants at loci such as *GCK*, *MTNR1B* and *TCF7L2* where we find a positive maternal effect on BW in the absence of a negative fetal effect. On the other hand in the Insulin secretion 2 cluster we find variants where the T2D risk allele has a negative fetal effect on BW, such as at *CCND2*, *CDKAL1*, *HHEX* and *KCNQ1*. This may reflect some basic difference between those two clusters. However, there are also variants such as at *ADCY5* that do not follow this pattern. This led us to simply claim that "The mode of action of individual variants, acting through the maternal and/or fetal genome, may give important clues to the underlying pathophysiology".

We have now expanded this paragraph on p. 14 in the Discussion: "Our data indicate that variants that associate with adult T2D through low insulin secretion at least in some cases already affect insulin secretion in the developing fetus, leading to compromised insulin response and reduced fetal growth. In contrast, there are variants that strongly associate with glycemic traits where the T2D risk allele in the mother associates with higher BW in the absence of a negative effect of the fetal allele".

12. Second paragraph of page 14: "This does exclude maternal effect" – I think the authors meant to say "This does not exclude…"

We thank the reviewer for pointing out this error "not" has now been added to the Discussion on p. 14.

13. In the discussion, the authors discuss the finding that a PRS for higher blood pressure was associated with lower BW through the fetal genome, but with less evidence for the maternal genome, and how this is consistent with a smaller study using phased haplotypes (Chen et al, PLoS Med), but inconsistent with findings from the large EGG&UKBiobank GWAS. Since the smaller study (as with the current study under review) found an effect of maternal blood pressure on shortening GA, and since the previous large GWAS was unadjusted for GA, could that explain some of the discrepancy? It might be a helpful point to add to this part of the discussion.

We agree with the reviewer that the effect of maternal risk of blood pressure on shortening GA is likely to affect the outcome of analyzing the effect of the risk score on BW that is unadjusted for GA. Other factors may also contribute to this discrepancy. We have added those points to the Discussion as outlined in the response regarding this issue to reviewer 1.

14. A table of basic characteristics (mean and SD birth weight, gestational duration, smoking etc) of study participants included in the new analyses would be helpful.

We have now added mean and SD for birth weight and birth length and median and IQR for gestational duration to the description of the Icelandic study population on p.16: "The mean BW was 3715 grams (s.d. 492), mean BL was 51.9 cm (s.d. 2.2) and the median GA was 280 days (inter quartile range 13)". We do not have information on maternal smoking.

Decision Letter, first revision:

Our ref: NG-A56560R

5th April 2021

Dear Dr. Steinthorsdottir,

Your revised manuscript "Distinction between the effects of parental and fetal genomes on fetal growth" (NG-A56560R) has been seen by the original referees. As you will see from their comments below, they are satisfied with the revision and therefore we will be happy in principle to publish it in Nature Genetics, pending final revisions to comply with our editorial and formatting guidelines.

We are now performing detailed checks on your paper and will send you a checklist detailing our editorial and formatting requirements in about a week. Please do not upload the final materials and make any revisions until you receive this additional information from us.

Thank you again for your interest in Nature Genetics Please do not hesitate to contact me if you have any questions.

Sincerely,
Kyle

Kyle Vogan, PhD
Senior Editor
Nature Genetics
https://orcid.org/0000-0001-9565-9665

Reviewer #1 (Remarks to the Author):

I am happy with the authors' revisions.

PS. I think strictly speaking there might be a term missing in the description of the log-likelihood in the Supplementary Note (i.e. n/2log(sigma^2) which of course cancels out in the log-likelihood ratio tests) but happy for this to be introduced in the proof stage to not delay acceptance.

Reviewer #2 (Remarks to the Author):

The authors have responded thoughtfully to all of my previous questions and comments, and have made changes that have improved this already excellent manuscript. I have no further comments, except to say that this is an extremely impressive study, which provides a major advance in our understanding of the genetics of fetal growth.

Rachel Freathy

Author Rebuttal, first revision:

Response to reviewer 1

Reviewer #1:
Remarks to the Author:

I am happy with the authors' revisions.

PS. I think strictly speaking there might be a term missing in the description of the log-likelihood in the Supplementary Note (i.e. n/2log(sigma^2) which of course cancels out in the log-likelihood ratio tests)

but happy for this to be introduced in the proof stage to not delay acceptance.

Response: The log likelihood should be kept as is, since sigma is not a parameter in the model and the log likelihood is determined up to a constant.

