nature portfolio

Peer Review File

Potentially causal associations between placental DNA methylation and schizophrenia and other neuropsychiatric disorders

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

In this work, Cilleros-Portet and colleagues aim at exploring how the placenta methylome can contribute to mediate the effect of genomic risk for neuropsychiatric disorders. They first characterize the landscape of placental mQTL, defining nominal, permuted, and conditional cis-mQTL, using placental genomic data from the INMA project. They define a catalog of imQTL, i.e., mQTL associated with the interaction between genetic variation and estimates of cell composition and gestational age. They then leverage SMR and colocalization to identify DNA methylation (DNAm) sites whose predicted level of methylation is associated with neuropsychiatric disorders. In this way, they prioritize DNAm sites associated with depression, bipolar disorders, and schizophrenia. Using the placenta genomic data form the RICHS cohort, Cilleros-Portet and colleagues then analyze whether methylation of these CpGs predict expression of the nearby genes. The analytical pipeline is finally enriched by a conditional SMR analysis, a SMR of imQTL, and by a comparison with SMR results for brain mQTL. With this work the authors identify placenta DNAm sites associated with risk for depression, bipolar disorder, and schizophrenia, some of which are associated with the expression of the nearby genes in placenta. They also provide a useful resource to the scientific community, in terms of datasets of mQTL and imQTL. Their main conclusion is that genomic risk of several neuropsychiatric disorders could also operate through DNAm and associated gene expression in placenta.

The study is interesting and well-written, and it provides novel and important knowledge about the possible role of DNAm in mediating the effects of genomic risk for neuropsychiatric disorders. The authors employ an appropriate methodology, and the availability of codes and data will ensure reproducibility of their results.

I think there are just few points that need to be clarified, and some analyses to be performed, to further improve the quality of this work.

1. The authors highlight the finding of a particular relevance of the placental methylome for bipolar disorders, depression, and schizophrenia. However, these are also the disorders with largest sample size in GWAS studies, as shown in Table 2, and the highest number of GWAS hits, so that it is not clear whether such relevance is driven by biology or by power issues. I suggest to add some other disorders to the analyses to help to answer this question (for example, height, BMI, IQ, diabetes, etc). As alternative, the authors could compare the number of DNAm sites across disorders, after adjusting for GWAS sample size, and GWAS hits.

2. In the reporting summary, the authors write that analyses stratified by sex were not performed for power issue. However, I think that it could be interesting to add those analyses, in light of the fact that the placenta epigenetic landscape could be very different in males and females. If they prefer, they could also search for sex-interacting mQTL.

3. In looking at Figure 3, I'm a bit surprised to see colocalization signals (blue dots) for SNPs that are much below the GWAS p-value threshold of significance (5e-08). Did the authors perform the colocalization analysis to identify the overlapping of causal meQTL and GWAS-significant SNPs, or to identify the overlapping of causal meQTL and SNPs in GWAS-significant loci? I think a GWAS significant threshold ($p<5e-08$, or $p<1e-06$) should be considered to include the SNPs in the colocalization analysis, and it's not an issue if the colocalization signals will be less. If the authors want to perform a locus-level colocalization analysis, they could also consider some recent work on this topic (PMID: 35523146).

4. Some more discussion about biology would benefit the manuscript. Do these results suggest any possible epigenetic mechanism through which genomic risk for neuropsychiatric disorder act in placenta? For example, any possible effect on placental imprinting? Any convergence between these methylation changes and other methylation changes associated with early life environmental exposures? Is there any convergence between the DNAm changes associated with depression, bipolar disorder, and schizophrenia?

Minor points:

5. #103,104: I suggest to add some more references supporting the link between prenatal insults and schizophrenia.

6. #152-154: a definition of cell type interacting eQTL and mQTL may be useful for the broad readership of this journal, here or also in the results section (#229) where I suggest to include also a definition of positive and negative imQTL.

7. #206: please provide a definition of "permuted and conditional cis-mQTLs". I know it's in the methods section, but here it would help too.

8. #223: please specify that STB are multinucleated

9. #Table 2: I suggest to specify N of cases and controls, number of GWAS-significant loci detected in each study, and also N of DNAm sites associated in this study.

10. #397-403: were, the overlapping placenta and brain mQTL, associated with schizophrenia with the same directionality in SMR? Also, the authors say “we cannot rule out that mostly in the case of these non-specific shared DNAm sites, the placenta could be mirroring what occurs in the brain without necessarily being the main effector tissue”. However, genetic and environmental variation could also contribute to risk for schizophrenia through pleiotropic epigenetic effects in placenta and brain.

11. #477-481: “Particularly in the latter work, they also studied the existence of placenta-specific PRS in other neuropsychiatric disorders, and if present, whether they interacted with early neurodevelopmental outcomes. BIP presented a PRS with a very high enrichment in genes that are expressed in placenta, although this placenta-specific PRS did not interact with neither brain volume nor cognition.” I suggest to edit these sentences to better specify that the authors studied the enrichment in placenta of risk genes for other neuropsychiatric disorders and the relationship between neurodevelopmental outcomes and placenta-specific PRS for the same neuropsychiatric disorders.

12. Why imQTL were calculated only for STB and TB? Were other cell proportion estimated, e.g. Hofbauer cells, endothelial cells, stromal cells?

13. How was the nominal p-value cutoff of 5x10-8 established?

14. A gene set enrichment analysis with the genes with placental expression associated with methylation of the CpG associated with schizophrenia, bipolar disorder, and depression, would be interesting.

15. Perhaps in reporting the results the authors should mention the number of loci / genes with DNAm sites associated to the disorders, in addition to the numbers of DNAm sites. Similarly, in Fig.7, the overlap could be also represented in terms of genes/loci (perhaps within parentheses).

16. I suggest to specify in the main text the number of associations detected in the MHC locus, since the high linkage disequilibrium in this region may enhance false discovery. As alternative, the authors could report separately the MHC results.

Reviewer #2 (Remarks to the Author):

I have shared this reviewing responsibility with one well-qualified colleague, with permission from the Nature Comm, so I use the pronoun “we” below.

The authors have studied methylation quantitative trait loci (mQTLs) in human term placentas and created a useful comprehensive placental mQTL database. They have also usefully linked mQTLs to cis-regulation of placental gene expression, based on data from the Rhode Island Child Health (RICHS) study, from which they derived lists of expression quantitative trait methylation sites (eQTMs). They then applied their findings as a “post-GWAS” approach for narrowing in on potentially functional SNPs located under broad GWAS peaks for major neuropsychiatric conditions, including schizophrenia (SCZ), major depression (MDD), and bipolar disorder (BIP).

GENERAL COMMENTS:

1. The main strengths of the study are in the rigorous and diverse statistical approaches employed, and the valuable mQTL database generated.

2. The usefulness of these data for homing in on strong candidate functional SNPs under GWAS peaks, for neuropsychiatric disorders and, for that matter, any common genetically complex human phenotype, is straightforward and builds significantly on previous work in this area. This aspect of the paper would benefit from adding one or two locus-specific CRISP-Cas9 mutagenesis/deletion experiments as definitive functional validations, as detailed below.

3. The authors’ conclusions regarding specific biological connections between placental mQTLs (and eQTMs) and neuropsychiatric disorders are intriguing, but somewhat indirect. We also make some suggestions below for how they might bolster this aspect.

SUGGESTIONS/REQUESTS:

1. Most of the exposures in pregnancy that are thought to increase the risk of psychopathology in the offspring, need to occur in early pregnancy (eg. Alan Brown's work with SCZ). However, the placentas in the current study were collected at term. Could the authors obtain pre-term placentas (obviously it would be a smaller series) and determine if their high-priority disease-linked mQTLs

and eQTMs are stronger at the earlier stage? This type of information, if positive, could bolster their argument about a connection to neuropsychiatric disorders, as per the DOHaD hypothesis.

2. The authors hypothesize that the putative connection to neuropsychiatric disorders might occur via the effects of mQTLs on the expression of genes in immune/inflammatory pathways. However, their data analysis considered proportions of cyto- and syncitio-trophoblast but not the (relatively abundant) populations of myeloid and lymphoid cell types in term placentas. Information from a similar cell type deconvolution approach evaluating the relative percentage of such cells (placental macrophages, for example) and the influence of that percentage on the observed mQTLs, particularly the disease-associated ones, could also bolster their argument.

3. Regarding the post-GWAS application of their data for pinpointing functional regulatory SNPs ("rSNPs"), the paper would be dramatically strengthened if they could show, using CRISPR-Cas9 in convenient human cell lines (trophoblast or other; cancer lines would be adequate for this aspect) that their highest priority candidate rSNPs in fact control the expression and methylation of their highest priority candidate disease-linked genes – for example, LRFN5 (linked to MDD) and ZSCAN23 (linked to SCZ). Such experiments are straightforward to do, with precise site-directed mutagenesis being the most elegant approach, but with creation of small deletions spanning the relevant SNPs, followed by gene expression analysis (RNA-seq and/or RT-Q-PCR) in deletion-bearing vs wild-type cells being adequate.

Reviewer #3 (Remarks to the Author):

Cilleros-Portet et al. have performed cis-mQTL and interaction cis-mQTL mapping in 368 fetal placenta samples. The (i)mQTL results were then integrated with the largest GWASes on 10 neuropsychiatric disorders using SMR and colocalization analyses. The manuscript highlights interesting data. However, the selected methods do not follow the best practices in the field. There are also several other concerns that should be addressed.

Major points:

1. Discovery of cis-mQTLs and interaction cis-mQTLs. The authors account for fetal sex, 5 genotype PCs and Planet-estimated cell types in their cis-mQTL analysis. As there are several technical and biological factors that could affect the methylation or gene expression levels, it is common in molQTL analysis to include latent factors (such as PEER factors or PCs) to account for unwanted

variation in the data. Also, to avoid outlier effects inverse normal transformation is applied on the molecular data. These steps are followed by the GTEx Consortium (that the authors refer to) and recommended by the Aguet et al 2023 Nature Reviews Methods Primer paper.

2. Multiple testing. The authors thoroughly characterize mQTLs with P_nominal < 5e-8. A proper multiple testing procedure should be incorporated to cis-mQTL and interaction cis-mQTL discovery. Only mQTLs and imQTLs with FDR < 0.05 should be characterized and included to downstream analysis.

3. Sharing between iQTLs. How were common STB-imQTLs and TB-imQTLs found? Are iVariants in high LD affecting the methylation levels of the same CpG site considered to be common to both cell types? Storey's pi1 estimate would be a better estimate of sharing between STB-imQTLs and TB-imQTLs (it is not influenced by p-value cut-offs to call an imQTL a significant imQTL. There might be TB-imQTLs that are shared with STB-imQTLs just below the significance level due to statistical power). Same comment for estimating the overlap between STB/TB-imQTLs and GA-imQTLs.

4. Colocalization analysis: there seems to be a mistake in counting the region-CpG pairs with support of colocalization. For example, Supplementary Data 6 (BIP tab) indicates that there is strong support for H3 in coloc analysis between region 43772251 - 43592251 and CpG cg24283270 (PP3 = 0.999938693 and PP4 = 4.78e-05). Please correct this mistake. Also, it would be informative if lead variants in imQTL and GWAS analysis are highlighted in the table together with the p-values from the respective analyses. Maybe also genomic position of the CpG site could be included.

5. Processing of DNA methylation data. Is applying both noob background correction with dye-bias correction followed by functional normalization and beta-mixture quantile normalization necessary? May it result in over-correction of the data? Do I understand correctly that probes with SNPs that have MAF < 5% in European samples (line 640) are excluded? One should want to remove probes with common SNPs in the probe sequence to avoid false positive mQTL hits. How are cross-hybridizing probes defined?

Minor points:

1. Line 67: what are those eight million mQTLs? Are they significant mQTLs at some significance threshold?

2. What are nominal mQTLs and nominal mQTL database? Please be more clear.

3. Figure 1a: how is the median distance calculated? I assume it is the median value of the absolute distance between the CpG site and the SNP. If yes, then showing median distance as a red line on Fig. 1b is misleading, because the plot summarizes both the cases where the SNP is located before or after the CpG site.

4. Figure 1b: I disagree with the conclusion that “the distribution of mQTLs across chromosomes was in line with the chromosome length, except for chr 6”. For example, there are considerably fewer mQTLs on chr 9 as compared to chr 10-12, which are smaller in size.

5. Figure 1c: based on the volcano plot is hard to evaluate whether the effect sizes of the mQTLs follow uniform distribution or not. Bar plot would be better for that purpose.

6. Figure 1d and 1e: I found these pie charts to be hard to read. Would it be easier for the reader if the data is replotted using bar plots, for example? Also, then one could highlight significance differences in distributions more easily.

7. Line 198: as there are enrichments in many other DNase I hotspots too, I would stress that the strongest signal is observed in fetal placenta. Same comment for Suppl. Fig. 1A.

8. Supplementary Figure 2c: is the adjusted p-value the same for all the tested diseases? It is an unexpected result.

9. Line 206: what are permuted and conditional cis-mQTLs?

10. Figure 2b and Figure 2c: labels are mixed up in the text (line 227 and line 233).

11. As there is such a strong negative correlation between STB and TB, it is expected that there is negative correlation between the effect sizes of common STB- and TB-imQTLs (and in general a high sharing between STB-imQTLs and TB-imQTLs). It seems reasonable to focus only on the cell type with the highest statistical power.

12. What data is used to calculate LD between SNPs on the locusZoom plots?

13. Figure 3 is not mentioned in the main text.

14. Figure 4 and lines 305-314: there seems to be very strong LD between the mQTL-SNPs rs1111179 and rs61990289 based on the panel a. Thus, I would suspect that the methylation levels of the two CpGs sites are also strongly correlated. Thus, they represent the same signal rather than two independent, pleiotropically associated CpG sites. What is the correlation between the two CpG sites?

15. What does it mean that the regression coefficients, their variances and SNP minor allele frequencies were imputed for the coloc analysis?

First of all, we would like to thank the reviewers since we feel that they have carried out a rigorous and constructive revision of our work. We believe to have improved the manuscript considerably following their suggestions.

In this context, apart from the thorough revision of the main text of the article, we have redone main Figures 1, 2, 3 and 6, as well as Table 2. Regarding Supplementary Figures, we have modified 3 and 4, and added new Supplementary Figure 6, as well as 11, 12 and 13. Additionally, most of Supplementary Data had minor modifications (as new highlights and sheets).

Among the improvements and novelties we would like to highlight the creation of two novel mQTL datasets (one corrected by DNAm PCs and another one showing sex-interaction) that will be made publicly available. Additionally, we have corrected the filtering of the colocalization approach and discovered new pathways related to placental DNAm, that will help deepen in the biological interpretation of the data. We have also tried our best clarifying the methodological aspects of our work. We hope that the reviewers will be satisfied with these changes and are open for discussion.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

In this work, Cilleros-Portet and colleagues aim at exploring how the placenta methylome can contribute to mediate the effect of genomic risk for neuropsychiatric disorders. They first characterize the landscape of placental mQTL, defining nominal, permuted, and conditional cis-mQTL, using placental genomic data from the INMA project. They define a catalog of imQTL, i.e., mQTL associated with the interaction between genetic variation and estimates of cell composition and gestational age. They then leverage SMR and colocalization to identify DNA methylation (DNAm) sites whose predicted level of methylation is associated with neuropsychiatric disorders. In this way, they prioritize DNAm sites associated with depression, bipolar disorders, and schizophrenia. Using the placenta genomic data form the RICHS cohort, Cilleros-Portet and colleagues then analyze whether methylation of these CpGs predict expression of the nearby genes. The analytical pipeline is finally enriched by a conditional SMR analysis, a SMR of imQTL, and by a comparison with SMR results for brain mQTL. With this work the authors identify placenta DNAm sites associated with risk for depression, bipolar disorder, and schizophrenia, some of which are associated with the expression of the nearby genes in placenta. They also provide a useful resource to the scientific community, in terms of datasets of mQTL and imQTL. Their main conclusion is that genomic risk of several neuropsychiatric disorders could also operate through DNAm and associated gene expression in placenta.

The study is interesting and well-written, and it provides novel and important knowledge about the possible role of DNAm in mediating the effects of genomic risk for neuropsychiatric disorders. The authors employ an appropriate methodology, and the availability of codes and data will ensure reproducibility of their results.

Response: We thank the reviewer for his/her positive feedback.

I think there are just few points that need to be clarified, and some analyses to be performed, to further improve the quality of this work.

1. The authors highlight the finding of a particular relevance of the placental methylome for bipolar disorders, depression, and schizophrenia. However, these are also the disorders with largest sample size in GWAS studies, as shown in Table 2, and the highest number of GWAS hits, so that it is not clear whether such relevance is driven by biology or by power issues. I suggest to add some other disorders to the analyses to help to answer this question (for example, height, BMI, IQ, diabetes, etc). As alternative, the authors could compare the number of DNAm sites across disorders, after adjusting for GWAS sample size, and GWAS hits.

Response: We agree with the reviewer that in fact, there seems to be a correlation between the statistical power of the GWAS studies and the number of significant hits obtained in SMR. We had discussed this point in the original manuscript as follows:

"*These results could arise from the sample sizes and power constraints of the available GWAS, as well as, from the higher heritability of certain disorders, including BIP and SCZ. However, very importantly, a recent article found out that MDD shows a very high polygenicity compared to other psychiatric disorders; that is, more genetic variants with weaker effects contribute to the overall genetic signal in MDD, and make the trait less annotable[75]. In contrast, ADHD, BIP and SCZ showed the highest discoverability and hence a more annotable genetic signal. Subsequently, the estimated sample size required to reach 90% SNP heritability was more than eight times larger for MDD than for ADHD, BIP and SCZ. Therefore, the lack of signal in ADHD and even more remarkably, the presence of a considerable pleiotropy between placental DNAm and MDD, suggest that our findings are guided not only by the strength and annotability of the genetic basis of the diseases studied, but rather by a genuine association with placental DNAm*".

However, we thought that the suggestions by the reviewer could help reinforce the idea of independence between the statistical power and the results obtained, and thus, performed several additional experiments and added Supplementary Figure 13. In this figure we show the correlation between SMR hits and i) the GWAS sample size, ii) the GWAS significant hits, and iii) the total number of SNPs. We observe that whereas there seems to be some correlation between the number of the significant results in SMR and the GWAS sample size, there is no correlation with the other two variables studied. Moreover, when Pearson correlation between the GWAS sample size and the number of SMR hits is computed we get a r=0.52, with a p-value of 0.13. Additionally, we have run the same SMR approach with two previous Schizophrenia GWAS that, although much smaller than the one employed initially, still provide significant SMR results, with a remarkable overlap among the different GWAS studies in the resulting SMR hits. The exact figures of these studies as well as their characteristics are summarized in the abovementioned supplementary figure, followed by the Venn Diagrams showing the overlap.

We have added these considerations in the Discussion of our manuscript as follows:
"*However, we wanted to make sure that our SMR results were not merely led by statistical power and very especially, by GWAS sample size. Therefore, we calculated the correlation between SMR hits and i) the GWAS sample size, ii) the GWAS significant hits, and iii) the total number of SNPs per GWAS, in all the neuropsychiatric disorders studied. We observed no correlation between the number of significant SMR results and the abovementioned variables (Supplementary Figure 13A, B and C). Additionally, in the specific case of SCZ, we ran the same SMR approach with two previous, smaller GWAS that still provided significant SMR hits, with a remarkable overlap among the different studies. The exact characteristics of these studies are summarized in Supplementary Figure 13D, followed by the Venn Diagrams showing the overlapping results (Supplementary Figure 13E and F)*".

2. In the reporting summary, the authors write that analyses stratified by sex were not performed for power issue. However, I think that it could be interesting to add those analyses, in light of the fact that the placenta epigenetic landscape could be very different in males and females. If they prefer, they could also search for sex-interacting mQTL.

Response: We thank the reviewer for his/her suggestion and in fact, we think that the idea of mapping sex-interacting mQTLs is a good approach to bypass the power issues that could arise from the stratification of our data. In this sense, we have discovered 521 sex-imQTLs and run SMR with these. Very interestingly, we have obtained a common hit for both BIP and SCZ (cg09779537-rs2398668), and an additional hit for SCZ. SNP rs2398668 is located on chromosome 7, in a region showing high gene density and more specifically, in intron 7 of *SNX8*. This gene has been reported to act as a modulator of innate immunity in response to infections by DNA virus. Its disruption has also been related to mild intellectual disability. We have decided to add these findings at different points of the manuscript as follows.

In Results, new section "Placental cell type-, GA- and sex- imQTLs": "*Finally, as placental methylation could differ between sexes, we also mapped sex-imQTLs and discovered 521 associations at a genome-wide suggestive significance level ($P < 5x10^{-8}$) (Supplementary Data 4).*".

