

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

Received: 14 March 2023

Accepted: 26 February 2025

Published online: 14 March 2025

 Check for updates

A list of authors and their affiliations appears at the end of the paper

Increasing evidence supports the role of the placenta in neurodevelopment and in the onset of neuropsychiatric disorders. Recently, mQTL and iQTL maps have proven useful in understanding relationships between SNPs and GWAS that are not captured by eQTL. In this context, we propose that part of the genetic predisposition to complex neuropsychiatric disorders acts through placental DNA methylation. We construct a public placental *cis*-mQTL database including 214,830 CpG sites calculated in 368 fetal placenta DNA samples from the INMA project, and run cell type-, gestational age- and sex-imQTL models. We combine these data with summary statistics of GWAS on ten neuropsychiatric disorders using summary-based Mendelian randomization and colocalization. We also evaluate the influence of identified DNA methylation sites on placental gene expression in the RICHS cohort. We find that placental *cis*-mQTLs are enriched in placenta-specific active chromatin regions, and establish that part of the genetic burden for schizophrenia, bipolar disorder, and major depressive disorder confers risk through placental DNA methylation. The potential causality of several of the observed associations is reinforced by secondary association signals identified in conditional analyses, the involvement of cell type-imQTLs, and the correlation of identified DNA methylation sites with the expression levels of relevant genes in the placenta.

The impact of the intrauterine environment on development and health, both fetal and long-term, has been known for many years¹. In this line, the developmental origins of health and disease (DOHaD) hypothesis postulates that perinatal and early life environments can impact fetal and later life health². Prenatal stress affects the quality of the intrauterine environment, and is highly related to cardiovascular and metabolic, as well as behavioral and neurodevelopmental disorders³. Thus, the placenta is an ephemeral fetal organ that is uniquely situated to evaluate prenatal exposures in the context of DOHaD, because it manages the transport of nutrients, oxygen, waste,

and endocrine signals between mother and fetus⁴. 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 traits⁵. 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 fetus⁶. Hence, the placenta has been described as the third brain that links the fetal brain with the mature maternal brain, thus becoming the cornerstone to understanding the prenatal environmental effects on neurodevelopment⁷,

✉ e-mail: nora.fernandez@ehu.eus

and potentially, on the appearance of neurodevelopmental and neuropsychiatric disorders later in life.

Particularly in the case of schizophrenia (SCZ), many susceptibility genes have been identified, demonstrating a remarkable genetic basis, but a considerable amount of research suggests that environmental factors may also play a role⁸. The neurodevelopmental hypothesis of SCZ was proposed by Daniel Weinberger in 1987 and has been reinforced in the last decades, with increasing evidence of neurodevelopmental abnormalities contributing to the pathophysiology of the disease^{9–11}. The central argument of this hypothesis states that abnormal fetal neurodevelopment creates a vulnerability to developing SCZ later in life. 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, pre-eclampsia, birth asphyxia, and uterine rupture are associated with SCZ in offspring^{12,13}. Specifically, MIA occurs when inflammatory markers rise above the normal range in pregnancy as a result of maternal inflammation and can be caused by psychosocial stress, infection, or other factors¹⁴.

Besides, there are many studies proposing that the intrauterine environment alters the placental function through epigenetic mechanisms, such as DNA methylation (DNAm)¹⁵. It is very important to note that, while DNAm is bimodally distributed in most tissues and cell types, it follows a trimodal distribution in the very specific case of the placenta, due to its high content of both partially methylated domains and CpG positions with intermediate methylation levels¹⁶. Although a very recent study has pointed out that DNAm of healthy donors could be highly cell type-specific, with the modest impact of either genetics or other factors¹⁷, it is worth mentioning that DNAm has been considered as a bridge between the environment and the genome that, at the same time, is under the control of both environmental and genetic factors. The peculiarity of the placenta, an ephemeral organ that connects two organisms, together with its very particular methylome that is remarkably enriched in CpG sites with intermediate methylation levels¹⁶, leads us to speculate that the genome-environment interaction and its impact on DNAm is very likely unique in the placenta and deserves further investigation.

In the past several years, different works have demonstrated that placental DNAm is sensitive to environmental factors surrounding the gestation. In 2021, the Pregnancy and Childhood Epigenetics (PACE) consortium conducted a meta-analysis on 1700 placental samples and found a placenta-specific DNAm signature of maternal smoking during pregnancy, with differentially methylated CpG sites located in active regions of the placental epigenome, and close to genes involved in the regulation of inflammatory activity, growth factor signaling, and response to environmental stressors¹⁸. One year later, the same consortium reported in the largest placental DNAm study conducted to date that maternal pre-pregnancy body mass index also impacts the placental methylome, specifically at CpG sites located close to obesity-related genes and altogether, enriched in oxidative stress and cancer pathways¹⁹.

Noting the impact of the genetic background on placental DNAm, several studies have reported a number of methylation quantitative trait loci (mQTL) in the placenta and have used them in different downstream applications^{20–23}. Among others, Tekola-Ayele and collaborators showed candidate functional pathways that underpin the genetic regulation of birth weight via placental epigenetic and transcriptomic mechanisms²². However, the catalog of mQTLs they provided was limited to birth weight-related genomic loci. More recently, Casazza et al. reported nearly 50,000 placental *cis*-mQTLs and 2489 sex-specific placental *cis*-mQTLs²³. They found out that placental mQTLs were enriched in genome-wide association study (GWAS) loci for both growth- and immune-related traits, and that male- and female-specific mQTLs were more abundant than non-sex-specific ones. Another remarkable discovery is that they found a modest enrichment

of placental mQTLs in proportion to the calculated SNP heritability in the case of neuropsychiatric disorders, compared to immune- and growth-related traits. However, they did not rule out the possibility that a part of the genetic susceptibility of neuropsychiatric disorders is conferred through the modification of the placental methylome, and if this were the case, it would be relevant not only in terms of etio-pathology, but also from a clinical viewpoint, due to the importance of localizing therapeutic targets in the right tissues, contexts and developmental stages.