Final Decision Letter:

In reply please quote: NG-A56560R1 Steinthorsdottir

11th June 2021

Dear Valgerdur,

I am delighted to say that your manuscript "Distinction between the effects of parental and fetal genomes on fetal growth" has been accepted for publication in an upcoming issue of Nature Genetics.

Prior to setting your manuscript, we may make minor changes to enhance the lucidity of the text and with reference to our house style. We therefore ask that you examine the proofs most carefully to ensure that we have not inadvertently altered the sense of your text in any way.

Once your manuscript is typeset and you have completed the appropriate grant of rights, you will receive a link to your electronic proof via email with a request to make any corrections within 48 hours. If, when you receive your proof, you cannot meet this deadline, please inform us at rjsproduction@springernature.com immediately.

Your paper will be published online after we receive your corrections and will appear in print in the next available issue. You can find out your date of online publication by contacting the Nature Press Office (press@nature.com) after sending your e-proof corrections. Now is the time to inform your Public Relations or Press Office about your paper, as they might be interested in promoting its publication. This will allow them time to prepare an accurate and satisfactory press release. Include your manuscript tracking number (NG-A56560R1) and the name of the journal, which they will need when they contact our Press Office.

Before your paper is published online, we will be distributing a press release to news organizations worldwide, which may very well include details of your work. We are happy for your institution or funding agency to prepare its own press release, but it must mention the embargo date and Nature Genetics. Our Press Office may contact you closer to the time of publication, but if you or your Press Office have any enquiries in the meantime, please contact press@nature.com.

Acceptance is conditional on the data in the manuscript not being published elsewhere, or announced in the print or electronic media, until the embargo/publication date. These restrictions are not intended to deter you from presenting your data at academic meetings and conferences, but any enquiries from the media about papers not yet scheduled for publication should be referred to us.

Please note that Nature Genetics is a Transformative Journal (TJ). Authors may publish their research with us through the traditional subscription access route or make their paper immediately open access through payment of an article-processing charge (APC). Authors will not be required to make a final decision about access to their article until it has been accepted. Find out more about Transformative Journals

<B>Authors may need to take specific actions to achieve compliance with funder and institutional open access mandates. For submissions from January 2021, if your research is supported by a funder that requires immediate open access (e.g. according to Plan S principles) then you should select the gold OA route, and we will direct you to the compliant route where possible. For authors selecting the subscription publication route our standard licensing terms will need to be accepted, including our self-archiving policies. Those standard licensing terms will supersede any other terms that the author or any third party may assert apply to any version of the manuscript.

Please note that Nature Research offers an immediate open access option only for papers that were first submitted after 1 January 2021.

You will not receive your proofs until the publishing agreement has been received through our system.

If you have any questions about our publishing options, costs, Open Access requirements, or our legal forms, please contact ASJournals@springernature.com

If you have posted a preprint on any preprint server, please ensure that the preprint details are updated with a publication reference, including the DOI and a URL to the published version of the article on the journal website.

To assist our authors in disseminating their research to the broader community, our SharedIt initiative provides you with a unique shareable link that will allow anyone (with or without a subscription) to read the published article. Recipients of the link with a subscription will also be able to download and print the PDF.

As soon as your article is published, you will receive an automated email with your shareable link.

You can now use a single sign-on for all your accounts, view the status of all your manuscript submissions and reviews, access usage statistics for your published articles and download a record of your refereeing activity for the Nature journals.

An online order form for reprints of your paper is available at https://www.nature.com/reprints/author-reprints.html. Please let your coauthors and your institutions' public affairs office know that they are also welcome to order reprints by this method.

If you have not already done so, we invite you to upload the step-by-step protocols used in this manuscript to the Protocols Exchange, part of our on-line web resource, natureprotocols.com. If you complete the upload by the time you receive your manuscript proofs, we can insert links in your article that lead directly to the protocol details. Your protocol will be made freely available upon publication of your paper. By participating in natureprotocols.com, you are enabling researchers to more readily reproduce or adapt the methodology you use. Natureprotocols.com is fully searchable, providing your protocols and paper with increased utility and visibility. Please submit your protocol to https://protocolexchange.researchsquare.com/. After entering your nature.com username and password you will need to enter your manuscript number (NG-A56560R1). Further information can be found at https://www.nature.com/nprot/.

Sincerely,
Kyle

Kyle Vogan, PhD
Senior Editor
Nature Genetics
https://orcid.org/0000-0001-9565-9665