In Results, new section "*SMR analysis of placental cell type-, GA- and sex- imQTLs in BIP, MDD and SCZ*": "*Finally, we employed SMR to combine sex-imQTLs with the BIP, MDD and SCZ GWAS, obtaining a common hit for BIP and SCZ, and an additional hit for SCZ (Supplementary Data 10). The common sex-imQTL involved cg09779537 and rs2398668 (Supplementary Figure 11). Interestingly, rs2398668 is borderline significant in the SCZ GWAS, and non-significant in the BIP GWAS. The presence of the minor allele T is associated with reduced DNAm of cg09779537, specifically in males. The imQTL-SNP is in a region with high gene density, in the gene body of SNX8. This gene has been reported to act as a modulator of innate immunity in response to viral infections[63] and its disruption has been related with mild intellectual disability[64]. Overall, the data suggest a potential causal relationship between the associated CpG site, and BIP and SCZ*".

3. In looking at Figure 3, I'm a bit surprised to see colocalization signals (blue dots) for SNPs that are much below the GWAS p-value threshold of significance (5e-08). Did the authors perform the colocalization analysis to identify the overlapping of causal meQTL and GWAS-significant SNPs, or to identify the overlapping of causal meQTL and SNPs in GWAS-significant loci? I think a GWAS significant threshold (p<5e-08, or p<1e-06) should be considered to include the SNPs in the colocalization analysis, and it's not an issue if the colocalization signals will be less. If the authors want to perform a locus-level colocalization analysis, they could also consider some recent work on this topic (PMID: 35523146).

Response: On the one hand, regarding the colocalization method used in this manuscript, we would like to confirm that we applied this analysis to identify the overlap between causal mQTLs and GWAS-significant loci. The reason why we did not apply a GWAS significance threshold is that the developers of *coloc* specifically request not to do so, and therefore, all SNPs available have to be included. This is because the software enumerates all possible configurations of shared and non-shared causal variants an evaluates their relative support. To clarify this, we have added a sentence in Methods as follows: "*Colocalization analysis was performed to identify the overlap between causal mQTLs and GWAS-significant loci, as described in Giambartolomei C. et al.2014, with the R coloc package*[49]".

On the other hand, we thank the reviewer for bringing to our attention the methodology from Hukku et al.2022. To be honest, we were not aware of this publication and thus, did not consider it in our analysis. We have carefully read it, as well as other works employing their approach, and consider it a good reference to take into consideration in future projects. We understand that the locus-level colocalization probability (LCP) is more powerful than the regional colocalization probability (RCP), since RCP does not account for possible events other than distinct causal variants for different traits coexisting in the same locus, but we believe that the conditional SMR analysis is somehow dealing with this situation, as stated by Zhu et al.2016. Nevertheless, we would like to highlight that we considered running colocalization as a complementary analysis after reading the 2015 Nature Neuroscience paper by Hannon et al., where the authors performed different analyses using fetal brain mQTLs. In fact, these are the mQTLs with which we ran SMR to compare the results to our own placental SMR hits. We felt that it was coherent to follow a similar pipeline also for colocalization.

4. Some more discussion about biology would benefit the manuscript. Do these results suggest any possible epigenetic mechanism through which genomic risk for neuropsychiatric disorder act in placenta? For example, any possible effect on placental imprinting? Any convergence between these methylation changes and other methylation changes associated with early life environmental exposures? Is there any convergence between the DNAm changes associated with depression, bipolar disorder, and schizophrenia?

Response: The point raised by the reviewer is central and warrants future work in the field. Following his/her suggestion, on the one hand, we have checked the overlap between our significant SMR hits for BIP, MDD and SCZ, and those CpGs that are significantly associated in the two largest placental EWAS performed to date in the PACE consortium, i.e. PMID:34429407 on maternal smoking and PMID:36446949 on pre-pregnancy maternal BMI. We have found no overlap and therefore, it seems that those exposures are not affecting the loci associated to any of the three disorders in our study. On the other hand, we have checked the overlap between the SMR results of the three main disorders studied and have found an interesting convergence of 5 hits in BIP and SCZ, that moreover show a common direction of effect in both phenotypes. This finding could be in line with the similar genetic architecture described for the two disorders and has been added in the manuscript as follows.

In Results section "*Multi-omics approaches to unravel the placental origin of neuropsychiatric traits and disorders*": "*Six hits were common to BIP and SCZ, with the same direction of effect, as highlighted in the abovementioned Supplementary Data file*".

In Discussion: "*...SCZ and BIP present the highest genetic correlation compared to any other pair of psychiatric traits[79]. This is in line with the convergence found in this study between the CpGs associated with both BIP and SCZ. Nevertheless, there are remarkable differences in outcome...*"

Finally, thanks to the minor point 14 raised by this reviewer, in which it was advised to deepen into the pathway analysis of all the SMR-eQTM genes identified, we have found many epigenetic regulatory pathways involved, including different histone modifications, DNA methylation and chromatin condensation and organization, suggesting that SCZ risk alleles could cause deep epigenetic changes in placenta as a result of DNAm modifications located in close proximity to different histone-coding genes. Additionally, we have observed even more immune pathways associated to SCZ including, among the top enriched biological routes, "Human cytomegalovirus (HCMV) Infection" and "Viral Infection", given the presence of different histones and the

VPS37B **gene, that encodes a component of the endosomal sorting complex required for viral budding. In this sense, it has been reported that the DNAm states of the hosts can make cells more or less prone to HCMV infections, and that histones participate in this process because HCMV is devoid of these proteins but still captures them from the host for the chromatinization of its own genome (Esteki-Zadeh A et al.2012). Regarding SCZ, it is known that congenital HCMV infection can cause clinical long-term manifestations such as hearing loss, neurodevelopmental disorders, ophthalmic complications and ASD, among others (Elgueta D et al.2022). The virus moves from mother to fetus by infecting TB in the placenta, and the infection alters the development and integrity of the organ, as well as its gene expression patterns, to eventually inhibit placental cell differentiation, self-renewal and migration. We think that these new findings are relevant and have included them in the manuscript. Again, we sincerely thank the reviewer for his/her advice, that from our point of view, has helped improve our work considerably.**

In Results section "*Multi-omics approaches to unravel the placental origin of neuropsychiatric traits and disorders": "Finally, we performed a Reactome gene-set analysis[58] of the 26 genes at the intersection between the SMR and the eQTM approaches in SCZ. We observed enrichment for 115 gene-sets (Supplementary Data 8), including many epigenetic regulatory pathways, such as different histone modifications, DNA methylation and chromatin condensation and organization, suggesting that SCZ risk alleles could change the epigenetic landscape of placenta through the regulation of different histone-coding genes. Additionally, we observed a remarkable enrichment in immune-related pathways such as interferon alpha/beta signaling, interferon gamma signaling, and interferon signaling, as well as human cytomegalovirus (HCMV) infection and overall viral infection, reinforcing the idea of MIA as a link between placenta and SCZ risk***".**

In Discussion: "*Regarding the genes affected by placental DNAm pleiotropically associated to SCZ, we want to highlight that they were enriched in epigenetic regulation pathways and immune-related terms, including interferon signaling and viral infection. On the one hand, the multiple epigenetic pathways involved suggest that SCZ risk alleles could cause deep epigenetic changes in placenta as a result of DNAm modifications in close proximity to different histone-coding genes. On the other hand, the previously mentioned histone-coding genes, together with others such as HLA-H, IRF3 and VPS37B, were some of the most relevant genes contributing to the enrichment in immune-related biological routes. This reinforces the idea of MIA being implicated in the neurodevelopmental origins of SCZ. Particularly for HCMV, it has been reported that the DNAm state of the host genome can make cells more or less prone to its infection, and that histones participate in this process because, although HCMV is devoid of these proteins, it captures them from the host for the chromatinization of its own genome[81]. Moreover, it is known that congenital HCMV infection can cause long-term clinical manifestations such as hearing loss, neurodevelopmental disorders, ophthalmic complications, and ASD, among others[82]. The virus moves from mother to fetus by infecting TB in the placenta, altering the development and integrity of the organ, and eventually, inhibiting placental cell differentiation, self-renewal, and migration***".**

In Bibliography:

"*Esteki-Zadeh., A, et al. Human cytomegalovirus infection is sensitive to the host cell DNA methylation state and alters global DNA methylation capacity. Epigenetics. 7, 585-593 (2012)***"**

"*Elgueta, D., Murgas, P., Riquelme, E., Yang, G., Cancino, G.I. Consequences of Viral Infection and Cytokine Production During Pregnancy on Brain Development in Offspring. Front. Immunol. 13, 816619 (2022)***"**

Minor points:

5. #103,104: I suggest to add some more references supporting the link between prenatal insults and schizophrenia.

Response: We have added the following line and reference: "*In fact, different pieces of evidence state that prenatal insults such as maternal immune activation (MIA) and other pregnancy and labor-related triggers such as hemorrhage, preeclampsia, birth asphyxia, and uterine rupture are associated with SCZ in offspring*[11,12]."

"Cannon, M., Jones, P. B. & Murray, R. M. Obstetric complications and schizophrenia: historical and meta-analytic review. *Am. J. Psychiatry* 159, 1080–1092 (2002)"

6. #152-154: a definition of cell type interacting eQTL and mQTL may be useful for the broad readership of this journal, here or also in the results section (#229) where I suggest to include also a definition of positive and negative imQTL.

Response: We have defined ieQTLs/imQTLs as proxies of cell-type specific eQTLs/mQTLs as we consider this definition the most intuitive. However, we agree that positive vs negative classification deserves an additional explanation in the Results section. Therefore, we have added the following sentence: "*In the case of cell type-imQTLs, positive imQTLs showed increased DNAm related to the main genotype as a function of cell proportion, while negative showed a decrease*".

7. #206: please provide a definition of "permuted and conditional cis-mQTLs". I know it's in the methods section, but here it would help too.

Response: We have added a brief explanation as requested: "*Apart from the nominal cis-mQTLs, in which all significant SNP-CpG combinations were reported, permuted and conditional cis-mQTLs were also calculated in our 368 placenta DNA samples. Briefly, the permuted cis-mQTL approach corrects for multiple correlated variants tested via a beta approximation, whereas the conditional database uses the permuted QTLs to perform a stepwise regression procedure and map conditionally independent cis-QTLs (see Methods for more details)*".

8. #223: please specify that STB are multinucleated

Response: Done.

9. #Table 2: I suggest to specify N of cases and controls, number of GWAS-significant loci detected in each study, and also N of DNAm sites associated in this study.

Response: We have updated the Table 2 with the information requested.

10. #397-403: were, the overlapping placenta and brain mQTL, associated with schizophrenia with the same directionality in SMR? Also, the authors say "we cannot rule out that mostly in the case of these non-specific shared DNAm sites, the placenta could be mirroring what occurs in the brain without necessarily being the main effector tissue". However, genetic and environmental variation could also contribute to risk for schizophrenia through pleiotropic epigenetic effects in placenta and brain.

Response: The overlapping placental and brain mQTLs associated with SCZ have the same SMR direction. Nevertheless, we agree with the reviewer that genetic and environmental variation could also contribute to the risk of SCZ through pleiotropic epigenetic effects in placenta and brain. We have added this consideration in the manuscript as follows: "*...mostly in the case of those non-specific, shared DNAm sites, we cannot rule out neither that the placenta could be mirroring what occurs in the brain without necessarily being the main effector tissue, nor that both placenta and brain could contribute to the risk for SCZ through real, pleiotropic epigenetic effects*".

11. #477-481: "Particularly in the latter work, they also studied the existence of placenta-specific PRS in other neuropsychiatric disorders, and if present, whether they interacted with early neurodevelopmental outcomes. BIP presented a PRS with a very high enrichment in genes that are expressed in placenta, although this placenta-specific PRS did not interact with neither brain volume nor cognition." I suggest to edit these sentences to better specify that the authors studied the enrichment in placenta of risk genes for other neuropsychiatric disorders and the relationship between neurodevelopmental outcomes and placenta-specific PRS for the same neuropsychiatric disorders.

Response: We have tried to clarify this point following the reviewer's advice. We hope it is clearer now: "*Particularly in the latter work, they also studied the enrichment in placenta of risk genes for other neuropsychiatric disorders, and the relationship between neurodevelopmental outcomes and placenta-specific PRS for the same disorders. BIP presented a very high enrichment in genes that are expressed in placenta, although its placenta-specific PRS did not interact with neither brain volume nor cognition*".

12. Why imQTL were calculated only for STB and TB? Were other cell proportion estimated, e.g. Hofbauer cells, endothelial cells, stromal cells?

Response: We understand the question of the reviewer but the point here is that iQTLs, as defined by GTEx, are only computed in the most abundant cell type per tissue, due to power constraints. In fact, while two of the three reviewers have suggested to extend the imQTL mapping to other cell types, one has suggested to report only those calculated for STB, as it is the most abundant cell type. As a consequence, we have decided not to report imQTLs for more cell types, other than STB and TB.

13. How was the nominal p-value cutoff of 5x10-8 established?

Response: As the default cutoff for QTLs in SMR is $5x10^{-8}$, we decided to keep this also as a suggestive genome-wide threshold of significance in the mQTL databases.

14. A gene set enrichment analysis with the genes with placental expression associated with methylation of the CpG associated with schizophrenia, bipolar disorder, and depression, would be interesting.

Response: Following the reviewer's advice, we have performed a gene set enrichment analysis with the eQTMs from BIP, MDD and SCZ. The point is that BIP and MDD are not contributing to the final enrichment, and therefore we have decided to extend the pathway analysis for SCZ of the previous version to all the SMR-eQTM genes, regardless of the overlap with *coloc*. This has improved the power of the approach, with a substantial enrichment of the results that moreover, allows us to make a wider biological interpretation. We thank the reviewer for his/her suggestion

and we feel this has helped improve considerably the manuscript (please, see major point 4 for more details).

15. Perhaps in reporting the results the authors should mention the number of loci / genes with DNAm sites associated to the disorders, in addition to the numbers of DNAm sites. Similarly, in Fig.7, the overlap could be also represented in terms of genes/loci (perhaps within parentheses).

Response: We have added Supplementary Figure 12, showing the overlap among the mQTL-SNPs of the three tissues (analogue to Figure 7) and cited it in the main text as follows: "*As shown in Figure 7, the trait with the largest overlap of pleiotropically associated DNAm sites among tissues is SCZ, although the overlap is very limited in all the traits. The overlap observed for mQTL-SNPs is even more limited (Supplementary Figure 12)*".

16. I suggest to specify in the main text the number of associations detected in the MHC locus, since the high linkage disequilibrium in this region may enhance false discovery. As alternative, the authors could report separately the MHC results.

Response: Within the MHC locus we detect 1,754, 1,165 and 1,152 unique mQTL-participaiting CpGs in the nominal, permuted and conditional mQTL databases, respectively. This observation is translated into a significant enrichment of the MHC locus within our databases. The contingency tables for this enrichment can be found in the Supplementary Data 1, together with the tables for the Relation To Island and UCSC RefGene annotations. These results have been added in the manuscript as follows:

In Results, section "*Placental cis-mQTL characterization*": "*They were also enriched in the HLA region with 1,754 CpGs out of the total 110,721 mQTL-CpGs of the nominal database ($P < 2.2x10^{-16}$)*" .

Reviewer #2 (Remarks to the Author):

I have shared this reviewing responsibility with one well-qualified colleague, with permission from the Nature Comm, so I use the pronoun "we" below.

The authors have studied methylation quantitative trait loci (mQTLs) in human term placentas and created a useful comprehensive placental mQTL database. They have also usefully linked mQTLs to cis-regulation of placental gene expression, based on data from the Rhode Island Child Health (RICHS) study, from which they derived lists of expression quantitative trait methylation sites (eQTMs). They then applied their findings as a "post-GWAS" approach for narrowing in on potentially functional SNPs located under broad GWAS peaks for major neuropsychiatric conditions, including schizophrenia (SCZ), major depression (MDD), and bipolar disorder (BIP).

GENERAL COMMENTS:
1. The main strengths of the study are in the rigorous and diverse statistical approaches employed, and the valuable mQTL database generated.
2. The usefulness of these data for homing in on strong candidate functional SNPs under GWAS peaks, for neuropsychiatric disorders and, for that matter, any common genetically complex human phenotype, is straightforward and builds significantly on previous work in this area. This aspect of the paper would benefit from adding one or two locus-specific CRISP-Cas9 mutagenesis/deletion experiments as definitive functional validations, as detailed below.

3. The authors' conclusions regarding specific biological connections between placental mQTLs (and eQTMs) and neuropsychiatric disorders are intriguing, but somewhat indirect. We also make some suggestions below for how they might bolster this aspect.

Response: We thank the reviewers for their positive comments and interesting

suggestions. SUGGESTIONS/REQUESTS:

1. Most of the exposures in pregnancy that are thought to increase the risk of psychopathology in the offspring, need to occur in early pregnancy (eg. Alan Brown's work with SCZ). However, the placentas in the current study were collected at term. Could the authors obtain pre-term placentas (obviously it would be a smaller series) and determine if their high-priority disease-linked mQTLs and eQTMs are stronger at the earlier stage? This type of information, if positive, could bolster their argument about a connection to neuropsychiatric disorders, as per the DOHaD hypothesis.

Response: This is a very interesting point but, unfortunately, ours is a pediatric prospective cohort in which placenta sample collection was carried out about 15 years ago, and the specimens were collected in order to follow the children during development. Thus, non-term placentas (those from spontaneous abortions that could have been a source of non-term placentas) were not included in the cohort. At the moment, we do not have access to such a sample set and therefore, we are not able to test this. Nevertheless, we have checked if the SMR significant hits from the three main disorders overlap with the list of CpGs provided in Y. Lee at al.2019, where the authors reported 10,827 CpG sites exhibiting a genome-wide significant correlation with gestational age (GA). We have found 1/29 and 6/188 SMR hits from BIP and SCZ, respectively, overlap with the list of CpG sites associated with GA. This could mean that the majority of the SMR hits obtained for the three main disorders do not change during pregnancy but, as we do not have first trimester placentas, we cannot confirm this. We have added this consideration in the manuscript as follows.

In Results section "*Multi-omics approaches to unravel the placental origin of neuropsychiatric traits and disorders*": "*Additionally, most of these hits did not overlap with CpGs reported to change as a function of GA[48], suggesting relative stability in the DNAm levels during pregnancy*".

In Bibliography:
"*Lee, Y., et al. Placental epigenetic clocks: estimating gestational age using placental DNA methylation levels. Aging (Albany NY) 11, 4238-4253 (2019)*"

2. The authors hypothesize that the putative connection to neuropsychiatric disorders might occur via the effects of mQTLs on the expression of genes in immune/inflammatory pathways. However, their data analysis considered proportions of cyto- and syncitio-trophoblast but not the (relatively abundant) populations of myeloid and lymphoid cell types in term placentas. Information from a similar cell type deconvolution approach evaluating the relative percentage of such cells (placental macrophages, for example) and the influence of that percentage on the observed mQTLs, particularly the disease-associated ones, could also bolster their argument.

Response: We understand the suggestion by the reviewers but, as responded to Reviewer 1, the point here is that iQTLs, as defined by GTEx, are only computed in the most abundant cell type in each tissue, due to power constraints. In fact, while two of the three reviewers have suggested to extend the imQTL mapping to other cell types, one has suggested to report only those

calculated for STB, as it is the most abundant cell type. As a consequence, we have decided not to report imQTLs for more cell types, other than STB and TB. We hope that the reviewers understand this decision as a midpoint, given that we are already being quite flexible interrogating another cell type other than the most abundant.

3. Regarding the post-GWAS application of their data for pinpointing functional regulatory SNPs ("rSNPs"), the paper would be dramatically strengthened if they could show, using CRISPR-Cas9 in convenient human cell lines (trophoblast or other; cancer lines would be adequate for this aspect) that their highest priority candidate rSNPs in fact control the expression and methylation of their highest priority candidate disease-linked genes – for example, LRFN5 (linked to MDD) and ZSCAN23 (linked to SCZ). Such experiments are straightforward to do, with precise site-directed mutagenesis being the most elegant approach, but with creation of small deletions spanning the relevant SNPs, followed by gene expression analysis (RNA-seq and/or RT-Q-PCR) in deletion-bearing vs wild-type cells being adequate.

Response: This is indeed a great idea but unfortunately we do not have this expertise in our lab and moreover, we consider that the publication of our data as it is now will encourage related, functional research from expert labs. However, we do think that our work will need to be supplemented with the employment of approaches as the ones proposed by the reviewers and thus, have added this idea as part of the conclusions: "*Future work, including functional approaches, as well as long term follow up in prospective studies, will be needed in order to better characterize our findings and their implication in disease development*".

Reviewer #3 (Remarks to the Author):

Cilleros-Portet et al. have performed cis-mQTL and interaction cis-mQTL mapping in 368 fetal placenta samples. The (i)mQTL results were then integrated with the largest GWASes on 10 neuropsychiatric disorders using SMR and colocalization analyses. The manuscript highlights interesting data. However, the selected methods do not follow the best practices in the field. There are also several other concerns that should be addressed.

Response: We thank the reviewer for such an accurate and rigorous revision. We are especially thankful because he/she noticed and pointed out an error in one of our supplementary tables that, although does not dramatically alter the results, was mandatory to correct. We consider this observation by the reviewer of outmost importance in order to improve the quality of our manuscript. We really hope that the reviewer will be satisfied with the improvements, and we are open to discussion.

