nature portfolio

Peer Review File

Multi-ancestry, trans-generational GWAS meta-analysis of gestational diabetes and glycaemic traits during pregnancy reveals limited evidence of pregnancy-specific genetic effects

Corresponding Author: Ms Caroline Brito Nunes

This manuscript has been previously reviewed at another journal. This document only contains information relating to versions considered at Nature Communications.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
I'd like to thank the authors for their consideration of my comments. I think the manuscript has definitely improved but I have two residual concerns:

1. I still find it very strange that 1/3rd of the known loci (9/27) described in the introduction have disappeared into the ether with no mention besides an appearance at the bottom of a supplementary table. What confidence would this give a reader that 1/3rd of the current loci don't suffer the same fate in the next iteration of GWAS should there be one? Several loci seem to be lost because of various analytical choices, for example to only use MR-MEGA or to not include X-chromosome. I would suggest that ascertaining what loci are real should be the key aim here. The danger is that with GWAS meta-analyses, the largest study will be taken as the benchmark that all subsequent studies are based on. My concern is that real things are being missed here for avoidable reasons that can be addressed. On the flip side, if there is compelling evidence that some of the past loci are false-positives (comments about very poorly imputed loci sound concerning!) this should be clearly articulated. This issue needs to be better addressed and discussed in the text as the field will ultimately benefit from clarity around what is real or not.

2. The issue of whether these GDM loci are known T2D/glycaemic trait loci (albeit with effect heterogeneity) is still more ambiguous and confusing than it needs to be. In the rebuttal the authors state "all" GDM loci overlap with GWAS of T2DM and gylcaemic traits outside of pregnancy, yet the abstract and other parts of the text say "most". I also don't understand why a simple beta/p-value lookup of the relevant GWASs can't be added to ST4 rather than a more qualitative assessment which is currently included. I'd also challenge the authors to think about their final sentence conclusion in the abstract – does this current study really suggest that larger more ancestrally diverse GDM studies are needed, or that more T2D and glycaemic trait studies are needed?

Reviewer #3

(Remarks to the Author)
I have no further comments.

Version 1:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
I am satisified with the responses to my concerns.

Reviewer #2 (Remarks to the Author)

I'd like to thank the authors for their consideration of my comments. I think the manuscript has definitely improved but I have two residual concerns:

1. I still find it very strange that 1/3rd of the known loci (9/27) described in the introduction have disappeared into the ether with no mention besides an appearance at the bottom of a supplementary table. What confidence would this give a reader that 1/3rd of the current loci don't suffer the same fate in the next iteration of GWAS should there be one? Several loci seem to be lost because of various analytical choices, for example to only use MR-MEGA or to not include X-chromosome. I would suggest that ascertaining what loci are real should be the key aim here. The danger is that with GWAS meta-analyses, the largest study will be taken as the benchmark that all subsequent studies are based on. My concern is that real things are being missed here for avoidable reasons that can be addressed. On the flip side, if there is compelling evidence that some of the past loci are false-positives (comments about very poorly imputed loci sound concerning!) this should be clearly articulated. This issue needs to be better addressed and discussed in the text as the field will ultimately benefit from clarity around what is real or not.

We thank the reviewer for this comment. Our study is a multi-ancestry meta-analysis, whereas several previous reports (e.g., MONN and FinnGen) were conducted within single ancestries. Moreover, in the case of MONN, the summary statistics available to us at the time of analysis were from an earlier and smaller release (Zhen et al., 2024; PMID: 38372780) than the now-published dataset (Gu et al., 2025; PMID: 40325049). Differences in sample size, ancestry composition, allele frequencies, LD structure, imputation quality, and statistical models can materially affect power and genome-wide significance. As such, loci identified in smaller or single-ancestry studies may not reach genome-wide significance in a multi-ancestry meta-analysis, even if they represent true underlying signals.

We agree that it is important to distinguish between signals that are not detected due to analytical constraints and those for which there is limited supporting evidence. In our evaluation of previously reported loci, for example, several variants were absent from the MONN summary statistics we had access to (2 SNPs), had poor imputation quality (2 SNPs), were present in too few cohorts to meet MR-MEGA requirements (4 SNPs), or located on the X chromosome, for which most cohorts lacked data (1 SNP). Some SNPs met more than one of these exclusion criteria (e.g., poor imputation and insufficient cohort representation). For a subset of loci, we observed limited supporting evidence (e.g., lack of nominal association in multi-ancestry and ancestry-specific analyses), and 3 loci were inconsistently replicated even within their original reports in Gu et al. Our analytical choices, including MR-MEGA, exclusion of low-quality variants, and

omission of the X chromosome, prioritized cross-cohort consistency and data quality, increasing confidence that the loci retained represent robust and reproducible associations rather than signals driven by single cohorts or low-quality variants.

To address this concern directly, we have expanded this section, which was previously presented as Supplementary Table 4 and is now presented as a Supplementary Table 5, to clearly categorize previously reported loci according to the reason they were not observed at genome-wide significance in our study (e.g., unavailable variant, insufficient imputation quality, lack of replication in the original study by Gu et al., 2025). We have also added a sentence to the limitations section in the Discussion mentioning the lack of replication of some of the reported GDM loci. All changes made to the manuscript are highlighted in yellow for ease of reference. We hope these additions provide clearer context for readers regarding which signals are well supported across studies and which remain uncertain.

2. The issue of whether these GDM loci are known T2D/glycaemic trait loci (albeit with effect heterogeneity) is still more ambiguous and confusing than it needs to be. In the rebuttal the authors state "all" GDM loci overlap with GWAS of T2DM and gylcaemic traits outside of pregnancy, yet the abstract and other parts of the text say "most". I also don't understand why a simple beta/p-value lookup of the relevant GWASs can't be added to ST4 rather than a more qualitative assessment which is currently included. I'd also challenge the authors to think about their final sentence conclusion in the abstract – does this current study really suggest that larger more ancestrally diverse GDM studies are needed, or that more T2D and glycaemic trait studies are needed?

We have now carefully revised the manuscript to ensure consistent language throughout. Specifically, we clarify that all GDM loci showing stronger effects in GDM than T2DM are also associated with T2DM and/or glycaemic traits outside of pregnancy. The text in the abstract has been updated to reflect this and now reads: "We classified 12 GDM variants with stronger effects on GDM than T2DM into five biologically informed categories, revealing distinct patterns of pleiotropy, pregnancy-dependent effect modification, and diagnostic heterogeneity. *All* these loci overlapped with T2DM and/or non-pregnant glycaemic traits."

In addition, as suggested, we have now included a lookup for all lead GDM SNPs in both Europeans and East Asians, in our GWAS of glycemic traits during pregnancy (FG, 1hG, 2hG and HbA1c), glycemic traits outside of pregnancy (FG, 2hG and HbA1c; PMID: 34059833) as well as T2DM (PMID: 38374256) and have added it to Supplementary Table 4. If variant was available in the GWAS used for the lookup, we report the beta, standard errors and p-values.

Our intention in the final sentence of the abstract was to emphasise the need for larger and more ancestrally diverse GDM and pregnancy-specific glycaemic trait studies, rather than additional T2DM or general glycaemic trait GWAS, which are already substantially larger and better

powered. We recognise that the previous wording was potentially ambiguous and may have implied a broader need across all glycaemic traits. To clarify this, we have revised the final sentence of the abstract to explicitly refer to larger, multi-ancestry genetic studies of GDM and glycaemic traits during pregnancy (changes are highlighted in yellow). This revision aligns with the limitation described in the Discussion regarding sample size imbalance and limited ancestry-specific power in current GDM datasets.