Major points:
1. Discovery of cis-mQTLs and interaction cis-mQTLs. The authors account for fetal sex, 5 genotype PCs and Planet-estimated cell types in their cis-mQTL analysis. As there are several technical and biological factors that could affect the methylation or gene expression levels, it is common in molQTL analysis to include latent factors (such as PEER factors or PCs) to account for unwanted variation in the data. Also, to avoid outlier effects inverse normal transformation is applied on the molecular data. These steps are followed by the GTEx Consortium (that the authors refer to) and recommended by the Aguet et al 2023 Nature Reviews Methods Primer paper.

Response: The reviewer is very right when pointing out that we have taken the GTEx Consortium as a reference, but we would like to highlight several points: first, we take GTEx as

our main reference in the sense that we took the idea of calculating iQTLs from them. Also, we learnt about TensorQTL by reading their paper (Kim-Hellmuth et al.2020) and talked to one of the authors to better classify and interpret our imQTLs (i.e., Dr. Diego Garrido-Martin, who is now a co-author of the present work). However, we did not base our model in theirs because they work with expression data. Our main reference in this sense was GoDMC, that is, the largest mQTL database to date, calculated in blood (doi: 10.1038/s41588-021-00923-x). In the Methods section of that paper, the authors indicate that "***for unrelated participants we regressed out age, sex, predicted cell counts, predicted smoking and genetic PCs***". They additionally declare that they used 5 DNAm PCs only when some relatedness was present in a given cohort, but this is not our case because: i) no related participants were admitted, and ii) identity-by-descent values were calculated with PLINK, and from those sample pairs that showed PI-HAT estimates above 0.18, the sample with higher proportion of missing genotypes was removed. Therefore, we have employed the same model as them except for age (all our samples come from term placentas and we control for gestational age specifically in the GA-imQTL approach), and the smoking status because our participants are fetuses. We considered to take into account the smoking information from mothers but we did not include it as a covariate because on the one hand, very few pregnant mothers in our cohort were smokers, and on the other hand, we run a sensitivity analysis including maternal smoking that revealed a very limited impact in our mQTLs:

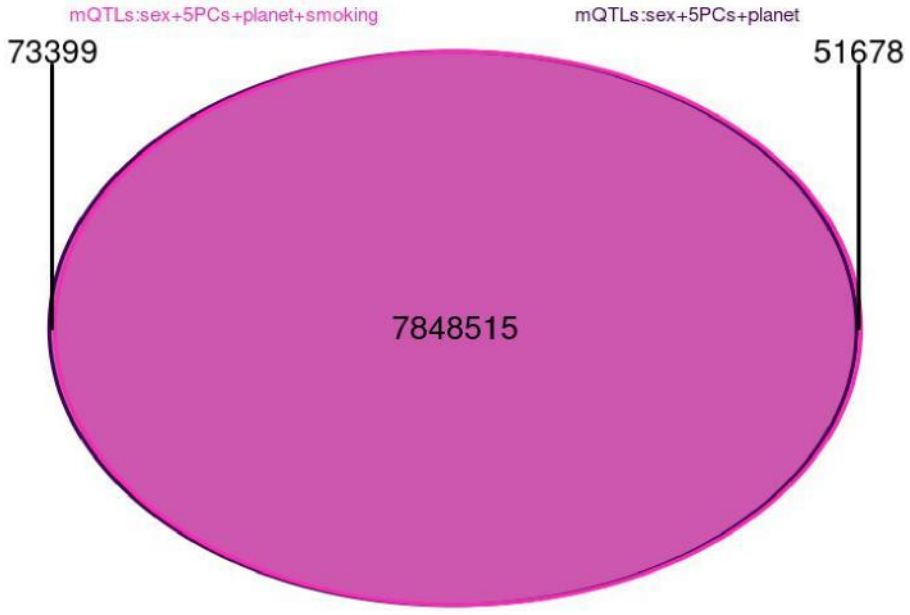

Figure 1 (for reviewers only, unless they ask us to include it in the manuscript): Venn Diagram showing the overlap between the mQTLs calculated w/o maternal smoking as a covariate.

However, we think that still, the concern by the reviewer is justified and we have calculated our mQTLs by adding 5 DNAm PCs to our model. As a result, we have obtained a big overlap although as expected, adding the 5 DNAm PCs yields more associations. Also, the Pearson correlation between effect sizes from the two models is extremely high and significant ($r=0.999$, $p<2.2x10^{-16}$). This was also observed in the abovementioned GoDMC paper, where they compared the results obtained w/o the inclusion of the DNAm PCs in one of the participating cohorts. They stated that "a***s adjustment for non-genetic DNAm PCs might have substantial benefits on power or an adverse effect by inducing collider bias, we explored the impact by comparing mQTLs not***

adjusted for non-genetic PCs with mQTLs adjusted for non-genetic PCs in ARIES. Specifically, we found 80,890 clumped mQTL associations in the PC-adjusted dataset and 74,402 clumped mQTL associations in the PC-unadjusted dataset. The Pearson correlation between effect sizes of the PC-unadjusted clumped mQTLs vs PC-adjusted mQTLs (cis r=0.998; trans r=0.998) and PC-adjusted clumped mQTLs (cis r=0.997; trans r=0.997) versus PC-unadjusted mQTLs was very high. These results suggest that if collider bias is impacting the results, it is extremely small. The simplest explanation for the minimal difference in effect sizes and slightly higher mQTL yield amongst the PC-adjusted mQTLs is that reduced residual variance has improved power". However, they preferred to be conservative and performed the downstream analyses without including the DNAm PCs, in the case of the cohorts with unrelated participants only.

In conclusion, we propose: i) to provide the full mQTL data with the 5 DNAm PCs as a downloadable file in our shiny database (please visit https://smari.shinyapps.io/shiny_mqtl_placenta/); ii) to add a new supplementary figure (new Supplementary Figure 6) with the overlap in terms of both CpGs and SNPs between the two models, and iii) to report in the main text the high correlation and overlap between the two models and the decision of continuing with the most conservative one but provided that both are publicly accessible online. We hope that this solution satisfies the reviewer.

Apart from adding new Supplementary Figure 6, we have included the following changes in the manuscript:

In results section "*Placental cis-mQTL characterization*": "*Finally, we calculated an additional nominal cis-mQTL database, this time including 5 non-genetic DNAm PCs apart from the abovementioned covariates. We obtained a total of 9,080,030 mQTLs (123,946 unique CpGs and 2,014,144 unique SNPs), most of which overlapped with the ones obtained in the prior nominal model but yielding 14.72% more associations (Supplementary Figure 6). Also, the Pearson correlation between effect sizes arising from the two models was extremely high and significant ($P < 2.2x10^{-16}$, $r = 0.999$). As previously stated by other authors, these results suggest a minimum collider bias (if any) and an improved statistical power very probably derived from the reduced residual variance after adjusting for DNAm PCs*[33]*. However, we decided to work with the most conservative set of results in all downstream analyses (nominal cis-mQTL database, without DNAm PCs), as done in previous large mQTL catalogues in cohorts of unrelated participants like ours*[33]*.*

The four complete placental cis-mQTL databases are publicly available online in the following address: https://smari.shinyapps.io/shiny_mqtl_placenta/".

In Methods, section "*Placental cis-mQTL analysis*": "*Five non-genetic DNAm PCs were also included in a secondary nominal cis-mQTL analysis*".

In Bibliography:

"*Min, J.L., et al. Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. Nat. Genet. 53, 1311-1321 (2021)*"

Second, we know that transforming the methylation beta-values is something very reasonable to do but again, mQTL papers doing so are relatively new (like Oliva et al.2023), and we based this study on prior literature such as GoDMC and Hannon et al.2015, that make use of untransformed DNAm data. Moreover, it is widely accepted to use DNAm beta-values in

different approaches given that using these values yields much more biologically interpretable results (i.e., most of the DNAm studies by the PACE consortium). It is true that we may have a more limited detection rate in extreme DNAm values (that on the other hand, are probably less prone to participate in mQTLs), but we are confident that our strict QC of DNAm data, together with the winsorization of outliers, will help avoid most false positive results. Therefore, we think that although this topic may be open for discussion, the use of beta-values is based on a large amount of solid literature and should not be a major problem.

2. Multiple testing. The authors thoroughly characterize mQTLs with P_nominal < 5e-8. A proper multiple testing procedure should be incorporated to cis-mQTL and interaction cis-mQTL discovery. Only mQTLs and imQTLs with FDR < 0.05 should be characterized and included to downstream analysis.

Response: The reason why we fixed a p-value threshold of $5e^{-8}$ is that SMR, the software that we use for our core analysis, recommends to include all QTLs below this value, as suggestive threshold of genome-wide significance. The software then gives the opportunity to fix a more stringent significance cutoff for the analysis, but results are more powered if all QTLs under this threshold are included. One of our main aims in this work is to create a resource that is useful to the scientific community and thus, we decided to provide and characterize all those mQTLs below $p=5e^{-8}$, but researchers will be able to adjust further, depending on the tools they want to use to run particular studies. In summary, our intention was never to overclaim our findings by fixing a too lax threshold, but to be as flexible as possible to provide a large dataset that can be further filtered depending on particular needs.

However, we performed FDR correction of our mQTLs as suggested by the reviewer. With the FDR<0.05 correction, out of the 1,132,596,401 possible CpG-SNP combinations tested in 0.5 Mb windows in the whole genome, 27,636,581 mQTL associations appeared to be significant, compared to the 7,921,914 that were suggestive of genome-wide significance under a nominal p-value of $5e^{-8}$. We humbly think that this is a huge number, even difficult to manage in any characterization tool, and that probably is hiding a number of false positive associations. We are obviously open to further discuss this issue with the reviewer but, in the present revision, we propose to keep our suggestive threshold.

3. Sharing between iQTLs. How were common STB-imQTLs and TB-imQTLs found? Are iVariants in high LD affecting the methylation levels of the same CpG site considered to be common to both cell types? Storey's pi1 estimate would be a better estimate of sharing between STB-imQTLs and TB-imQTLs (it is not influenced by p-value cut-offs to call an imQTL a significant imQTL. There might be TB-imQTLs that are shared with STB-imQTLs just below the significance level due to statistical power). Same comment for estimating the overlap between STB/TB-imQTLs and GA-imQTLs.

(and minor 11. As there is such a strong negative correlation between STB and TB, it is expected that there is negative correlation between the effect sizes of common STB- and TB-imQTLs (and in general a high sharing between STB-imQTLs and TB-imQTLs). It seems reasonable to focus only on the cell type with the highest statistical power).

Response: Common STB-imQTLs and TB-imQTLs were found by overlapping significant imQTLs at a suggestive significance threshold of $5e^{-8}$ that had been separately calculated. However, we find the idea put forward by the reviewer interesting and have redone our analysis following his/her suggestion. First of all, we tried to redo the analysis by filtering by Storey's p1

estimate calculated with the qvalue() function from the R package qvalue, as in PMID: 23671422 and the original paper in which the method was described (PMID: 12883005). Unfortunately, our dataset was too large to run in this package. Alternatively, we used the p.adjust function in R and adjusted for FDR (after making sure that estimates in the two packages are equivalent, as shown in Figure 2, where both estimates for mQTLs of chromosome 22 were compared).

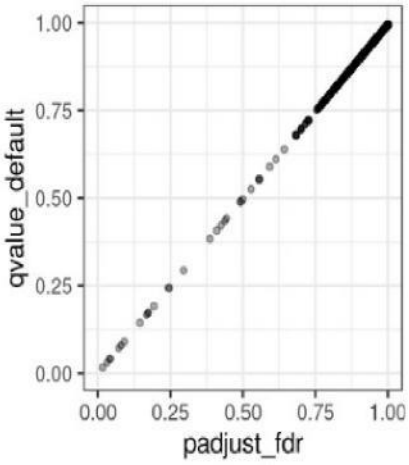

Figure 2 (for reviewers only, unless they ask us to include it in the manuscript): comparison between the estimates calculated with the qvalue() function from the qvalue R package and with the FDR correction of the p.adjust() R function for chromosome 22.

Afterwards, we reported the imQTLs filtered by FDR<0.05. Compared to our previous imQTLs (filtered by the suggestive genome-wide significance threshold $5e^{-8}$), we found a higher number of STB-imQTLs, but a lower number of both TB- and GA-imQTLs. This was expected due to the fact that STB is the most abundant cell-type (and therefore, the best powered for interaction analyses). Another observation is that we still find a relatively moderate overlap between STB-and TB-imQTLs (with a strong negative correlation between effect sizes of r = - 0.99, as with our previous method), while the overlap with GA-imQTLs keeps being non-existent. Convinced by the increase of power in the main cell type, we decided to proceed with the new correction and threshold, including the SMR approach. We have redone all the figures and supplementary material accordingly, and changed the manuscript as follows:

In Results section "*Placental cell type-, GA- and sex-imQTLs*": "*The most abundant cell type in fetal placenta are syncytiotrophoblasts (STB). During gestation, undifferentiated trophoblasts (TB) change into fully differentiated STB, a continuous, multinucleated and specialized layer of epithelial cells*[34]. *We estimated the cell-type proportions of our samples by using the reference-based method Planet*[26]. *As expected, the estimated content of STB was negatively correlated with the estimated proportion of TB in our placental samples ($P < 2.2x10^{-16}$, r = -0.89). In turn, STB content was positively correlated with gestational age (GA) ($P < 4.5x10^{-5}$, r = 0.21) (Figure 2A and 2B). We calculated STB-, TB- and GA-imQTLs, as proxies for STB-, TB- and GA-specific cis-mQTLs, and obtained 772 positive and 420 negative STB-imQTLs, 243 positive and 187 negative TB-imQTLs,*

and 186 GA-imQTLs ($FDR_{interaction}$ < 0.05) (Supplementary Data 4). We decided to filter by FDR < 0.05 with the aim of using a better estimate of sharing between STB-, TB- and GA-imQTLs. In the case of cell type-imQTLs, positive imQTLs showed increased DNAm related to the main genotype as a function of cell proportion, while negative imQTLs showed a decrease. The higher amount of STB-imQTLs, revealed a higher statistical power for the most abundant cell type, as previously stated by Kim-Hellmuth et al[24]. Out of the 1,192 STB-imQTLs and 430 TB-imQTLs, 136 were common to both cell types (Figure 2C). Considering the same allele for both STB and TB, the effect sizes of the shared cell type-imQTLs were negatively correlated ($P < 2.2x10^{-16}$, r = -0.99) with opposite effect directions as in the example in Figure 2D, E and F. Of note, most of the overlapping imQTLs were positively correlated with STB content, and negatively correlated with the TB abundance (120/136). In contrast, no overlap was observed with GA-imQTLs (Figure 2C). The effect of GA on the genetic regulation of placental DNAm seemed to be independent of the cell type composition of each sample and therefore, rather a matter of time of gestation regardless of the TB to STB transition. Finally, as placental DNAm could differ between sexes, we also mapped sex-imQTLs and discovered 521 associations at a genome-wide suggestive significance level of $P_{interaction} < 5x10^{-8}$ (Supplementary Data 4).".

In Results section "*SMR analysis of placental cell type-, GA- and sex- imQTLs in BIP, MDD and SCZ*": "*We used SMR to combine imQTLs with the BIP, MDD and SCZ GWAS, and obtained three STB-imQTLs (one each in BIP, MDD and SCZ), and three TB-imQTLs (one each in BIP, MDD and SCZ) (Supplementary Data 10). Remarkably, we found an imQTL involving cg27130493 and rs11637581 (Figure 6). cg27130493 appeared to be pleiotropically associated with BIP while it changed its methylation levels as a function of TB proportion, in a rs11637581 genotype-dependent manner. This CpG is located in the gene body of SMAD3. The placental overexpression of this gene has been described to activate the ability of TB to form endothelial-like networks, while its defect has been associated with pre-eclampsia[61]. Additionally, it is a well-known target for lithium treatment in BIP[62]. Interestingly, rs11637581 is not directly associated with BIP in the original GWAS, suggesting a possible causal association between placental DNAm and BIP at this CpG site.*

Finally, we employed SMR to combine sex-imQTLs with the BIP, MDD and SCZ GWAS, obtaining a common hit for BIP and SCZ, and an additional hit for SCZ (Supplementary Data 10). The common sex-imQTL involved cg09779537 and rs2398668 (Supplementary Figure 11). SNP rs2398668 is borderline significant in the SCZ GWAS, and non-significant in the BIP GWAS. The presence of the minor allele T is associated with reduced methylation of cg09779537, specifically in males. The imQTL-SNP is in a region with high gene density, in the gene body of SNX8. This gene has been reported to act as a modulator of innate immunity in response to viral infections[63] and its disruption has been related with mild intellectual disability[64]. Overall, the data suggest a potential causal relationship between the associated CpG site, and BIP and SCZ".

Please note that we are losing some of our prior SMR results but still keep the same CpG site for BIP. Moreover, we have included sex-interacting mQTLs as requested by reviewer 1, with interesting results that are now included as Supplementary Figure 11.

Finally, we have included minor point 11 here due to its relationship with major point 3. We sincerely understand the point of the reviewer and would like to explain our motivation to include TB, as well as our doubts after carefully reading the suggestions by the different reviewers. Reviewers 1 and 2 asked us to include more cell types but we argued that, in principle, cell type-interaction should only be run with the most abundant cell type. Thus, the reviewer is right when pointing out that probably, the most correct, straightforward thing to do would be to

remove TB-imQTLs. Our aim when adding TB-imQTLs was to compare STB-, TB- and GA-interaction effects, to see whether GA has an effect itself or whether its effect relies solely on the TB to STB transition across gestation. Given that GA seems to have its own impact, and also given that there are some (lesser than for STB but still some) positive TB-imQTL hits, we decided to report the results with the main aim of making public our datasets for the scientific community. In this sense, we have no objection if the reviewer thinks that the most conservative option is to exclude the TB-imQTLs, but we are not persuaded that this would be the best solution and would like to know his/her opinion.

4. Colocalization analysis: there seems to be a mistake in counting the region-CpG pairs with support of colocalization. For example, Supplementary Data 6 (BIP tab) indicates that there is strong support for H3 in coloc analysis between region 43772251 - 43592251 and CpG cg24283270 (PP3 = 0.999938693 and PP4 = 4.78e-05). Please correct this mistake. Also, it would be informative if lead variants in imQTL and GWAS analysis are highlighted in the table together with the p-values from the respective analyses. Maybe also genomic position of the CpG site could be included.

Response: We really thank the reviewer for this comment since we were not aware of this error in the filtering of our colocalization results. We very sincerely apologize for this technical error and hope that, although some of the hits at the intersection between the three approaches (SMR, *coloc* and eQTMs) have fallen as a consequence of the correction, the reviewer will be satisfied with the modifications included in the manuscript.

**First, we have changed the colocalization results in Results section "*Multi-omics approaches to unravel the placental origin of neuropsychiatric traits and disorders*" as follows: "*To replicate pleiotropic associations from the SMR analyses, we performed a colocalization test with the same data. The Bayesian colocalization analysis implemented in the coloc R package[49] focuses on finding the intersection between significant variants independently associated with two phenotypes, comparing the association patterns in the GWAS and QTL analyses across genomic regions, and thus combining the summary statistics into posterior probabilities for five hypotheses (see Methods for more information). Colocalization was performed across 64, 102, and 287 genomic regions defined in each GWAS associated with BIP, MDD and SCZ, respectively. These regions spanned 110,721 placental DNAm sites from the nominal cis-mQTL database. In BIP, the posterior probabilities for 46 regions, involving 374 DNAm sites in 420 region-CpG pairs, were supportive of a colocalization signal (PPA4>0.8) (Figure 3A and Supplementary Data 6). In MDD, the posterior probabilities for 51 regions, involving 217 DNAm sites in 228 region-CpG pairs, were supportive of a colocalization signal (PPA4>0.8) (Figure 3B and Supplementary Data 6). Finally, in SCZ, the posterior probabilities for 182 regions, involving 1,351 DNAm sites in 1,446 region-CpG pairs, were supportive of a colocalization signal (PPA4>0.8) (Figure 3C and Supplementary Data 6). When the overlap with SMR hits was assessed, out of the 29 SMR hits in BIP, 19 DNAm sites (65%) showed evidence of colocalization with BIP. In MDD, from the 27 SMR hits, 12 DNAm sites (44%) showed evidence of colocalization with MDD. And, from the 188 SMR hits in SCZ, 63 DNAm sites (33%) showed evidence of colocalization with SCZ.*

[...]

Next, we identified the intersect between SMR, colocalization and eQTM results for each trait (Supplementary Figure 7). In the case of BIP, two placental DNAm sites that colocalized with two GWAS loci on chromosomes 3 and 15 showed DNAm levels correlated with the expression of two nearby genes (ITIH4 and ZNF592). For MDD, two placental DNAm sites colocalized with a single

MDD-associated genomic locus on chromosome 14, with DNAm levels correlating with the placental expression of LFRN5. Finally, in SCZ, 8 placental DNAm sites colocalized with 10 SCZ-associated loci in chromosomes 3, 6, 7, 11, 18, 19 and 22, with DNAm levels that correlated with the placental expression of 8 nearby genes (FXR1, GLB1L3, IRF3, NAGA, PSMG3, SLC6A16, TXNL4A and VARS2)".

We want to highlight that, out of the 8 genes at the intersection of the three approaches, four have been previously linked to SCZ through placental expression by other experts in the field. We have added this appreciation in Discussion as follows: "*Remarkably, four out of the eight genes whose expression correlates with DNAm sites associated to SCZ according to the different approaches presented here, were among the 248 placenta-enriched SCZ genes described by Ursini et al. (i.e. FXR1, IRF3, NAGA, and PSMG3). Therefore, we propose that these genes could not only mediate SCZ risk through placental expression but, more precisely, through DNAm-regulated placental expression patterns*".

Second, due to the limited number of genes remaining at the intersection among the three approaches, the significance of immune-related terms has fallen to 0.051 in the Reactome pathway analysis. However, and inspired by the suggestion by reviewer 1 of performing the pathway analysis with all the SMR-eQTM-genes, we have provided new results that enable to improve the biological interpretation of the data. Please see changes in the main text, as well as new Supplementary Data 8.

In Results section "*Multi-omics approaches to unravel the placental origin of neuropsychiatric traits and disorders": "Finally, we performed a Reactome gene-set analysis[58] of the 26 genes at the intersection between the SMR and the eQTM approaches in SCZ. We observed enrichment for 115 gene-sets (Supplementary Data 8), including many epigenetic regulatory pathways, such as different histone modifications, DNAm, and chromatin condensation and organization, suggesting that SCZ risk alleles could change the epigenetic landscape of placenta through the regulation of different histone-coding genes. Additionally, we observed a remarkable enrichment in immune-related pathways such as interferon alpha/beta signaling, interferon gamma signaling, and interferon signaling, as well as human cytomegalovirus (HCMV) infection and overall viral infection, reinforcing the idea of MIA as a link between placenta and SCZ risk*".

In Discussion: "*Regarding the genes affected by placental DNAm pleiotropically associated to SCZ, we want to highlight that they were enriched in epigenetic regulation pathways and immune-related terms, including interferon signaling and viral infection. On the one hand, the multiple epigenetic pathways involved suggest that SCZ risk alleles could cause deep epigenetic changes in placenta as a result of DNAm modifications in close proximity to different histone-coding genes. On the other hand, the previously mentioned histone-coding genes, together with others such as HLA-H, IRF3 and VPS37B, were some of the most relevant genes contributing to the enrichment in immune-related biological routes. This reinforces the idea of MIA being implicated in the neurodevelopmental origins of SCZ. Particularly for HCMV, it has been reported that the DNAm state of the host genome can make cells more or less prone to infection, and that histones participate in this process because, although HCMV is devoid of these proteins, it captures them from the host for the chromatinization of its own genome[81]. Moreover, it is known that congenital HCMV infection can cause long-term clinical manifestations such as hearing loss, neurodevelopmental disorders, ophthalmic complications, and ASD, among others[82]. The virus moves from mother to fetus by infecting TB in the placenta, altering the development and integrity of the organ, and eventually, inhibiting placental cell differentiation, self-renewal, and migration*".

In Bibliography:
"*Esteki-Zadeh., A, et al. Human cytomegalovirus infection is sensitive to the host cell DNA methylation state and alters global DNA methylation capacity. Epigenetics. 7, 585-593 (2012)*"

"*Elgueta, D., Murgas, P., Riquelme, E., Yang, G., Cancino, G.I. Consequences of Viral Infection and Cytokine Production During Pregnancy on Brain Development in Offspring. Front. Immunol. 13, 816619 (2022)*"

Third, after the correction of the colocalization results, we realized that *coloc* PP4 > 0.8 was not a useful threshold anymore in our search for secondary signals by means of the conditional analysis. This was because the overlap between SMR hits failing HEIDI and *coloc* PP4 > 0.8 decreased dramatically after the new correction in relation with our computing potential. In fact, we were only capturing a 6% of the total SMR hits that failed HEIDI (29/469). In this context, we contacted Dr. Claudia Giambartolomei from UCLA (developer of the *coloc* package) and asked directly how would she relax this threshold if her aim was to fine-tune the selection among those SMR hits that failed the HEIDI test. She told us that filtering for PP3+PP4 > 0.8 would be reasonable as this threshold is suggestive for association with both traits and could contain independent signals other than the principal. With this new threshold, we were able to include 372/469 SMR hits failing HEIDI, that is, the 79%. We have indicated this new parameter in the manuscript, particularly in Results, section "*Conditional SMR analysis of placental cis-mQTLs, and BIP, MDD and SCZ*": "*We applied this approach to those CpGs that passed the SMR test but failed to pass the HEIDI test. Nevertheless, due to power constrains, only those CpGs with suggestive support for association with both traits according to coloc (PP3+PP4 > 0.8) were considered*". The results of the conditional analysis keep being the same as in the original manuscript.

Finally, we have added the lead variants and the genomic positions of the CpG sites in Supplementary Data 6, as requested by the reviewer.

5. Processing of DNA methylation data. Is applying both noob background correction with dye-bias correction followed by functional normalization and beta-mixture quantile normalization necessary? May it result in over-correction of the data? Do I understand correctly that probes with SNPs that have MAF < 5% in European samples (line 640) are excluded? One should want to remove probes with common SNPs in the probe sequence to avoid false positive mQTL hits. How are cross-hybridizing probes defined?

Response: On the one hand, the pipeline for the DNAm QC used in the present work is the pipeline commonly used in the PACE Consortium. Among others, our last two EWAS studies published in Nat Commun and Commun Biol have used the same pipeline (PMID: 34429407 and PMID: 36446949, respectively). Additionally, other important works that are being conducted in the Consortium right now are also using this approach, and it has recently being compiled into a R package called PACEAnalysis (https://github.com/epicenteredresearch/PACEanalysis).

However, we are aware that many authors and experts in the field use different combinations of packages with comparable results. In our case and given that: i) Noob is a background correction method with dye-bias normalization for Illumina Infinium methylation arrays that works in an intrasample manner; ii) functional normalization removes unwanted technical variation using control probes in an intersample manner, and iii) BMIQ is an intrasample normalization procedure, specifically correcting the bias of type-2 probes, we considered the three packages

complementary to one another, as they do different things and operate in different ways. In fact, Noob + functional normalization is probably one of the most widely used combinations, as it is the one recommended by Minfi (https://pubmed.ncbi.nlm.nih.gov/25599564/). We do not expect the addition of BMIQ to have a great impact since it is a very specific and widely used package. In summary, we believe that our pipeline is not expected to over-correct the data. Maybe it is more strict than others but it is robust and widely supported by literature in the field.

On the other hand, we apologize because in fact we did exclude SNPs with a MAF in European samples ≥ 5% (instead of MAF < 5%). We have corrected this mistake in the text. Finally, we apologize for not including the reference from which we took the cross-hybridizing probes excluded from our analysis. The reference is the following and has been included in the bibliography (and cited in the corresponding part in Methods):

"*McCartney, D.L., et al. Identification of polymorphic and off-target probe binding sites on the Illumina Infinium MethylationEPIC BeadChip. Genom. Data 9, 22-24 (2016)*"

Minor points:

1. Line 67: what are those eight million mQTLs? Are they significant mQTLs at some significance threshold?

Response: Yes, these eight million mQTLs are the significant *cis*-mQTLs considering a $P_{nominal} <$ 5e-08, as defined in the third line of Results. To clarify this, we have specified it in the Abstract, as follows: "*We constructed the first public placental cis-mQTL database including nearly eight million mQTLs ($P_{nominal} < 5x10^{-8}$) calculated in 368 fetal placenta DNA samples...*".

2. What are nominal mQTLs and nominal mQTL database? Please be more clear.

Response: Nominal mQTLs and nominal mQTL database are essentially the same. When we use the term "database" it is usually to talk specifically about the resource. We tried to make sure this is the case across the manuscript.

3. Figure 1a: how is the median distance calculated? I assume it is the median value of the absolute distance between the CpG site and the SNP. If yes, then showing median distance as a red line on Fig. 1b is misleading, because the plot summarizes both the cases where the SNP is located before or after the CpG site.

Response: In Figure 1a, the distribution of the distance is two-sided, reflecting the two possible scenarios (mQTL-SNPs located up- or down-stream the CpGs). The median has been computed with the raw distances, instead of with the absolute values. This has been clarified both in the text and in the figure legends.

4. Figure 1b: I disagree with the conclusion that "the distribution of mQTLs across chromosomes was in line with the chromosome length, except for chr 6". For example, there are considerably fewer mQTLs on chr 9 as compared to chr 10-12, which are smaller in size.

Response: The reviewer is right, we have change the manuscript accordingly.

Results section "*Placental cis-mQTL characterization*": "*The distribution of mQTLs across chromosomes was in line with chromosome length, with exceptions such as chromosome 6, where a*

higher number of CpG-SNP pairs was observed (883,548 CpG-SNP pairs with 8,173 and 153,743 unique CpGs and SNPs, respectively) (Figure 1B), and chromosome 9, where there were less mQTLs than in chromosomes 10 and 11. A higher number of mQTLs in chromosome 6 was expected...".

5. Figure 1c: based on the volcano plot is hard to evaluate whether the effect sizes of the mQTLs follow uniform distribution or not. Bar plot would be better for that purpose.

Response: We have tried to avoid speculation in this sentence and replace it by this one: "*Additionally, the Haplotype Reference Consortium (HRC) reference alleles in placental cis-mQTLs presented both positive and negative effect directions (Figure 1C)*".

6. Figure 1d and 1e: I found these pie charts to be hard to read. Would it be easier for the reader if the data is replotted using bar plots, for example? Also, then one could highlight significance differences in distributions more easily.

Response: We have redone the figure with barplots, as suggested by the reviewer (also analogue Supplementary Figures 3 and 4).

7. Line 198: as there are enrichments in many other DNase I hotspots too, I would stress that the strongest signal is observed in fetal placenta. Same comment for Suppl. Fig. 1A.

Response: The reviewer is right. We apologize for this mistake. We have corrected the sentence as follows: "*Using eFORGE[29,30,31], we were able to detect an enrichment in DNase I hotspots and H3K4me1 broadPeaks of different tissues, fetal placenta being the one showing the strongest signal (Figure 1G, Supplementary Figure 1A and Supplementary Data 2)*".

8. Supplementary Figure 2c: is the adjusted p-value the same for all the tested diseases? It is an unexpected result.

Response: This is due to the fact that this significance level is the highest reported by the package. All the diseases shown in this figure share the lowest p-value. However, p-values are different from row 100 on in Supplementary Data 3 (but were not included in the Supplementary Figure 2C).

9. Line 206: what are permuted and conditional cis-mQTLs?

Response: We have added a short explanation in Results (apart from Methods) as follows: "*Briefly, the permuted cis-mQTL approach corrects for multiple correlated variants tested via a beta approximation while, the conditional database uses the permuted QTLs to perform a stepwise regression procedure and map conditionally independent cis-QTLs (see Methods for more details)*".

10. Figure 2b and Figure 2c: labels are mixed up in the text (line 227 and line 233).

Response: We thank the reviewer for pointing this out. We have corrected this error in new Figure 2.

12. What data is used to calculate LD between SNPs on the locusZoom plots?

Response: We used 1000G as suggested by the locusZoom web page. We have detailed this in the main figure legends as well as in Methods section "*Genome Wide Association Studies*" as follows: "*All locusZoom plots were done as suggested in their web page (http://locuszoom.org/), using the LD panel of 1000G*[119]".

In Bibliography: "Boughton, A.P., et al. LocusZoom.js: interactive and embeddable visualization of genetic association study results. *Bioinformatics* 37, 3017-3018 (2021)".

13. Figure 3 is not mentioned in the main text.

Response: We apologize for this error. We have mentioned Figure 3 in the main text.

14. Figure 4 and lines 305-314: there seems to be very strong LD between the mQTL-SNPs rs1111179 and rs61990289 based on the panel a. Thus, I would suspect that the methylation levels of the two CpGs sites are also strongly correlated. Thus, they represent the same signal rather than two independent, pleiotropically associated CpG sites. What is the correlation between the two CpG sites?

Response: The correlation between the methylation values of the two CpGs (cg10318063 and cg32217097) is negative (r= -0.17 and P = 0.0007).

15. What does it mean that the regression coefficients, their variances and SNP minor allele frequencies were imputed for the coloc analysis?

Response: It means that the statistics were taken from the original data; for instance, the SNP minor allele frequency from the mQTLs for the *coloc* analysis were subtracted from the mQTL data, and the same statistic from the GWAS data was taken from the GWAS itself. This is also applied for regression coefficients and variances. But, we agree that this sentence is quite confusing, therefore we have decided to reword it as follows: "*We retrieved the regression coefficients, their variances and the SNP minor allele frequencies from the original GWAS data and our mQTLs*".

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

I think the manuscript is improved and the authors have performed most of the required work to address the reviewers' comments. I have just few remaining comments and suggestions:

1. I appreciate that the authors have performed further analyses to test the relationship between GWAS sample size, GWAS hits, and SMR hits. However, I do not agree about their interpretation of the results, and some details of these analyses are not clear. First, the authors comment that there is no correlation about GWAS sample size and SMR hits, although when they test this relationship the r is 0.52. This is indicative of an almost moderate correlation. The p-value is affected by sample size that in this analysis is 10, so it's recommendable to not base any conclusion of the fact that a p-value of 0.13 is higher than the 0.05 threshold. It is evident in figure supplementary 13 a and b that the 3 diseases with the bigger GWAS sample size, and higher number of GWAS hits, have also the higher number of SMR hits.

2. Also, from supplementary figure 13b it seems that there are '-10,000 GWAS hits for depression and '-2,500 GWAS hits for bipolar disorder, however bipolar GWAS identified 64 loci (ref.35) and depression GWAS identified 102 loci (ref.37). It's important that the authors double check everything to be sure that they are reporting correct information. Indeed, the number of GWAS hits reported in Table 2 does not seem right, at least for ADHD, autism, bipolar disorder, and depression.

3. I suggest to add labels withing the plots of supplementary figure 13, so that each disease can be clearly identified.

4. I appreciate the explanation provided by the author about the colocalization analyses. In the response they say that "we applied this analysis to identify the overlap between causal mQTLs and GWAS-significant loci", however in the main text they write that "The Bayesian colocalization analysis implemented in the coloc R package focuses on finding the intersection between significant variants independently associated with two phenotypes, comparing the association patterns in the GWAS and QTL analyses across genomic regions, and thus combining the summary statistics into posterior probabilities for five hypotheses (see Methods for more information)." So, I think they should write in the main text they applied this analysis to identify the overlap between causal mQTLs and GWAS-significant loci, not only GWAS-significant variants. Also, it would be nice if they also report the overlap with significant GWAS variants.

I think the manuscript is improved in the current version. I suggest the authors to also discuss their results in the context of the findings from recent placental TWAS/SMR studies that have been published on this journal (PMID: 35121757 and PMID: 37188697).

Reviewer #2 (Remarks to the Author):

I have read the revised manuscript carefully, and looked at the authors' rebuttal letter as well.

I see that the authors have responded substantively to some, though not all, of the requests by the other reviewers, but not to mine.

So, their conclusions remain based on their careful analysis of descriptive data, still without direct experimental support.

The experiment using CRISPR in a human cell line to test their functional predictions for one locus, as I had recommended, is really not hard to do.

The strengths of the study that I originally noted (as did the other reviewers) remain.

Perhaps importantly, it appears from the authors' cover letter that they would be willing to try harder to address the reviewers' requests, so a third revision might be considered.

Reviewer #3 (Remarks to the Author):

I would like to thank the authors for a careful consideration of my comments/suggestions. While I agree with the authors that the manuscript has improved following the suggestions from the reviewers, there are still some aspects of the analysis that I do not agree with.

Major points:

1. Discovery of cis-mQTLs and interaction cis-mQTLs.

The authors note that their main reference for selecting the model for discovery of cis-mQTLs was the manuscript by the GoDMC consortium. The authors point out that the paper by GoDMC also uses similar model with age, sex, predicted cell counts, predicted smoking and genetic PCs as covariates (for unrelated participants). However, the authors have failed to notice that this was the first step in DNAm data adjustment. As stated in the second paragraph of the “DNAm data adjustment” in the GoDMC paper:

“For unrelated participants we regressed out age, sex, predicted cell counts, predicted smoking and genetic PCs (adjustment 1).”... “We took the residuals from the first adjustment forward to regress out the nongenetic DNAm PCs on the adjusted DNAm β values (adjustment 2). The residuals from these analyses were rank transformed and centered to have mean 0 and variance 1.”

After reading the GoDMC paper, my conclusions are that their mQTL discovery pipeline adjusted for nongenetic DNAm PCs and they chose to be conservative by including index SNPs from clumping procedures for all subsequent analyses. Thus, I can’t agree with the following sentence by the authors in the response to the reviewers letter:

“However, they preferred to be conservative and performed the downstream analyses without including the DNAm PCs, in the case of the cohorts with unrelated participants only.”

All in all, a recommended practice in molQTL studies is to incorporate latent factors computed from the normalized molecular data, such as PCs or PEEER factors. As exemplified by the GoDMC and the GTEx consortium, there are different ways how to achieve this. In addition to accounting for known and unknown confounders, this results in increased power due to reduced residual variance to detect molQTL effects as demonstrated in the section “Evaluation of DNAm data adjustment” by the GoDMC paper that the authors also refer to.

I thank the authors for the additional analyses performed. However, in my opinion, the discovery of (interaction) molQTL should include latent factors (note that 5 PCs may not be enough). This is the standard practice in the field now, which is also followed by large consortiums, such as GTEx, GoDMC, and eQTLGen.

I agree with the authors that the use of DNAm beta-values in different approaches are widely accepted. However, I would like to correct the authors by pointing out that the GoDMC project chose to model inverse normalized values in the molQTL model. This is clearly outlined in their WIKI "... and to reduce the possibility of type 1 errors we will be transforming the methylation values using inverse normal transformation" (https://github.com/MRCIEU/godmc/wiki/Process-methylation-data). Inverse normal transformation also ensures that the assumptions of the linear regression model (e.g., Gaussian residuals and homoscedasticity) are met.

2. Multiple testing.

I would like to clarify that a multiple testing procedure for molQTL mapping should account for multiple testing due to multiple variants per molecular trait and multiple testing due to multiple molecular traits across the genome. This is typically achieved by permutations (i.e., the permuted cis-mQTL approach as the authors call this). Then at a set FDR threshold one could estimate genome-wide significant molQTL features (e.g., eGenes in eQTL analysis and mSites or mProbes in mQTL analysis).

The authors point out that:

"However, we performed FDR correction of our mQTLs as suggested by the reviewer. With the FDR<0.05 correction, out of the 1,132,596,401 possible CpG-SNP combinations tested in 0.5 Mb windows in the whole genome, 27,636,581 mQTL associations appeared to be significant, compared to the 7,921,914 that were suggestive of genome-wide significance under a nominal p-value of 5e-8. We humbly think that this is a huge number, even difficult to manage in any characterization tool, and that probably is hiding a number of false positive associations. We are obviously open to further discuss this issue with the reviewer but, in the present revision, we propose to keep our suggestive threshold."

Indeed 27,636,581 mQTL associations is a huge number. But more important question is how many mSites were discovered at FDR < 0.05? There's an increase in the total number of mQTL associations due to the fact that each mSite has it's own significance level for SNPs. There's no

need to characterize all the significant mQTLs, it is sufficient to focus on the mSites (with FDR < 0.05) together with their top associated mQTL variants.

I thank the authors for clarifying their choice for choosing to focus on mQTLs with P_nominal < 5e8. After checking the paper by Zhu et al. that describes the SMR method, I noticed the following two sentences:

“We included in the analysis only probes for which the P value of the top associated cis-eQTL was <5 × 10−8. This is because one of the basic assumptions for an MR analysis is that the instrumental variable has a strong effect on exposure, that is, the SNP is strongly associated with gene expression.”

Thus, the authors of SMR recommend to include *probes* with the top associated eQTL p < 5e-8. This is not the same as including all QTLs below this threshold as the authors interpret this (note the focus on probes vs eQTLs). Moreover, the authors of SMR use this threshold as an indication of a strong association between a SNP and an expression probe and not a “suggestive threshold of genome-wide significance” as the authors interpret this.

All in all, I suggest the authors to focus in the main text on the permuted mQTL database. Later the SMR analysis could be applied to mSites with FDR < 0.05 and with the top associated cis-mQTL (nominal) p < 5e-8. Nominal mQTL associations are valuable in downstream analysis, such as SMR and coloc, but multiple testing on two levels is needed for discovery of significant mQTLs.

3. Sharing between iQTLs.

I got the impression that the authors are not familiar how the Storey’s pi1 value could be used to estimate sharing among iQTLs. Briefly, one takes the top mQTL variant per mSite from one set (e.g., test set) and estimates the association p-value of the given CpG sites and variant in the other set (e.g., replication set). Then one can use the q-value package to estimate p1 (the proportion of true positives) using the p-values corresponding to the CpG site-variant pairs from the test set calculated based on the replication set.

Minor points.

1. Figure 1a – I'm confused by the edits. Is the median distance mentioned in the main text calculated based on the raw distances or absolute distance? I think there's no point in highlighting the raw median distance between SNPs and CpG sites, because the distribution of the distance is two-sided as the authors also note.

2. Definition of positive and negative imQTLs. As introduced by Kim-Hellmuth et al., the direction of effect of iQTL shows whether the QTL effect increases (positive) or decreases (negative) as a function of cell type proportion. How did the authors classify their imQTLs to be positive or negative?

REVIEWER COMMENTS

General comment: *First of all, we would like to thank the reviewers and the editor for their suggestions and help. We believe to have updated the manuscript with all the methodological improvements suggested by the reviewers, very especially including the (i)mQTL mapping from rank-based inverse normal transformed (RNT) DNAm values with 18 DNAm PCs from residualized DNAm data as covariates to avoid multicollinearity (following the Min et al 2021 - GoDMC pipeline). Additionally, we have performed two-level multiple testing and focused on the permuted database. These modifications have led to minor changes in some sections and more deep changes in others, but we can now say that the main message as well as many of the main hits of the previous versions remain. In this context, not all the changes have been literally copied here, but only those that directly respond to the reviewers' comments or are crucial to understand our decisions. The rest of modifications have been tracked in the manuscript. Finally, please note that most of the figures have been modified and, although RNT values have been consistently used across the manuscript, beta-values are still depicted for an easier biological interpretation. We believe that the manuscript is ready for publication and sincerely hope that the reviewers will enjoy reading this new version.*

Reviewer #1 (Remarks to the Author):

I think the manuscript is improved and the authors have performed most of the required work to address the reviewers' comments.

Response: *We very sincerely thank the reviewer for his/her constructive comments and help in this and the previous revision. We hope he/she will enjoy reading the new version of our manuscript and will find our work ready to be published this time.*

I have just few remaining comments and suggestions:

1. I appreciate that the authors have performed further analyses to test the relationship between GWAS sample size, GWAS hits, and SMR hits. However, I do not agree about their interpretation of the results, and some details of these analyses are not clear. First, the authors comment that there is no correlation about GWAS sample size and SMR hits, although when they test this relationship the r is 0.52. This is indicative of an almost moderate correlation. The p-value is affected by sample size that in this analysis is 10, so it's recommendable to not base any conclusion of the fact that a p-value of 0.13 is higher than the 0.05 threshold. It is evident in figure supplementary 13 a and b that the 3 diseases with the bigger GWAS sample size, and higher number of GWAS hits, have also the higher number of SMR hits.

2. Also, from supplementary figure 13b it seems that there are ~10,000 GWAS hits for depression and ~2,500 GWAS hits for bipolar disorder, however bipolar GWAS identified 64 loci (ref.35) and depression GWAS identified 102 loci (ref.37). It's important that the authors double check everything to be sure that they are reporting correct information. Indeed, the number of GWAS hits reported in Table 2 does not seem right, at least for ADHD, autism, bipolar disorder, and depression.

Response: *We are responding to the first two points at the same time. We appreciate these observations from the reviewer because in fact, he/she is very right. We really apologize for any misunderstanding (e.g. we had reported the number of significant SNPs, sometimes under a suggestive cutoff, instead of GWAS significant loci, as noted by the reviewer, and correcting this*

has impacted our results). We have changed Table 2, Supplementary Figure 13 (now new Supplementary Figure 9) and modified the manuscript accordingly:

"However, we wanted to make sure that our SMR results were not merely led by statistical power. Therefore, on the one hand, we calculated the correlation between the SMR hits and i) the GWAS sample size, ii) the total number of SNPs per GWAS, and iii) the number of GWAS-significant loci, in all the neuropsychiatric disorders studied. We observed a correlation of r = 0.48 between the number of significant SMR results and the GWAS sample size, although it did not reach statistical significance, probably due to the limited number of SMR experiments performed (Table 2, Supplementary Figure 9A). On the other hand, even though the total number of SNPs included in each of the GWAS was not correlated at all with the results obtained in SMR, there was a highly significant correlation between the latter and the number of significant loci reported in each GWAS ($P < 1.7x10^{-6}$, r = 0.97) (Table 2, Supplementary Figure 9B and C). This is intuitively logical since the higher the number of genomic regions with disease-associated SNPs, the more likely it will be to find an overlap with SNPs associated with placental DNAm or with anything else. In the specific case of SCZ, we ran the same SMR approach with two previous, smaller GWAS and still obtained significant SMR hits, with a remarkable overlap among the three studies. The characteristics of these studies are summarized in Supplementary Figure 9D, together with the Venn diagrams showing the overlapping results (Supplementary Figure 9E and F). In summary, our results reveal an at least moderate correlation between the number of SMR hits and the GWAS sample size, that is surpassed by far by the number of GWAS significant loci, with a vast effect on the capacity of SMR to find potentially interesting results."

3. I suggest to add labels withing the plots of supplementary figure 13, so that each disease can be clearly identified.

Response: *We tried to implement this suggestion by the reviewer which in fact is a good idea but unfortunately the dots close to 0 in the X and Y axes are too close to one another to be adequately labeled. We hope the reviewer understands our decision.*

4. I appreciate the explanation provided by the author about the colocalization analyses. In the response they say that "we applied this analysis to identify the overlap between causal mQTLs and GWAS-significant loci", however in the main text they write that "The Bayesian colocalization analysis implemented in the coloc R package focuses on finding the intersection between significant variants independently associated with two phenotypes, comparing the association patterns in the GWAS and QTL analyses across genomic regions, and thus combining the summary statistics into posterior probabilities for five hypotheses (see Methods for more information)." So, I think they should write in the main text they applied this analysis to identify the overlap between causal mQTLs and GWAS-significant loci, not only GWAS-significant variants. Also, it would be nice if they also report the overlap with significant GWAS variants.

Response: *We have corrected this in the main text as suggested by the reviewer:*

"The Bayesian colocalization analysis implemented in the coloc R package[52] focuses on finding the intersection between significant loci independently associated with two phenotypes, comparing the association patterns in the GWAS and QTL analyses across genomic regions, and thus combining the summary statistics into posterior probabilities for five hypotheses (see Methods for more information)."

Since we perform colocalization between mQTLs and GWAS-significant loci (not variants), we have specified it, and therefore standardized this explanation.

I think the manuscript is improved in the current version. I suggest the authors to also discuss their results in the context of the findings from recent placental TWAS/SMR studies that have been published on this journal (PMID: 35121757 and PMID: 37188697).

Response: *We thank the reviewer for bringing these two interesting papers to our attention. We had already read them but this was after writing the article in its first version. However, they are remarkable literature in the field that we agree deserve to be mentioned.*

Regarding the first reference (PMID: 35121757), we have mentioned this work in the Introduction:

"Additionally, fetal genetics-mediated placental expression has been reported to have large effects on early-life traits, that may persist later-in-life as etiologic antecedents for complex traits[5]. Finally, the placenta constitutes the interface between mother and child during pregnancy and is the maximum regulator of the prenatal milieu, having a key role in the growth and the neurodevelopment of the fetus[6]."

And in the Bibliography:

"5. Bhattacharya, A. et al. Placental genomics mediates genetic associations with complex health traits and disease. Nat Commun. 13, 706 (2022)."

Regarding the second reference (PMID: 37188697), we have mentioned this work in the Discussion:

"Moreover, it is known that congenital HCMV infection can cause long-term clinical manifestations such as hearing loss, neurodevelopmental disorders, ophthalmic complications, and ASD, among others[83]. The virus moves from mother to fetus by infecting TB in the placenta, altering the development and integrity of the organ, and eventually, inhibiting placental cell differentiation, self-renewal, and migration. Last but not least, in a very recent transcriptome-wide association study (TWAS) by Ursini and colleagues, the placental genes and transcripts associated with SCZ risk were shown to be enriched for pathways related with the activation of the pathogenesis of both Coronavirus and influenza, further reinforcing the idea of the implication of viral infections[84]."

And in the Bibliography:

"84. Ursini, G., et al. Prioritization of potential causative genes for schizophrenia in placenta. Nat. Commun. 14, 2613 (2023)."

Reviewer #2 (Remarks to the Author):

I have read the revised manuscript carefully, and looked at the authors' rebuttal letter as well.

I see that the authors have responded substantively to some, though not all, of the requests by the other reviewers, but not to mine.

So, their conclusions remain based on their careful analysis of descriptive data, still without direct experimental support.

The experiment using CRISPR in a human cell line to test their functional predictions for one locus, as I had recommended, is really not hard to do.

The strengths of the study that I originally noted (as did the other reviewers) remain.

Perhaps importantly, it appears from the authors' cover letter that they would be willing to try harder to address the reviewers' requests, so a third revision might be considered.

Response: *We thank the reviewer for his/her comments and suggestions. Unfortunately, we insist that we have not the expertise nor the material means to carry out the in vitro experiments requested. In the previous revision, we had already added an explanation of how this could be a limitation of the study that additionally, requires further investigation. We have extended and highlighted this idea in the limitations and conclusions of the present version as follows:*

"Finally, at this point we lack direct experimental support of our results. Further research, especially regarding in vitro validation of the most likely causal candidates, would enormously help supplement our findings."

And:

"Future work, including functional approaches such as in vitro modulation of expression in potential effector cell lines, as well as long term follow up in prospective studies, will be needed in order to better characterize our findings and their implication in disease development."

Lastly, as the reviewer will notice, we have also made deep changes following the suggestions of the other reviewers. We very sincerely hope that the reviewer will appreciate this effort and enjoy reading this new version.

Reviewer #3 (Remarks to the Author):

I would like to thank the authors for a careful consideration of my comments/suggestions. While I agree with the authors that the manuscript has improved following the suggestions from the reviewers, there are still some aspects of the analysis that I do not agree with.

Response: *We very much thank the reviewer for his/her thorough revision of our work, and sincerely hope that he/she will find our work ready to be published this time.*

Major points:

1. Discovery of cis-mQTLs and interaction cis-mQTLs.

The authors note that their main reference for selecting the model for discovery of cis-mQTLs was the manuscript by the GoDMC consortium. The authors point out that the paper by GoDMC also uses similar model with age, sex, predicted cell counts, predicted smoking and genetic PCs as covariates (for unrelated participants). However, the authors have failed to notice that this was the first step in DNAm data adjustment. As stated in the second paragraph of the "DNAm data adjustment" in the GoDMC paper:

"For unrelated participants we regressed out age, sex, predicted cell counts, predicted smoking and genetic PCs (adjustment 1)."... "We took the residuals from the first adjustment forward to regress out the nongenetic DNAm PCs on the adjusted DNAm β values (adjustment 2). The residuals from these analyses were rank transformed and centered to have mean 0 and variance 1."

After reading the GoDMC paper, my conclusions are that their mQTL discovery pipeline adjusted for nongenetic DNAm PCs and they chose to be conservative by including index SNPs from clumping procedures for all subsequent analyses. Thus, I can't agree with the following sentence by the authors in the response to the reviewers letter:

"However, they preferred to be conservative and performed the downstream analyses without including the DNAm PCs, in the case of the cohorts with unrelated participants only."

All in all, a recommended practice in molQTL studies is to incorporate latent factors computed from the normalized molecular data, such as PCs or PEER factors. As exemplified by the GoDMC and the GTEx consortium, there are different ways how to achieve this. In addition to accounting for known and unknown confounders, this results in increased power due to reduced residual variance to detect molQTL effects as demonstrated in the section "Evaluation of DNAm data adjustment" by the GoDMC paper that the authors also refer to.

I thank the authors for the additional analyses performed. However, in my opinion, the discovery of (interaction) molQTL should include latent factors (note that 5 PCs may not be enough). This is the standard practice in the field now, which is also followed by large consortiums, such as GTEx, GoDMC, and eQTLGen.

I agree with the authors that the use of DNAm beta-values in different approaches are widely accepted. However, I would like to correct the authors by pointing out that the GoDMC project chose to model inverse normalized values in the molQTL model. This is clearly outlined in their WIKI "... and to reduce the possibility of type 1 errors we will be transforming the methylation values using inverse normal transformation" (https://github.com/MRCIEU/godmc/wiki/Process-

methylation-data). Inverse normal transformation also ensures that the assumptions of the linear regression model (e.g., Gaussian residuals and homoscedasticity) are met.

Response: *We would like to thank Reviewer #3 for this extensive revision of the methodology. We have made different comparisons and have finally decided 1) to use the rank-based inverse normal transformation (RNT) of the DNAm data and 2) to adjust our model by DNAm principal components (PCs), exactly as described by GoDMC (Min et al. 2021). The Methods section has been modified as follows:*

"***Methylation data***

1...]

To correct for the possible outliers, we Winsorized the extreme values to the 1% percentile (0.5% in each side), where percentiles were estimated with the empirical beta-distribution. Finally, the rank-based inverse normal transformation (RNT) was applied on the beta-values, and these estimations were the DNAm values considered for mQTL mapping. The final dataset consisted in 368 samples and 747,486 DNAm probes (CpGs).

1...]

Placental cis-mQTL analysis

A total of 4,171,035 SNPs and 747,486 CpGs from 368 samples with paired genotype and DNAm data were considered for the cis-mQTL analysis in TensorQTL[110]. TensorQTL nominal modality performs linear regressions between the genotype and the normalized DNAm RNT-values, as implemented in FastQTL[111]. The covariates included in the regression model were the sex of the fetuses, the first five PCs derived from the genotype data (genotype PCs), the first 18 non-genetic DNAm PCs, and the cell type proportions claculated with the Planet methylation panel. Genotype and DNAm PCs were included in the model as covariates to remove hidden and/or technical confounders affecting the DNAm data. To avoid multicollinearity between the non-genetic DNAm PCs and the other covariates, the DNAm PCs were calculated on the residuals from a multiple linear regression adjusting the normalized DNAm RNT-values by the known covariates (sex of the fetuses, the first five genotype PCs and the five estimated cell types). Following Min et al. 2021, we kept all DNAm PCs that cumulatively explained 80% of the variance with a maximum number of 20 PCs for subsequent steps[28]. Then we performed a GWAS on each of the DNAm PCs and retained those PCs that were not associated with the genotype at a suggestive threshold ($P > 1 \times 10^{-7}$). This procedure returned 18 non-genetic DNAm PCs."

Regarding Results, in summary, RNT had little impact in this section while adding the 18 DNAm PCs did make a difference at different points. We are not including all of them here, but we have marked them with track changes in the manuscript.

2. Multiple testing.

I would like to clarify that a multiple testing procedure for molQTL mapping should account for multiple testing due to multiple variants per molecular trait and multiple testing due to multiple molecular traits across the genome. This is typically achieved by permutations (i.e., the permuted cis-mQTL approach as the authors call this). Then at a set FDR threshold one could estimate genome-wide significant molQTL features (e.g., eGenes in eQTL analysis and mSites or mProbes in mQTL analysis). The authors point out that:

“However, we performed FDR correction of our mQTLs as suggested by the reviewer. With the FDR<0.05 correction, out of the 1,132,596,401 possible CpG-SNP combinations tested in 0.5 Mb windows in the whole genome, 27,636,581 mQTL associations appeared to be significant, compared to the 7,921,914 that were suggestive of genome-wide significance under a nominal p-value of 5e-8. We humbly think that this is a huge number, even difficult to manage in any characterization tool, and that probably is hiding a number of false positive associations. We are obviously open to further discuss this issue with the reviewer but, in the present revision, we propose to keep our suggestive threshold.”

Indeed 27,636,581 mQTL associations is a huge number. But more important question is how many mSites were discovered at FDR < 0.05? There’s an increase in the total number of mQTL associations due to the fact that each mSite has it’s own significance level for SNPs. There’s no need to characterize all the significant mQTLs, it is sufficient to focus on the mSites (with FDR < 0.05) together with their top associated mQTL variants.

Response: *We understand the two-level correction pointed out by the reviewer and therefore, we have focused only in the permuted database (please see the next response for the specific corrections that we have carried out in the interaction models). As a result, we have performed the following changes in Methods:*

*“**Characterization of placental cis-mQTLs***

As previously mentioned, the characterization of the cis-mQTLs was performed on the permuted database in different steps. First, we plotted the distribution of both the distance between mQTL-CpGs and paired mQTL-SNPs and the mQTLs along the chromosomes, followed by a volcano plot to check the distribution of the negative and positive effect sizes.”

And in Results:

“*The permuted mQTL database was calculated in 368 fetal placenta DNA samples from the Gipuzkoa, Sabadell and València cohorts of the Infancia y Medio Ambiente (INMA) project[26], and contained 214,830 cis-mQTLs (FDR P_{beta} < 0.05) with 214,830 and 150,649 unique CpGs and SNPs, respectively. Briefly, placental DNAm was modeled as a function of genotype in ±0.5 Mb windows, with fetal sex, 5 genotype principal components (PC), 18 DNAm PCs, and Planet-estimated cell types[27] as covariates (See Methods section for more details). Phenotypic information of the donor mothers is summarized in Table 1. In the vast majority of the cis-mQTLs, the SNP and the CpG were located relatively close to each other, with an absolute median distance of 13,87 kb, indicating that genetically modulated DNAm is typically close to the implicated regulatory variant, as previously described by other authors[28] (Figure 1A). The distribution of mQTLs across chromosomes tended to be in line with chromosome length (Figure 1B). Additionally, the Haplotype Reference Consortium (HRC)[29] reference alleles in placental cis-mQTLs presented both positive and negative effect directions (Figure 1C), as expected.*

In general, mQTL-CpGs were depleted from CpG islands and promoters ($P < 2.2x10^{-16}$), and enriched within gene bodies ($P < 2.2x10^{-16}$) and genomic features with intermediate methylation values such as open sea, and CpG island shelf and shore regions ($P < 2.2x10^{-16}$) (Figure 1D, E and F, and Supplementary Data 1). They were also enriched in the HLA region with 2,896 CpGs out of the total 214,830 mQTL-CpGs of the permuted database ($P = 1.45x10^{-10}$). Using eFORGE[30,31,32], we were able to detect an enrichment in DNase I hotspots and H3K4me1 broadPeaks in different tissues, fetal placenta being the one showing the strongest signal (Figure

1G, Supplementary Figure 1 and Supplementary Data 2). DNase I hotspots mark accessible chromatin while H3K4me1 broadPeaks appear in enhancer regions. Afterwards, we annotated the mQTL-CpGs to their closest genes, performed a gene-set enrichment analysis with the Disease Ontology database[33], and obtained 219 enriched gene sets, including neuropathy (Benjamini-Hochberg P = 1.88x10^{-9}), mood disorder (Benjamini-Hochberg P = 1.22x10^{-7}), demyelinating disease (Benjamini-Hochberg P = 2.48x10^{-6}), and BIP (Benjamini-Hochberg P = 6.96x10^{-6}) (Supplementary Data 3 and Supplementary Figure 2).”

Finally, we have remodeled Figure 1 and erased the (Supplementary) Figures in which the characterization of the rest of the databases were shown. We have modified our online database accordingly and thus, only permuted and conditional databases appear in the browser, although we do provide a link to download the nominal database filtered by a nominal P-value < 5x10^{-8}.

I thank the authors for clarifying their choice for choosing to focus on mQTLs with P_nominal < 5e-8. After checking the paper by Zhu et al. that describes the SMR method, I noticed the following two sentences:

“We included in the analysis only probes for which the P value of the top associated cis-eQTL was <5 × 10−8. This is because one of the basic assumptions for an MR analysis is that the instrumental variable has a strong effect on exposure, that is, the SNP is strongly associated with gene expression.”

Thus, the authors of SMR recommend to include *probes* with the top associated eQTL p < 5e8. This is not the same as including all QTLs below this threshold as the authors interpret this (note the focus on probes vs eQTLs). Moreover, the authors of SMR use this threshold as an indication of a strong association between a SNP and an expression probe and not a “suggestive threshold of genome-wide significance” as the authors interpret this.

All in all, I suggest the authors to focus in the main text on the permuted mQTL database. Later the SMR analysis could be applied to mSites with FDR < 0.05 and with the top associated cis-mQTL (nominal) p < 5e-8. Nominal mQTL associations are valuable in downstream analysis, such as SMR and coloc, but multiple testing on two levels is needed for discovery of significant mQTLs.

Response: *We understand the concerns raised by the reviewer since the sentence is quite confusing. Given that we have published previous work using SMR with the databases available in the SMR webpage, and we have never had problems when stablishing the nominal P-value cutoff of 5x10^{-8}, we decided to write to Prof. Jian Yang, the corresponding author of the paper in which the method was described. Please see below our specific question (in which we confront our approach with the suggestion of the reviewer) and the answer of Prof. Yang:*

From: ariadna cilleros portet
Date: Thursday, 7 September 2023 at 21:03
To: Jian Yang
Subject: Doubt: P-value threshold SMR tool

Dear Dr. Yang,

My name is Ariadna Cilleros, and I am a PhD student at the University of the Basque Country. I am currently using the SMR software to perform Multi-SNP-based Mendelian Randomization analysis with an mQTL database. This mQTL database contains all the possible associations between CpGs and SNPs with a nominal p-value < 5e-8. Therefore, all the probes (or DNA methylation sites) have at least one associated mQTL p-value < 5e-8. We wondered whether this approach is correct.
At the moment, we are under review and one of the reviewers have suggested the following approach instead: "I suggest the authors to focus in the main text on the permuted mQTL database. Later the SMR analysis could be applied to mSites with FDR < 0.05 and with the top associated cis-mQTL (nominal) p < 5e-8. Nominal mQTL associations are valuable in downstream analysis, such as SMR and coloc, but multiple testing on two levels is needed for discovery of significant mQTLs."
We would be grateful if you could provide us with some advice on this matter.
Thank you in advance!

Yours sincerely,

De: Jian YANG 杨剑
Enviat el: diumenge, 10 de setembre de 2023 11:13
Per a: ariadna cilleros portet
Tema: Re: Doubt: P-value threshold SMR tool

Dear Ariadna,

The threshold of 5e-8 is very likely to be more stringent than the threshold obtained from permutation. For an SMR analysis, I would always suggest that users stick to the p-value threshold of 5e-8.

Best regards,
Jian

We have corroborated that this is true and in fact, the probes with a permuted FDR<0.05 are more abundant, and all the CpGs under the nominal P-value of $5x10^{-8}$ are included among them. Therefore, we are keeping the threshold suggested by the developer for all SMR approaches. We hope that the reviewer will understand this decision.

3. Sharing between iQTLs.

I got the impression that the authors are not familiar how the Storey's pi1 value could be used to estimate sharing among iQTLs. Briefly, one takes the top mQTL variant per mSite from one set (e.g., test set) and estimates the association p-value of the given CpG sites and variant in the other set (e.g., replication set). Then one can use the q-value package to estimate p1 (the proportion of true positives) using the p-values corresponding to the CpG site-variant pairs from the test set calculated based on the replication set.

Response: *We sincerely thank the reviewer for his/her explanation of this method with which we were not familiar. We hope to have correctly understood and implemented it this time.*

We have changed the Methods section as follows:

"Placental cis-imQTL analysis

The cell type-imQTLs were computed with the interaction modality in TensorQTL, consisting in nominal associations for a linear model that includes a genotype per interaction term[110]. Additionally, we used the run_eigenmt=True option, to compute eigenMT-adjusted P-values. This method runs faster than permutations, calculates adjusted P-values that closely approximate empirical ones[35] and thus, has been implemented by TensorQTL as an alternative to permutation

in imQTL mapping. The same SNPs, CpGs and samples from the permuted database (without interaction terms) were considered for the interaction analyses. The covariates used were the sex of the fetuses, the first five genotype PCs and 18 DNAm PCs. Following the recommendations by Kim-Hellmuth et al. 2021[25], estimations of STBs and TBs per sample were defined separately as the interaction terms. We considered significant only those imQTLs that showed an eigenMT FDR < 0.05.

We estimated sharing between STB- and TB- imQTLs with Storey's p1 method[36]. Briefly, we took the eigenMT FDR < 0.05 STB-imQTLs and retrieved the nominal P-values of these SNP-CpG pairs from the TB-imQTL set. Then we used the qvalue package[123] to estimate p1 (the proportion of true positives). As STB and TB proportions are negatively correlated in term placentas, we wanted to know whether the interaction effect sizes of the overlapping imQTLs were also negatively correlated. Therefore, we took the same list of eigenMT FDR < 0.05 STB-imQTLs and recovered all the imQTLs associated to these CpG sites with $P_{nominal} \times N_{eff}$ (effective number of independent variants in the cis window estimated by eigenMT) < 0.05[37]. Then we retrieved these SNP-CpG pairs from the TB-imQTL set only if their $P_{nominal} \times N_{eff} < 0.05$ and drew a Venn diagram. Finally, we calculated the correlation of the interaction effect sizes of the overlapping imQTLs.

Additionally, both GA and sex were also considered as interaction terms. In the first case, the covariates included were the sex of the fetuses, the first five genotype PCs, the 18 DNAm PCs and the cell type proportions from Planet. In the second case, the covariates included were the first five genotype PCs, the 18 DNAm PCs and the cell type proportions from Planet. For sharing calculations, the same methods as with TB were used for GA."

And added the following reference to Bibliography:

"35. Davis, J.R., Fresard, L., Knowles, D.A., Pala, M., Bustamante, C.D., Battle, A., Montgomery, S.B. An Efficient Multiple-Testing Adjustment for eQTL Studies that Accounts for Linkage Disequilibrium between Variants. Am. J. Hum. Genet. 98, 216-224 (2016).

36. Storey, J.D., Tibshirani, R. Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences of the United States of America 100, 9440–9445 (2003).

37. Garrido-Martín, D. et al. A fast non-parametric test of association for multiple traits. Genome Biol. 24, 230 (2023)."

Additionally, the results have changed considerably, being this section the most affected as a consequence of the reviewer's suggestions (probably due to the fact that employing both interaction and eigenMT correction at the same time is a quite restrictive approach). In conclusion, we have modified the Results section accordingly and tracked the changes but we are not including them in this letter.

Minor points.

1. Figure 1a – I'm confused by the edits. Is the median distance mentioned in the main text calculated based on the raw distances or absolute distance? I think there's no point in highlighting the raw median distance between SNPs and CpG sites, because the distribution of the distance is two-sided as the authors also note.

Response: *We have taken into account Reviewer #3 comment and we have only highlighted the absolute median distance as follows:*

"In the vast majority of the cis-mQTLs, the SNP and the CpG were located relatively close to each other, with an absolute median distance of 13,87 kb, indicating that genetically modulated DNAm is typically close to the implicated regulatory variant, as previously described by other authors[28] (Figure 1A)."

2. Definition of positive and negative imQTLs. As introduced by Kim-Hellmuth et al., the direction of effect of iQTL shows whether the QTL effect increases (positive) or decreases (negative) as a function of cell type proportion. How did the authors classify their imQTLs to be positive or negative?

Response: *As we state in the Methods section "Placental cis-imQTL analysis", genotype main effects at low (<25th percentile) vs high (>75th percentile) cell type proportion was compared. imQTLs with positive cell type correlation show an increase of the genotype main effect from low to high cell proportions ($1i_{mQTL}$ low < $1i_{mQTL}$ high). imQTLs with negative cell type correlations show a decrease ($1i_{mQTL}$ low > $1i_{mQTL}$ high) and the uncertain group contain imQTLs for which the sign flipped between low and high cell type proportions ($1i_{mQTL}$ low ≠ $1i_{mQTL}$ high). This classification was defined in the Supplementary Methods from Kim-Hellmuth et al. 2020. We have moved this paragraph to the Multi-SNP-based Mendelian Randomization section in Methods, since in the current version of the manuscript, the classification is performed solely with the aim of retrieving potentially cell type-specific mQTLs for SMR (that is, positive imQTLs) rather than those that more likely reflect interactions with another cell population that is decreasing (that is, negative imQTLs), according to Kim-Hellmuth.*

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

I think the manuscript is improved and suitable for publication. I have just one minor suggestions:

Table 2 clearly shows that schizophrenia has the highest number of SMR hits and the highest ratio between SMR hits and GWAS significant loci. This may suggest that the placental methylome is particularly relevant for schizophrenia, also in comparison with other disorders, and it is consistent with previous findings (PMID: 37188697). The authors may consider mentioning this in the discussion.

Reviewer #1 (Remarks on code availability):

I think the code is well written, and thins important for reproducibility.

Reviewer #3 (Remarks to the Author):

I would like to thank the authors for applying state-of-the-art methodology in their study by including DNAm PCs as covariates in (i)mQTL mapping and performing two-level multiple testing to discover CpG sites with an mQTL. I have some additional questions and suggestions regarding the analysis presented in the revised manuscript.

Major comments

1. Terminology used in molQTL studies. The authors use mQTL-CpGs and mQTL-SNPs in the text. It might be desirable to use the proper molQTL terminology to make the text easier to understand for a reader.

For example, mQTL is a genetic loci that affects the methylation levels of a CpG site. CpG sites with at least one genome-wide significant mQTL at a chosen FDR threshold are called mSites or mProbes and the corresponding significant variants are called mVariants (see Aguet et al, 2023, Nature Reviews Methods Primers). Importantly, the discovery of mQTL focuses on the number of

significant CpG sites (and (conditionally) independent mVariants). Usually, the variant with the lowest association p-value is called the lead mVariant (this is the variant reported in the permuted database).

As a suggestion, the first sentence of the section “Placental cis-mQTL characterization” could be rephrased as: “We discovered cis-mQTLs for 214,830 CpG sites (mSites) with FDR < 0.05 (the permuted mQTL database) in 368 fetal placenta DNA samples from the Gipuzkoa, Sabadell and València cohorts of the Infancia y Medio Ambiente (INMA) project26.”

2. Sharing of the results. The authors note that only permuted and conditional databases appear in their browser, and they provide a link to download the nominal database filtered by a nominal p-value < 5e-8. If possible, it would be great if the authors could share a link to download the full nominal database without a p-value filter. This will ensure that the results of this study can be used in downstream analyses, such as colocalization or pi1 analyses by other groups.

3. P-value threshold in the SMR analysis. I just wanted to clarify the main message of my previous request regarding the choice of p-value threshold in the SMR analysis. In molQTL studies, the first task is to find the molecular traits (e.g., genes or CpG sites) that have a significant mQTL effects at a chosen FDR threshold, which are called eGenes and mSites/mProbes for eQTL or mQTL studies, respectively. This is achieved by applying permutations (based on the permutation database as the authors call it). The next task may involve using eGenes and mSites in SMR or colocalization analysis.

Thus, my main message was to focus on mSites throughout the manuscript (i.e., the permuted database). Given the type of downstream analysis (pi1 analysis, SMR, colocalization), one may need to use the summary statistics based on nominal database to include all variants or all associated variants at a certain p-value threshold. Now reading the manuscript I can understand the authors have set the focus on GWAS loci and they try to use mQTLs to give mechanistic insights of GWAS SNPs. Could the authors at least report the number of mSites with FDR < 0.05 spanning the GWAS regions tested in the SMR/colocalization analysis? This would make a connection between the SMR/colocalization analysis and the permuted database of mQTLs described before.

While the suggested threshold of nominal p-value of 5e-8 is more stringent than the permutation-based threshold for mQTLs, it is likely not true for imQTLs. Please include also the FDR of the imQTLs included to SMR analysis for an assessment of the significance of the imQTL effect.

4. Sharing between imQTLs using the Storey’s pi1 method. Please correct the typo in Storey’s pi1 method. It should be pi1 (not p1), where pi stands for the symbol $'^{1}$pi$'^{1}$. Also, please plot the p-value

histograms corresponding to the significance of the genotype*TB and genotype*GA effect of the STB-imQTLs as supplementary figures. These figures give additional visual support of sharing between imQTLs. In my opinion Figure 2C is misleading. The figure legend says that the intersection between STB- and TB-imQTLs is represented as Venn diagrams in the figure, but the set of CpG site-variant pairs include SNPs in LD (if I have understood the analysis correctly) and statistically nonsignificant TB-imQTLs. If the authors wish to compare the direction of genotype main effect, then they should at least focus on independent SNPs per CpG site.

5. Colocalization analysis. Colocalization analysis is performed using all the genetic variants in the specified region (including also non-significant ones for a given molecular phenotype or trait of interest). Looking at the Supplementary Data 6, the number of SNPs analyzed (column nsps) is very low for some tested loci, e.g., nsps = 4 for region 98357153-96396153, where the lead BIP GWAS SNP is rs4619651 and CpG site is cg19932515. Could the authors please clarify how the colocalization analysis was performed?

Minor comments.

1. Figure 1a legend: I'm still confused by the Figure 1a. The legend says that x-axis represents the absolute distance in Mb, but the red line represents the raw median distance. This does not make sense. Is the red line representing the median absolute distance? It should be noted that the distance is calculated between the lead mVariant and mSite (see the comment about the terminology used in molQTL studies).

2. absolute median distance – would it be more correct to say "median absolute distanace"?

3. Thank you for clarifying the definition of positive and negative imQTLs. The following sentence in the main text needs modification: "In summary, positive imQTLs show increased DNAm related to the main genotype as a function of cell proportion, while negative imQTLs show a decrease." As noted in the Methods section, imQTLs with positive direction of effect (positively correlated with the cell type proportion) showed an increase in the main genotype effect from low to high cell proportions, i.e., the mQTL effect size is increasing with the cell type abundance.

4. Please include p-values of the (i)mQTL effect to Figure 2D-F (and to other similar plots).

5. The authors note that "Out of the 28 and 214 SMR-significant DNAm sites in MDD and SCZ, we found that 3 and 43 correlated with the expression levels of 4 and 26 significant eQTM-genes (FDR < 0.05), respectively, strongly supporting the placental function of the CpGs identified

(Supplementary Data 7 and Supplementary Figure 3).” – does 3 out 28 and 26 out 214 SMR-significant DNAm sites that are also correlated with gene expression levels indicate strong support of the CpG sites?

6. Could the authors include the locuszoom plot of mQTLs for cg10318063 on Figure 4 (to visually represent the colocalization between GWAS of MDD and mQTLs for cg10318063)? Of note, were there a cell type imQTL for cg10318063?

7. The locuszoom plot on Figure 5A looks odd. Why is the association peak not in the center of the region tested? Please also include a locuszoom plot of mQTLs for cg11165035 to Figure 5.

8. Please also include a locuszoom plot of cell type imQTLs to Figure 6 and Supplementary Figure 7.

9. Why were only the top 10,000 mQTL-CpGs tested for enrichment and depletion of overlap with tissue-specific and cell-type-specific regulatory features? This is not mentioned in the main text.

REVIEWER COMMENTS

We want to thank the editor and the reviewers for their useful comments and advice. We believe to have improved the manuscript considerably and hope it is ready for publication this time. We would like to highlight that one of the reviewers have detected an error in the colocalization analysis that we have successfully corrected without major modifications of the results. We very sincerely apologize for this mistake.

Reviewer #1 (Remarks to the Author):

I think the manuscript is improved and suitable for publication. I have just one minor suggestions: Table 2 clearly shows that schizophrenia has the highest number of SMR hits and the highest ratio between SMR hits and GWAS significant loci. This may suggest that the placental methylome is particularly relevant for schizophrenia, also in comparison with other disorders, and it is consistent with previous findings (PMID: 37188697). The authors may consider mentioning this in the discussion.

Reviewer #1 (Remarks on code availability):

I think the code is well written, and thins important for reproducibility.

Response: We would like to thank the reviewer for his/her comments and suggestion. Indeed, it is a good point to highlight that, even though the strong correlation between the number of SMR results and the number of GWAS significant loci, SCZ is by far the disorder showing the highest ratio between SMR hits and GWAS loci. We have added this consideration to the *Discussion* as follows:

"On the other hand, even though the total number of SNPs included in each of the GWAS was not correlated at all with the results obtained in SMR, there was a highly significant correlation between the latter and the number of significant loci reported in each GWAS ($P < 1.7x10^{-6}$, $r = 0.97$) (Table 2, Supplementary Figure 9B and C). This is intuitively logical since the higher the number of genomic regions with disease-associated SNPs, the more likely it will be to find an overlap with SNPs associated with placental DNAm or with anything else. Anyway , it is worthy of mention that SCZ is by far the disorder showing the highest ratio between SMR hits and GWAS loci, suggesting that placental methylome may be particularly relevant for this disorder. Additionally, we ran the same SMR approach with two previous, smaller SCZ GWAS and still obtained significant SMR

***hits, with a remarkable overlap among the three studies. The characteristics of these studies are summarized in Supplementary Figure 9D, together with the Venn diagrams showing the overlapping results (Supplementary Figure 9E and F). In summary, our results reveal an at least moderate correlation between the number of SMR hits and the GWAS sample size, that is surpassed by far by the number of GWAS significant loci, with a vast effect on the capacity of SMR to find potentially interesting results. Nevertheless , the abundant results in SCZ are not fully explained by this effect"*.**

Reviewer #3 (Remarks to the Author):

I would like to thank the authors for applying state-of-the-art methodology in their study by including DNAm PCs as covariates in (i)mQTL mapping and performing two-level multiple testing to discover CpG sites with an mQTL. I have some additional questions and suggestions regarding the analysis presented in the revised manuscript.

Major comments

1. Terminology used in molQTL studies. The authors use mQTL-CpGs and mQTL-SNPs in the text. It might be desirable to use the proper molQTL terminology to make the text easier to understand for a reader.

For example, mQTL is a genetic loci that affects the methylation levels of a CpG site. CpG sites with at least one genome-wide significant mQTL at a chosen FDR threshold are called mSites or mProbes and the corresponding significant variants are called mVariants (see Aguet et al, 2023, Nature Reviews Methods Primers). Importantly, the discovery of mQTL focuses on the number of significant CpG sites (and (conditionally) independent mVariants). Usually, the variant with the lowest association p-value is called the lead mVariant (this is the variant reported in the permuted database).

As a suggestion, the first sentence of the section “Placental cis-mQTL characterization” could be rephrased as: “We discovered cis-mQTLs for 214,830 CpG sites (mSites) with FDR < 0.05 (the permuted mQTL database) in 368 fetal placenta DNA samples from the Gipuzkoa, Sabadell and ValeÌncia cohorts of the Infancia y Medio Ambiente (INMA) project26.”

Response: We thank the reviewer for his/her comment. We have updated the terminology across the manuscript following the reviewer’s advice.

2. Sharing of the results. The authors note that only permuted and conditional databases appear in their browser, and they provide a link to download the nominal database filtered by a nominal p-value < 5e-8. If possible, it would be great if the authors could share a link to download the full nominal database without a p-value filter. This will ensure that the results of this study can be used in downstream analyses, such as colocalization or pi1 analyses by other groups.

Response: We thank the reviewer for his/her comment. We have added a link that enables the download of the unfiltered results (the full TensorQTL output). Please have a look at the website in its final version: irlab.shinyapps.io/shiny_mqtl_placenta/. We honestly think that it has notably improved and that it will be a valuable resource.

3. P-value threshold in the SMR analysis. I just wanted to clarify the main message of my previous request regarding the choice of p-value threshold in the SMR analysis. In molQTL studies, the first task is to find the molecular traits (e.g., genes or CpG sites) that have a significant mQTL effects at a chosen FDR threshold, which are called eGenes and mSites/mProbes for eQTL or mQTL studies, respectively. This is achieved by applying permutations (based on the permutation database as the authors call it). The next task may involve using eGenes and mSites in SMR or colocalization analysis.

Thus, my main message was to focus on mSites throughout the manuscript (i.e., the permuted database). Given the type of downstream analysis (pi1 analysis, SMR, colocalization), one may need to use the summary statistics based on nominal database to include all variants or all associated variants at a certain p-value threshold. Now reading the manuscript I can understand the authors have set the focus on GWAS loci and they try to use mQTLs to give mechanistic insights of GWAS SNPs. Could the authors at least report the number of mSites with FDR < 0.05 spanning the GWAS regions tested in the SMR/colocalization analysis? This would make a connection between the SMR/colocalization analysis and the permuted database of mQTLs described before.

While the suggested threshold of nominal p-value of 5e-8 is more stringent than the permutation-based threshold for mQTLs, it is likely not true for imQTLs. Please include also the FDR of the imQTLs included to SMR analysis for an assessment of the significance of the imQTL effect.

Response: From the total amount of 214,830 mSites with FDR < 0.05 in the permuted database, 12,228; 10,343 and 38,412 mSites were located in the BIP, MDD and SCZ-associated regions, respectively, and therefore, were considered for the colocalization analyses. In the case of SMR, the 130,736 mSites considered ($P_{nominal} < 5x10^{-8}$)

were among the mSites of the permuted database and therefore had an FDR < 0.05 (please have a look at Figure 1 in this letter). These figures have been included in the *Results* section.

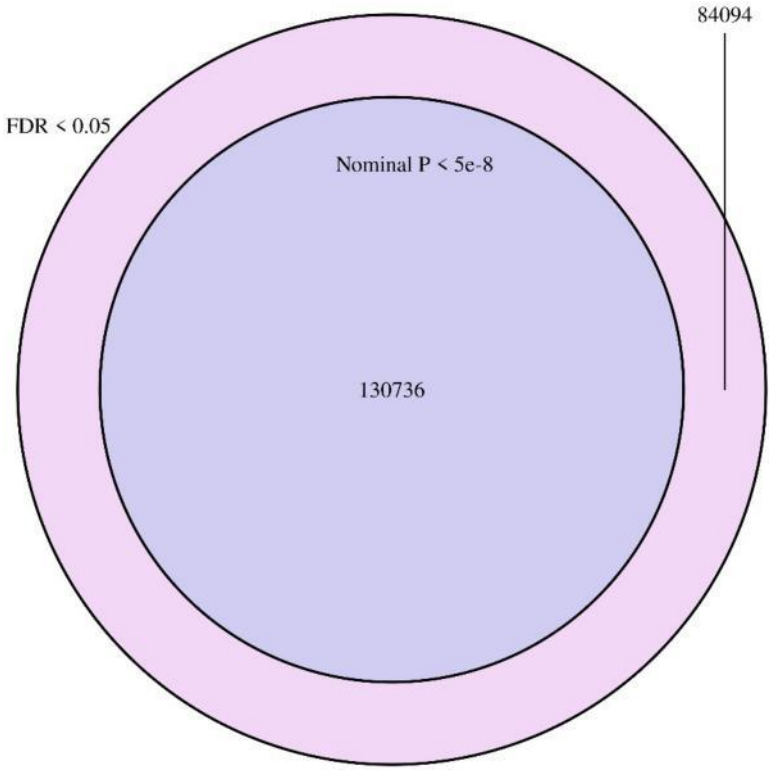

Figure 1: Venn Diagram showing how all the mSites of the nominal database under a p-value of $5x10^{-8}$ are among the FDR-significant mSites in the permuted database.

On the other hand, FDR P-values have been added in Supplementary Data 10 in STB- and TB-imQTLs. Briefly, all STB-imQTLs employed for SMR, with $P_{nominal} < 5x10^{-8}$, have an FDR < 0.05. In turn, 337/557 TB-imQTLs have an FDR < 0.05, with a maximum FDR = 0.09 among those above the threshold. The TB-imQTL resulting from the SMR approach for BIP that has been reported in Figure 6, for example, has a FDR = 0.05.

4. Sharing between imQTLs using the Storey's pi1 method. Please correct the typo in Storey's pi1 method. It should be pi1 (not p1), where pi stands for the symbol "pi".

Response: We thank the reviewer for the comment. The typo has been corrected.

Also, please plot the p-value histograms corresponding to the significance of the genotype*TB and genotype*GA effect of the STB-imQTLs as

supplementary figures. These figures give additional visual support of sharing between imQTLs.

Response: We thank the reviewer for the comment, but we do not find it necessary to add any other figures regarding imQTL sharing. The point is that our aim never was to focus on the global sharing between imQTL sets, but to find the overlapping imQTLs to compare the effect directions between cell type-imQTLs, or between imQTLs of a given cell type and those interacting with the time of gestation. We think that the suggestion of the reviewer has helped understand that the sharing can go beyond a simple overlap between significant QTLs, but this is not the story in which we would like to go deeper.

In my opinion Figure 2C is misleading. The figure legend says that the intersection between STB- and TB-imQTLs is represented as Venn diagrams in the figure, but the set of CpG site-variant pairs include SNPs in LD (if I have understood the analysis correctly) and statistically non-significant TB-imQTLs. If the authors wish to compare the direction of genotype main effect, then they should at least focus on independent SNPs per CpG site.

Response: Following our previous response and as noted by the reviewer, our aim at this point has always been to compare the direction of the main genotype effect in the different imQTL sets. For this purpose, the first step has consistently been to identify the particular overlapping imQTLs.

In the first version of the manuscript, we identified the overlapping imQTLs between the cell type-imQTLs with $P_{nominal} < 5x10^{-8}$ and drew the correlation between the interaction effect sizes that those particular imQTLs had for each cell type. This resulted in a very strong and significant negative correlation.

In the second version of the manuscript, we were asked to correct QTLs at two levels and used eigenMT as a surrogate of permutation in the imQTLs. This resulted in a very limited number of significant imQTLs, particularly for TB, as expected. Therefore, we took the list of eigenMT FDR < 0.05 STB-imSites and recovered all the associated imQTLs with $P_{nominal} \times N_{eff}$ (effective number of independent variants in the *cis* window estimated by eigenMT) < 0.05. Then we retrieved those same imQTLs from the TB-imQTL set, only if they showed $P_{nominal} \times N_{eff} < 0.05$, and found the overlap. The correlation between the interaction effect sizes of these imQTLs in the two cell types was again negative and highly significant but it is true that our attempt to compute the correlation only with imSites that pass the two

levels of correction, led us to select a number of imVariants that are in linkage disequilibrium.

In the present version, we have tried to reach a midway: if we aim to make a comparison with independent imVariants, we cannot limit to imQTLs filtered by eigenMT and FDR, as these are too restrictive for TB-imQTLs. Thus, we have considered the list of unique imSites with $P_{nominal} < 5x10^{-8}$ common between STB- and TB-imQTLs (12 imSites). For each of these imSites we have considered the top TB-imQTL, and we have calculated the correlation between their interaction effect sizes and the effect sizes reported for those same imQTLs in the STB-imQTL set. As a result, and despite the limited number of points in the regression, we have obtained a highly significant negative correlation ($P = 4.1x10^{-10}$ and $r = -0.99$). These results reveal that the correlation is strong and consistent, regardless of the filters and the inclusion / exclusion by linkage, or other criteria. From our point of view, there is no perfect approach but this effect seems to be biological and deserves to be mentioned in the article.

We have explained this in *Results*:

"Beyond the global estimated sharing between cell types, and as STB and TB proportions are negatively correlated in term placentas, we wanted to know whether the genotype-cell type interaction had opposite effect directions in the overlapping imQTLs. Due to the limited number of eigenMT-significant imQTLs, we retrieved the unique interacting mSites (imSites) for each one of the two cell types analyzed, as well as for GA, with $P_{nominal} < 5x10^{-8}$. We found an overlap of 12 imSites between the two cell types, and no overlap between STB- and GA-imSites. For each of the cell type-overlapping imSites, we considered the top TB-imQTL, and we calculated the correlation between their interaction effect sizes and the effect sizes reported for those same imQTLs in the STB-imQTL set. As a result, and despite the limited number of points in the regression, we obtained a highly significant negative correlation ($P = 4.1x10^{-10}$ and $r = -0.99$) , as in the example in Figure 2D, E and F. Of note, the imQTL in Figure 2 is the only eigenMT-significant TB-imQTL (FDR = 0.05)".

And in *Methods*:

"Therefore, we retrieved the imSites with a $P_{nominal} < 5x10^{-8}$ in STB-and TB-imQTLs to draw a Venn diagram. Finally, with this previous list of imSites, we calculated the correlation of the interaction effect

sizes of the overlapping imQTLs considering only the top imVariant from the TB-imQTLs database".

5. Colocalization analysis. Colocalization analysis is performed using all the genetic variants in the specified region (including also non-significant ones for a given molecular phenotype or trait of interest). Looking at the Supplementary Data 6, the number of SNPs analyzed (column nsps) is very low for some tested loci, e.g., nsps = 4 for region 98357153-96396153, where the lead BIP GWAS SNP is rs4619651 and CpG site is cg19932515. Could the authors please clarify how the colocalization analysis was performed?

Response: We would like to very sincerely thank the reviewer for his/her advice in the colocalization approach. He/she has helped us correct mistakes related to this methodology, and we are really grateful for the accurate detection of errors in the procedure. We apologize and this time, honestly believe to have solved all the errors: we have repeated the colocalization analyses with all mVariants, significant or not, associated with the mSites (FDR < 0.05 in the permuted database) that overlap the GWAS regions of BIP, MDD and SCZ.

We copy below the modifications derived from the new colocalization analysis in the different sections of the manuscript:

In *Results*:

"To replicate the pleiotropic associations identified in the SMR analyses, we performed a colocalization test with the same data. The Bayesian colocalization analysis implemented in the coloc R package[52] focuses on finding the intersection between significant loci independently associated with two phenotypes, comparing the association patterns in the GWAS and QTL analyses across genomic regions, and thus combining the summary statistics into posterior probabilities for five hypotheses (see Methods for more information). Colocalization was performed across 64, 102, and 287 genomic regions defined in the GWAS associated with BIP, MDD and SCZ, respectively. Regarding mQTLs, from the total amount of 214,830 mSites with FDR < 0.05 in permuted database, 12,228, 10,343 and 38,412 mSites were considered for BIP, MDD and SCZ colocalization analyses, respectively. . In BIP, the posterior probabilities for 64 regions, involving 6,487 DNAm sites in 7,123 region-CpG pairs, were supportive of a colocalization signal (PPA3+PPA4 > 0.9) (Figure 3A and Supplementary Data 6). In MDD, the posterior probabilities for 79 regions, involving 3,850 DNAm

sites in 3,895 region-CpG pairs, were supportive of a colocalization signal (PPA3+PPA4 > 0.9) (Figure 3B and Supplementary Data 6). Finally, in SCZ, the posterior probabilities for 274 regions, involving 19,661 DNAm sites in 21,071 region-CpG pairs, were supportive of a colocalization signal (PPA3+PPA4 > 0.9) (Figure 3C and Supplementary Data 6). When the overlap with SMR hits was assessed, out of the 30 SMR hits in BIP, 28 DNAm sites (93%) showed evidence of colocalization with BIP; in MDD, 20 DNAm sites (71%) out of the 28 SMR hits showed evidence of colocalization; and from the 214 SMR hits in SCZ, 198 DNAm sites (92%) colocalized with SCZ".

[...]

"*Next, we identified the intersect between SMR, colocalization and eQTM results for each trait (Supplementary Figure 3). In the case of MDD, one placental DNAm site colocalized with a MDD-associated genomic locus on chromosome 14, and DNAm levels correlated with the placental expression of LRFN5. In SCZ, 40 placental DNAm sites colocalized with 12 SCZ-associated loci on chromosomes 1, 3, 6, 7, 8, 11, 12, 18, 19 and 22, with DNAm levels that correlated with the placental expression of 23 nearby genes (GLB1L3, H2BC6, H3C4, KCNG2, LETM2, NAGA, PGBD1, PSMG3, RNF39, SDCCAG8, SFMBT1, SLC6A16, TOB2P1, TRIM27, TXNL4A, VARS2, VPS37B, ZKSCAN4, ZNF165, ZNRD1-AS1, ZSCAN12, ZSCAN12P1, and ZSCAN23) (Supplementary Figure 5)*".

And in *Methods*:

"*Recalling the definition of cis-mQTLs in our analysis, all mSites with FDR < 0.05 and located within ±0.5 Mb of the regions defined in the original neuropsychiatric GWAS were tested. Colocalization analysis was performed to identify the overlap between causal mQTLs and GWAS-significant loci, as described in Giambartolomei C. et al. 2014, with the R coloc package[52]. In total 12,228, 10,343, and 38,412 mSites were tested for BIP (n=63 autosomal regions), MDD (n=98 autosomal regions) and SCZ (n=279 autosomal regions), respectively. Because data for all SNPs (regardless of significance) are required for this analysis, first, the mQTL analysis was rerun for these mSites so that all association statistics could be recorded for all mVariants within ±0.5 Mb of the DNAm site. Then, we retrieved the regression coefficients, their variances and the SNP minor allele frequencies from the original GWAS data and our mQTLs, and the prior probabilities were left as their default values. This*

methodology quantifies the support across the results of each GWAS for five hypotheses by calculating the posterior probabilities, denoted as PPi for hypothesis Hi".

Minor comments

1. Figure 1a legend: I'm still confused by the Figure 1a. The legend says that x-axis represents the absolute distance in Mb, but the red line represents the raw median distance. This does not make sense. Is the red line representing the median absolute distance? It should be noted that the distance is calculated between the lead mVariant and mSite (see the comment about the terminology used in molQTL studies).

Response: The red line from Figure 1a represents the median absolute distance in Mb calculated between the lead mVariant and the mSite of the permuted database adjusted by FDR. The legend from Figure 1 has been corrected.

2. absolute median distance – would it be more correct to say "median absolute distance"?

Response: We have corrected this in the manuscript.

3. Thank you for clarifying the definition of positive and negative imQTLs. The following sentence in the main text needs modification: "In summary, positive imQTLs show increased DNAm related to the main genotype as a function of cell proportion, while negative imQTLs show a decrease." As noted in the Methods section, imQTLs with positive direction of effect (positively correlated with the cell type proportion) showed an increase in the main genotype effect from low to high cell proportions, i.e., the mQTL effect size is increasing with the cell type abundance.

Response: We would like to thank the reviewer for this correction. We have modified the sentence in the main text.

4. Please include P-values of the (i)mQTL effect to Figure 2D-F (and to other similar plots).

Response: We have decided not to include the P-values in the (i)mQTL figures for two reasons that we hope the reviewer will understand. First, because as different (i)mQTL sets have been used for each approach, it would be difficult to decide which P-value is the most interesting in each case, and this would probably led to increase the heterogeneity among the figures. For instance, in Figure 2 we would put eigenMT FDR for the imQTLs, and $P_{nominal}$ for

the standard mQTL while, in Figure 6, we would very likely put $P_{nominal}$ for the imQTL because those with a $P_{nominal} < 5x10^{-8}$ are the ones that have been selected for SMR. Second, because this high heterogeneity would probably result in less balanced and harmonic figures. In conclusion, we have decided to include this information in the main figure legends instead of in the figures themselves.

5. The authors note that "Out of the 28 and 214 SMR-significant DNAm sites in MDD and SCZ, we found that 3 and 43 correlated with the expression levels of 4 and 26 significant eQTM-genes (FDR < 0.05), respectively, strongly supporting the placental function of the CpGs identified (Supplementary Data 7 and Supplementary Figure 3)." – does 3 out 28 and 26 out 214 SMR-significant DNAm sites that are also correlated with gene expression levels indicate strong support of the CpG sites?

Response: We have corrected the main text:

"Out of the 28 and 214 SMR-significant DNAm sites in MDD and SCZ, we found that 3 and 43 correlated with the expression levels of 4 and 26 significant eQTM-genes (FDR < 0.05), respectively, supporting the placental function of these CpG sites (Supplementary Data 7 and Supplementary Figure 3)".

6. Could the authors include the locuszoom plot of mQTLs for cg10318063 on Figure 4 (to visually represent the colocalization between GWAS of MDD and mQTLs for cg10318063)? Of note, were there a cell type imQTL for cg10318063?

Response: The locuszoom is already there in Figure 4. cg10318063 has not been detected as a cell type imQTL in our study.

7. The locuszoom plot on Figure 5A looks odd. Why is the association peak not in the center of the region tested? Please also include a locuszoom plot of mQTLs for cg11165035 to Figure 5.

Response: The association peak is not centered because the SNP and the histone genes of interest are quite distant. We cannot center the peak more and visualize all these elements at the same time. Does the reviewer want us to add all the mVariants for this CpG site in another locuszoom plot? Honestly, we do not see the point of doing this: SNPs are used as instrumental variables in SMR, while the output of the colocalization approach consists on CpG and disorder-associated region pairs. From our point of view, the locuszoom of the GWAS region is enough and the most common plot in works like ours in the current literature.

8. Please also include a locuszoom plot of cell type imQTLs to Figure 6 and Supplementary Figure 7.

Response: Same as above.

9. Why were only the top 10,000 mQTL-CpGs tested for enrichment and depletion of overlap with tissue-specific and cell-type-specific regulatory features? This is not mentioned in the main text.

Response: Dr. Charles Breeze, the developer of eFORGE, is our coauthor and suggested to do it that way because when an input CpG set is larger than 10,000 probes (and taking into consideration the total amount of probes in the EPIC array), it risks limiting the randomization in the background bins, as, depending on the annotation, background bins may be more limited in their choice. For example, if we input the majority of CpGs that are in the promoter-CpG Island bin, then the method can only select a limited set of CpGs from the promoter-CpG Island bin in the background, and therefore the background is not as rich as it otherwise could be. In summary, selecting only the top 10,000 probes is a good way to avoid saturating the background bins, especially those related to CpG Island annotation categories that may be more limited.

We have added a brief explanation in *Methods* as follows:

"Briefly, we selected only the top 10,000 probes following the advice of the developers to avoid saturating the background bins, especially those related to CpG Island annotation categories that may be more limited, taking into consideration the total amount of probes per category in the Illumina EPIC array".

REVIEWER COMMENTS

Reviewer #3 (Remarks to the Author):

The manuscript has improved during the revisions, but there are still parts of the analysis reported that need improvement to ensure the quality of the results.

1. Colocalization analysis – thank you for fixing the error in the colocalization analysis. However, the presentation of the colocalized signals needs to be corrected. The authors now use the sum of PP3 and PP4 to claim support for colocalization. This is not correct! As stated in the methods section PP3 denotes the posterior support for hypothesis H3 (there exist two distinct causal variants, one for each trait) and PP4 denotes the posterior support for hypothesis H4 (there exist a single causal variant common to both traits). With colocalization analysis we are interested in the examples, where both the traits (GWAS signal and mQTL signal) in the given locus seem to be driven by the same causal variants, i.e., a high PP4 value. Different thresholds for claiming support for colocalization are used, with PP4 > 0.8 being often used. Please clearly note in the methods section the criteria used for claiming support for colocalization. Also, please correct the number of region-CpG pairs showing colocalization in the main text. Please note that the example highlighted on Figure 5 is not supportive of colocalization between SZC GWAS and mQTLs for cg11165035 (PP3 = 0.999946306358956, PP4 = 1.27581938404575e-05 from Supplementary Data 6).

After statistical colocalization analysis, visualization of the colocalized loci is recommended (Aguet et al., 2023, https://doi.org/10.1038/s43586-022-00188-6). This is done by visualizing the p-value landscapes in the locus for the GWAS trait and mQTL association (i.e., so called locuszoom plots for both traits). Alternatively, one could also plot the GWAS p-values and mQTL p-values as a scatter plot (Liu et al., 2019, https://doi.org/10.1038/s41588-019-0404-0). Thus, I disagree with the authors that locuszoom of the GWAS region is enough and the most common plot in similar works. For instance, please see the supplementary materials accompanied with the GTEx v8 publication to see the examples. Please add locuszoom plots for both the traits (and/or the scatter plot) and add PP4 value in the figure legend, when showing examples of colocalization.

In the methods section, the authors note that:

“Because data for all SNPs (regardless of significance) are required for this analysis, first, the mQTL analysis was rerun for these mSites so that all association statistics could be recorded for all mVariants within ±0.5 Mb of the DNAm site.“

Why is this step needed? The association statistics for the selected mSites and genetic variants in a given locus could be retrieved from the nominal database.

2. Presentation of sharing between imQTLs using the Storey's pi1 method – even though the authors never aimed to focus on global sharing between imQTLs sets, then reporting the pi1 value together with the nominal p-value histograms as supplementary figures gives additional visual support for sharing. Most of all it adds validity to the reported statistical estimate as the p-values used for input for pi1 analysis are plotted. There's no need to go deeper in the story of global sharing; it is just a way to convince the reader of the correctness of this analysis. For instance, please see the supplementary materials accompanied with the GTEx v8 publication to see the examples of similar analysis, where the pi1 statistic is used.

Reviewer #4 (Remarks to the Author):

I have examined the manuscript, reviewer comments, and corresponding authors answers, for all three rounds of reviews. My insights:

Summary:

As per editor request, I have focused on the comments related to methodology from reviewer #3. Overall, all the comments are spot on. While I think that the manuscript has significantly improved over review rounds, I would not publish it in its current form, until the PP3+PP4 issue is addressed as per reviewer guidelines (see comments below). I would suggest to the authors to discuss cases with significant mendelian randomization AND significant colocalization (based solely on PP4 threshold), as suggested in PMC7612737. Notice that this involves substantial changes, e.g. for Fig. 5 and related text.

Comments:

. It is indeed a standard practice in the field to include hidden factors – being PEERs, PCs, or others - as it has been shown repeatedly that increased replication is observed when adjusting for them. Also, two rounds of multiple testing correction – per locus and across genes – are needed to determine molecules with at least 1 significant QTL, independently of the molecular phenotype considered (being expression, methylation, protein, etc.). The fact that the authors misinterpreted the methodology of the seminal work they are referring to (GoDMC) is concerning; it was only when reviewer #3 quoted explicitly that GoDMC was indeed adjusting for hidden factors, that the authors mapped mQTLs accordingly. Having said that, the authors addressed this issue and to the best of my knowledge, the current version of mQTL maps follow the field best practices.

. Overall, the authors do not correctly assess interpretable results in the context of colocalization, and mendelian randomization given a particular colocalization scenario. Indeed, the sum of PP3 and PP4 is not, by any means, indicative of colocalization, defined as one variant affecting both analyzed traits. Consequently, loci tested for colocalization with PP3+PP4>0.9 are not indicative of mQTL-GWAS signal colocalization. As noted by the reviewer, the example highlighted on Figure 5 is not supportive of colocalization between SZC GWAS and mQTLs for cg11165035, because the posterior probabilities indicate that there exist two distinct causal variants, one impacting methylation and another one impacting SCZ. In the locus zoom plot, one can observe variants in partial LD with the instrument mQTL variant that have a smaller GWAS p-values (than the instrument variant). To say that the GWAS p-value landscape for that locus suggests lack of colocalization. Regarding authors answer about the choice of the colocalization threshold “because we observed that being more restrictive was killing a lot of potentially functional signals.” It is unclear what they mean ‘being more restrictive’ as they do not specify. I assume they refer to applying a conventional PP4>0.8 threshold to identify colocalizations. The “potentially functional signals” that pass PP3+PP4>0.9 but do not pass PP4>0.8 (or even a more lenient colocalization threshold, i.e. PP4>0.5) are likely false positive mendelian randomization findings. See PMC7612737 for interpretation of evidence from Mendelian randomization and colocalization. “This suggests that the positive estimate in the polygenic Mendelian randomization analysis including the APOE variants arises from violation of the instrumental variable assumptions due to linkage disequilibrium. This emphasizes the importance of checking for outliers in polygenic Mendelian randomization analyses and using colocalization to test the Mendelian randomization assumptions at a particular locus.” ...”Without this, cis-Mendelian randomization analyses can provide false positive findings similarly to candidate gene studies” ...” We therefore strongly recommend that positive cis-Mendelian randomization analyses be accompanied by a corresponding colocalization analysis, as is becoming more common in the literature.

. I agree with rev. # 3 and disagree with the authors that locuszoom of the GWAS region is enough and the most common plot in similar works. An informative plot needs to contain the GWAS and QTL signal, being piled up or in scatterplot form. That helps the reader to visually assess the veracity of the colocalization by observing the presence of 2ary signals in any of the two traits and concordance of signals given LD patterns in both traits.

Reviewer #3 (Remarks to the Author):

The manuscript has improved during the revisions, but there are still parts of the analysis reported that need improvement to ensure the quality of the results.

1. Colocalization analysis – thank you for fixing the error in the colocalization analysis. However, the presentation of the colocalized signals needs to be corrected. The authors now use the sum of PP3 and PP4 to claim support for colocalization. This is not correct! As stated in the methods section PP3 denotes the posterior support for hypothesis H3 (there exist two distinct causal variants, one for each trait) and PP4 denotes the posterior support for hypothesis H4 (there exist a single causal variant common to both traits). With colocalization analysis we are interested in the examples, where both the traits (GWAS signal and mQTL signal) in the given locus seem to be driven by the same causal variants, i.e., a high PP4 value. Different thresholds for claiming support for colocalization are used, with PP4 > 0.8 being often used. Please clearly note in the methods section the criteria used for claiming support for colocalization. Also, please correct the number of region-CpG pairs showing colocalization in the main text. Please note that the example highlighted on Figure 5 is not supportive of colocalization between SZC GWAS and mQTLs for cg11165035 (PP3 = 0.999946306358956, PP4 = 1.27581938404575e-05 from Supplementary Data 6).

Response: We thank the reviewer for his/her comments and suggestions. We have changed the *coloc* criterium and based solely on PP4 threshold, as suggested by the reviewer. This has involved substantial changes as pointed by reviewer 3 and also 4, including main Figures 3 and 5, previous Supplementary Figure 5 (now 6) and related text in Results, as follows:

"*To replicate the pleiotropic associations identified in the SMR analyses, we performed a colocalization test with the same data. The Bayesian colocalization analysis implemented in the coloc R package*[52] *focuses on finding the intersection between significant loci independently associated with two phenotypes, comparing the association patterns in the GWAS and QTL analyses across genomic regions, and thus combining the summary statistics into posterior probabilities (PPA) for five hypotheses (see Methods for more information). Colocalization was performed across 64, 102, and 287 genomic regions defined in the GWAS associated with BIP, MDD and SCZ, respectively. Regarding mQTLs, from the total amount of 214,830 mSites with FDR < 0.05 in permuted database, 12,228, 10,343 and 38,412 mSites were considered for BIP, MDD and SCZ colocalization analyses, respectively. In BIP, the posterior probabilities for 47 regions, involving 280 DNAm sites in 307 region-CpG pairs, were supportive of a colocalization signal (PPA4 > 0.8) (Figure 3A and Supplementary Data 6). In MDD, the posterior probabilities for 55 regions, involving 246 DNAm sites in 253 region-CpG pairs, were supportive of a colocalization signal (PPA4 > 0.8) (Figure 3B and Supplementary Data 6). Finally, in SCZ, the posterior probabilities for 161 regions, involving 865 DNAm sites in 916 region-CpG pairs, were supportive of a colocalization signal (PPA4 > 0.8) (Figure 3C and Supplementary Data 6). When the overlap with SMR hits was assessed, out of the 30 SMR hits in BIP, 15 DNAm sites (50%) showed evidence of colocalization with BIP; in MDD, 11 DNAm sites (39.28%) out of the 28 SMR hits showed evidence of colocalization; and from the 214 SMR hits in SCZ, 62 DNAm sites (28.91%) colocalized with SCZ.*"

The main finding highlighted in SCZ is now *NAGA* (PP4 = 0.901), and this has involved rewriting of part of the Discussion:

"*In the case of NAGA, in 2021, Li and colleagues reinforced the hypothesis of the neurodevelopmental origins of SCZ by demonstrating that this gene modulates the dendritic spine density, a process with a pivotal role in cognitive function, learning and memory[58]. Deficits in dendritic spines could produce alterations in the neuronal circuitry within and across multiple brain regions. More precisely, lower spine density has been observed in the neocortex of SCZ patients[97]. Interestingly, in the context of pregnancy, gestational stress leads to significant changes of spine density and dendritic complexity in the prefrontal cortex, located within the neocortex. These effects have been reported to occur in the offspring of stressed mothers, in a sex-specific manner[98]."*

After statistical colocalization analysis, visualization of the colocalized loci is recommended (Aguet et al., 2023, https://doi.org/10.1038/s43586-022-00188-6). This is done by visualizing the p-value landscapes in the locus for the GWAS trait and mQTL association (i.e., so called locuszoom plots for both traits). Alternatively, one could also plot the GWAS p-values and mQTL p-values as a scatter plot (Liu et al., 2019, https://doi.org/10.1038/s41588-019-0404-0). Thus, I disagree with the authors that locuszoom of the GWAS region is enough and the most common plot in similar works. For instance, please see the supplementary materials accompanied with the GTEx v8 publication to see the examples. Please add locuszoom plots for both the traits (and/or the scatter plot) and add PP4 value in the figure legend, when showing examples of colocalization.

Response: We have added the locuszoom plots of the mQTLs to illustrate colocalization as requested, as well as PP4 in the figure legends. Please note these changes in both Figure 4 and new Figure 5.

In the methods section, the authors note that: "Because data for all SNPs (regardless of significance) are required for this analysis, first, the mQTL analysis was rerun for these mSites so that all association statistics could be recorded for all mVariants within ±0.5 Mb of the DNAm site. "Why is this step needed? The association statistics for the selected mSites and genetic variants in a given locus could be retrieved from the nominal database.

Response: The nominal database is very extensive, and it is much harder retrieving the mQTLs from the raw data than recalculating them. For instance, this methodology is the one that was followed in Hannon e*t al.* 2016 and as far as we know, is common practice. In addition, it gives identical results.

2. Presentation of sharing between imQTLs using the Storey's pi1 method – even though the authors never aimed to focus on global sharing between imQTLs sets, then reporting the pi1 value together with the nominal p-value histograms as supplementary figures gives additional visual support for sharing. Most of all it adds validity to the reported statistical estimate as the p-values used for input for pi1 analysis are plotted. There's no need to go deeper in the story of global sharing; it is just a way to convince the reader of the correctness of this analysis. For instance, please see the supplementary materials accompanied with the GTEx v8 publication to see the examples of similar analysis, where the pi1 statistic is used.

Response: This is the first time we apply the Storey's method. As the reviewer already know, our aim was never to talk about sharing but to find overlapping imQTLs and to see whether we could find some "patterns" derived from the complex biology of pregnancy and placental development. Having said that, please find below the histograms, as requested:

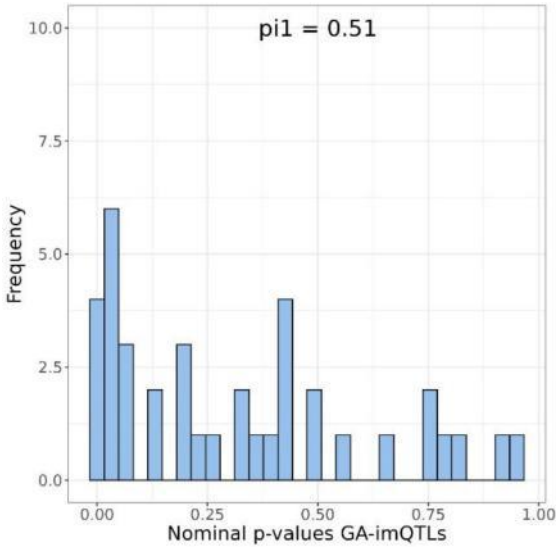

Figure 1. Distribution of the nominal p-values of the GA-imQTLs that were used to compute the Storey's pi1 value of sharing with the significant STB-imQTLs.

In the case of TB-imQTLs, we encountered an error using the qvalue() function. This error occurred because the p-values of the significant STB-imQTLs in the TB-imQTL set were extremely small (or significant), and this violated the assumption that the p-values should follow a uniform distribution between 0 and 1. To address this, we followed John D. Storey's recommendation from a Bioconductor discussion and used the qvalue_trunc() function instead. This alternative method rescales the p-values before applying the same statistical test as qvalue(). The result is a pi1 value of 0.68, as reported in the newest versions of the article.

Here the histograms:

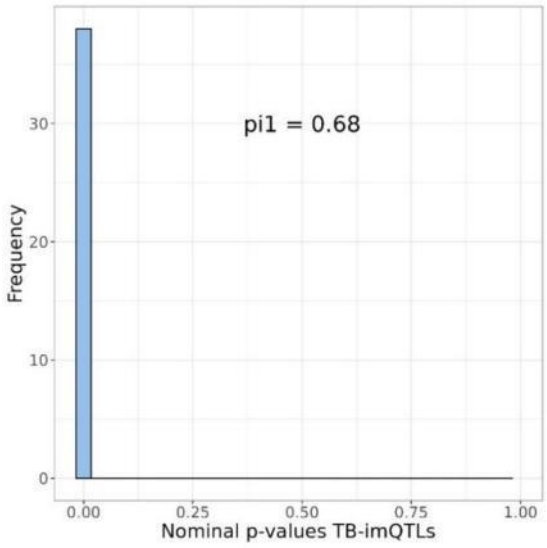

Figure 2. Distribution of the nominal p-values of the TB-imQTLs that were used to compute the Storey's pi1 value of sharing with the significant STB-imQTLs (X axis 0-1).

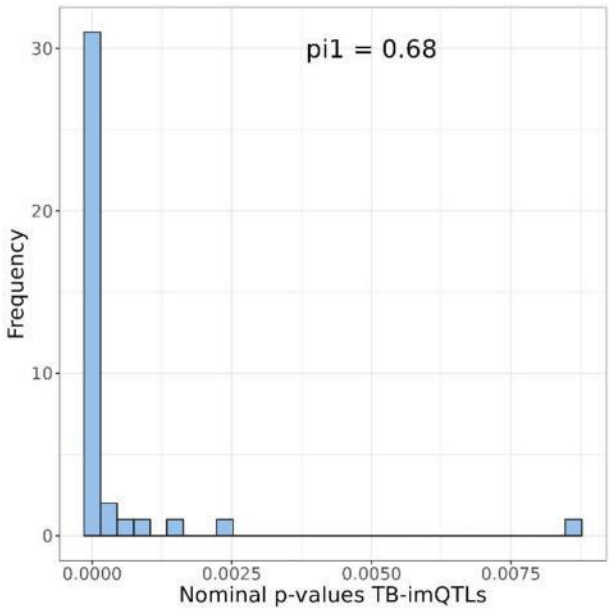

Figure 3. Distribution of the nominal p-values of the TB-imQTLs that were used to compute the Storey's pi1 value of sharing with the significant STB-imQTLs. X axis from 0 to the maximum p value, as suggested by the developers of qvalue (StoreyLab/qvalue source: R/qvalue_trunc.R (rdrr.io)).

We are confident that the tools provided by the developers may be adequate but we have no objections if the reviewer thinks that the adaptation of the Storey's pi1 method is not correct for the evaluation of sharing in the present particular case, and can erase the related sentence in the results section. For the moment, histograms have been added as Supplementary Figure 3.

Reviewer #4 (Remarks to the Author):

I have examined the manuscript, reviewer comments, and corresponding authors answer's, for all three rounds of reviews. My insights:

Summary:

As per editor request, I have focused on the comments related to methodology from reviewer #3. Overall, all the comments are spot on. While I think that the manuscript has significantly improved over review rounds, I would not publish it in its current form, until the PP3+PP4 issue is addressed as per reviewer guidelines (see comments below). I would suggest to the authors to discuss cases with significant mendelian randomization AND significant colocalization (based solely on PP4 threshold), as suggested in PMC7612737. Notice that this involves substantial changes, e.g. for Fig. 5 and related text.

Response: We thank the reviewer for his/her comments and suggestions. We have changed the *coloc* criterium and thus, discuss only cases with significant Mendelian Randomization AND significant colocalization (based solely on PP4 threshold), as suggested. This has involved substantial changes as pointed by reviewer 3 and 4, including several figures and related text. Please see the response to the first point of reviewer 3 for more details.

Comments:

1. It is indeed a standard practice in the field to include hidden factors – being PEERs, PCs, or others - as it has been shown repeatedly that increased replication is observed when adjusting for them. Also, two rounds of multiple testing correction – per locus and across genes – are needed to determine molecules with at least 1 significant QTL, independently of the molecular phenotype considered (being expression, methylation, protein, etc.). The fact that the authors misinterpreted the methodology of the seminal work they are referring to (GoDMC) is concerning; it was only when reviewer #3 quoted explicitly that GoDMC was indeed adjusting for hidden factors, that the authors mapped mQTLs accordingly. Having said that, the authors addressed this issue and to the best of my knowledge, the current version of mQTL maps follow the field best practices.

Response: We understand that in some fields PEERs and PCs may be extremely common but we are researchers at the Pregnancy and Childhood Epigenetics (PACE) consortium and in the Placenta section, it is now for the first time that we are considering including these kind of latent factors in our Epigenome-Wide Association Studies or EWAS. This is because methylation microarrays are by far less "noisy" than other sources of data such as transcriptome data, and very especially, RNA sequencing data.

We are happy that the reviewer thinks that we are now following the best practices in the field.

Overall, the authors do not correctly assess interpretable results in the context of colocalization, and mendelian randomization given a particular colocalization scenario. Indeed, the sum of PP3 and PP4 is not, by any means, indicative of colocalization, defined as one variant affecting both analyzed traits. Consequently, loci tested for colocalization with PP3+PP4>0.9 are not indicative of mQTL-GWAS signal colocalization. As noted by the reviewer, the example highlighted on Figure 5 is not supportive of colocalization between SZC GWAS and mQTLs for cg11165035, because the posterior probabilities indicate that there exist two distinct causal variants, one impacting methylation and another one impacting SCZ. In the locus zoom plot, one can observe variants in partial LD with the instrument mQTL variant that have a smaller GWAS p-values (than the instrument variant). To say that the GWAS p-value landscape for that locus suggests lack of colocalization.

Response: We have replaced the hit in Figure 5 by another hit that shows evidences of colocalization considering only PP4. Similarly, we have changed figures and text accordingly thorough the manuscript, as explained before.

Regarding authors answer about the choice of the colocalization threshold "because we observed that being more restrictive was killing a lot of potentially functional signals." It is unclear what they mean 'being more restrictive' as they do not specify. I assume they refer to applying a conventional PP4>0.8 threshold to identify colocalizations.

The "potentially functional signals" that pass PP3+PP4>0.9 but do not pass PP4>0.8 (or even a more lenient colocalization threshold, i.e. PP4>0.5) are likely false positive mendelian randomization findings. See PMC7612737 for interpretation of evidence from Mendelian randomization and colocalization. "This suggests that the positive estimate in the polygenic Mendelian randomization analysis including the APOE variants arises from violation of the instrumental variable assumptions due to linkage disequilibrium. This emphasizes the importance of checking for outliers in polygenic Mendelian randomization analyses and using colocalization

to test the Mendelian randomization assumptions at a particular locus." ..."Without this, cis-Mendelian randomization analyses can provide false positive findings similarly to candidate gene studies" ..." We therefore strongly recommend that positive cis-Mendelian randomization analyses be accompanied by a corresponding colocalization analysis, as is becoming more common in the literature.

Response: Yes, exactly, we refer to the fact that, when applying PP4>0.8 instead of PP3+PP4>0.8, many Mendelian Randomization results that seem to have a functional impact on placental expression disappear from the list. As both reviewer 3 and 4 have already noticed, we are not expert statisticians, but we wondered whether discordances between Mendelian Randomization and colocalization (mostly after using methods like HEIDI, designed to avoid problems with instrumental variables in linkage disequilibrium) could not be false negative results of the second, rather than false positive of the first. As biologists, and considering the expression results, we though that it was worth and reasonable to report the findings above PP3+PP4=0.8 but we may have been wrong and thus, we have finally corrected the manuscript accordingly, as requested by reviewers 3 and 4.

I agree with rev. # 3 and disagree with the authors that locuszoom of the GWAS region is enough and the most common plot in similar works. An informative plot needs to contain the GWAS and QTL signal, being piled up or in scatterplot form. That helps the reader to visually assess the veracity of the colocalization by observing the presence of 2ary signals in any of the two traits and concordance of signals given LD patterns in both traits.

Response: We have added the mQTL locuszoom plots in every situation in which it applies, as requested. Please note these changes in both Figure 4 and new Figure 5. Please also note that in the case of Figure 6 we are not reporting colocalization between the BIP GWAS and the "regular" placental mQTLs, but a pleiotropic association between a trophoblast-interacting mQTL and BIP, according only to Mendelian Randomization.

REVIEWERS' COMMENTS

Reviewer #3 (Remarks to the Author):

1. Coloc analysis and representation of the results: thank you for making the edits! Some further comments on this topic:

a) In the text the authors use posterior probabilities (PPA) and PPA4, but in the figure legends PP4 is used. Please be consistent. Usually the posterior probabilities are denoted as PP_i, e.g., PP4.

b) On Figure 4c, there seems to be a data point that is not a full circle on the top left corner. When I skimmed through the supplementary figures, I noticed a similar issue with Supplementary Figure 6. On panel a)-d) there seems to be like a white box on the top left corner of the scatter plots. Please show all the data points. Please check all similar plots.

c) Please add a locuszoom plot of trophoblast-interacting mQTL in the locus by plotting the interaction p-values to Figure 6. Also, it rather seems that there are two different genetic variants in the BIP GWAS and imQTL analysis (Figure 6). Could you please elaborate more how you reached the suggestion of causal association between placental DNAm and BIP at this CpG site - “Interestingly, rs11637581 is not directly associated with BIP in the original GWAS, suggesting a possible causal association between placental DNAm and BIP at this CpG site”? What was the p-value from the SMR analysis? Add it to the main text too. Also, the lead SNP in the BIP GWAS is below the genome-wide significance threshold (p = 5e-8). Were there any filters based on the p-values in the GWAS analysis similar to filtering of imQTLs for SMR?

2. Storey’s pi1 value: thank you for adding the Supplementary Figure 3 and finding a solution to overcome the issue with very small p-values (STB-imQTLS and TB-imQTLs comparison).

Reviewer #4 (Remarks to the Author):

The comments regarding the colocalization threshold criteria, locuszoom plots, and Fig 5 illustrated case have been satisfactorily addressed.

REVIEWERS' COMMENTS

Reviewer #3 (Remarks to the Author):

1. Coloc analysis and representation of the results: thank you for making the edits! Some further comments on this topic:
a) In the text the authors use posterior probabilities (PPA) and PPA4, but in the figure legends PP4 is used. Please be consistent. Usually the posterior probabilities are denoted as PP_i, e.g., PP4.
Response: We have replaced all PPAs for PPs.

b) On Figure 4c, there seems to be a data point that is not a full circle on the top left corner. When I skimmed through the supplementary figures, I noticed a similar issue with Supplementary Figure 6. On panel a)-d) there seems to be like a white box on the top left corner of the scatter plots. Please show all the data points. Please check all similar plots.
Response: Thank you for the correction. Indeed, there were two white boxes that have already been removed. We have checked all similar plots.

c) Please add a locuszoom plot of trophoblast-interacting mQTL in the locus by plotting the interaction p-values to Figure 6. Also, it rather seems that there are two different genetic variants in the BIP GWAS and imQTL analysis (Figure 6). Could you please elaborate more how you reached the suggestion of causal association between placental DNAm and BIP at this CpG site - "Interestingly, rs11637581 is not directly associated with BIP in the original GWAS, suggesting a possible causal association between placental DNAm and BIP at this CpG site"? What was the p-value from the SMR analysis? Add it to the main text too. Also, the lead SNP in the BIP GWAS is below the genome-wide significance threshold (p = 5e-8). Were there any filters based on the p-values in the GWAS analysis similar to filtering of imQTLs for SMR?
Response: On the one hand, we have not added a locuszoom in this case because we did not run colocalization with interaction mQTLs, and adding it could be misleading, and interpreted as a claim of colocalization between signals. However, we want to make it clear that the SNP involved in the imQTL and the one highlighted in the GWAS locuszoom are the same.

On the other hand, we have added the SMR p-value to the main text as follows:
"*We found an imQTL involving rs11637580 and cg27130493 (Figure 6) in which cg27130493 appeared to be associated with BIP and changed its DNAm levels as a function of TB proportion, in a rs11637580 genotype-dependent manner (Bonferroni PSMR = 0.015).*"

Our mention of causality is related to the fact that Mendelian Randomization originally prioritizes as causal the associations of SNPs that exclusively relate to the outcome through the exposure, excluding those that show direct effects on the outcome.

Finally, regarding the GWAS data, we have followed the advice of the SMR developers and kept all the SNPs. Please see the tutorial at the SMR portal for more details.

2. Storey's pi1 value: thank you for adding the Supplementary Figure 3 and finding a solution to overcome the issue with very small p-values (STB-imQTLS and TB-imQTLs comparison).
Response: Thank you.