In cross-tissue analyses, mQTLs have been highlighted as powerful instruments revealing molecular links to traits otherwise missed by expression quantitative trait loci (eQTL). In fact, recent work by Oliva et al. reported that trait-associated variants are more likely to result in detectable changes in DNAm rather than gene expression, highlighting the relevance of multi-tissue mQTL maps²⁴. On the other hand, in 2021, cell type-interacting eQTLs were defined as proxies of cell type-specific eQTLs and were found not to be covered by standard eQTLs, while they appeared to be highly valuable for gaining a mechanistic understanding of complex trait associations²⁵. Taken together, placenta-specific mQTLs, as well as placenta cell type-interacting mQTLs (imQTLs) can be useful tools in the search for the etiology of genetic associations. In this sense, it is of utmost importance to make this type of resource public and available to the scientific community.

Taking into consideration (i) the increasing evidence supporting the role of the placenta in neurodevelopment and potentially, in the onset of neuropsychiatric traits and disorders, (ii) the peculiarity of the placental methylome, and (iii) the potential of multi-tissue mQTL and imQTL maps to clarify the etiology of complex traits, we proposed that part of the genetic predisposition to complex neuropsychiatric disorders acts through the placental methylome. Thus, we constructed a publicly available placental *cis*-mQTL database calculated in 368 fetal placenta samples, investigated cell type-, gestational age (GA)- and sex-imQTLs, and integrated all these data with summary statistics of the largest GWAS on ten neuropsychiatric disorders, using summary-based Mendelian Randomization (SMR) and colocalization approaches. Finally, we evaluated the functional role of the identified DNAm sites on placental gene expression. We found that placental *cis*-mQTLs are enriched in placental active genomic regions, and are useful to map the etiology of neuropsychiatric disorders to prenatal stages. Specifically, part of the bipolar disorder (BIP), major depressive disorder (MDD), and SCZ genetic risk could act through placental DNAm.

Results

Placental *cis*-mQTL characterization

We discovered *cis*-mQTLs for 214,830 CpG sites (mSites) with False Discovery Rate (FDR) < 0.05 (permuted mQTL database) in 368 fetal placenta DNA samples from the Gipuzkoa, Sabadell, and València cohorts of the Infancia y Medio Ambiente (INMA) project²⁶. Briefly, placental DNAm was modeled as a function of genotype in ± 0.5 Mb windows, with fetal sex, five genotype principal components (PC), 18 DNAm PCs, and Planet-estimated cell types²⁷ as covariates (See Methods section for more details). The phenotypic information of the donor mothers is summarized in Table 1. In the vast majority of the *cis*-mQTLs, the SNP (mVariant) and the mSite were located relatively close to each other, with a median absolute distance of 13.87 kb, indicating that genetically modulated DNAm is typically close to the implicated regulatory variant, as previously described by other authors²⁸ (Fig. 1a). The distribution of mQTLs across chromosomes tended to be in line with chromosome length (Fig. 1b). Additionally, the Haplotype Reference Consortium (HRC)²⁹ reference alleles in placental *cis*-mQTLs presented both positive and negative effect directions (Fig. 1c), as expected.

In general, mSites were depleted from CpG islands and promoters ($P < 2.2 \times 10^{-16}$), and enriched within gene bodies ($P < 2.2 \times 10^{-16}$) and genomic features with intermediate methylation values such as open

Table 1 | Distribution of maternal smoking during pregnancy, demographic variables, birth outcomes, and covariates of the INMA cohort

Variable	Category	INMA (n = 368)		
		N	Mean ± SD or %	N missing
Maternal smoking during pregnancy	Yes	52	14.13%	5
	No	311	84.51%	
Maternal age (continuous)	Mean (SD)	368	30.85 ± 3.95	0
Parity	0	208	56.52%	0
	≥1	160	43.47%	
Maternal education	Primary or without education	68	18.47%	0
	Secondary	172	46.73%	
	University	128	34.78%	
Birth weight (g)	Mean (SD)	368	3278 ± 440.55	0
Gestational age (weeks)	Mean (SD)	368	39.71 ± 1.30	0
Preterm birth (<37 weeks)	Yes	10	2.71%	0
	No	358	97.28%	
Ancestry	White	347	94.29%	5
	Non-white	16	4.34%	
Sex of child	Female	189	51.35%	0
	Male	179	48.64%	
Socioeconomic status	I + II	130	35.32%	0
	III	100	27.17%	
	IV + V	138	37.50%	

SD standard deviation

sea, and CpG island shelf and shore regions ($P < 2.2 \times 10^{-16}$) (Fig. 1d–f, and Supplementary Data 1). They were also enriched in the HLA region with 2896 out of the total 214,830 mSites of the permuted database ($P = 1.45 \times 10^{-10}$). Using eFORGE^{30–32}, we were able to detect enrichment in DNase I hotspots and H3K4me1 broadPeaks in different tissues, the fetal placenta being the one showing the strongest signal (Fig. 1g, Supplementary Fig. 1 and Supplementary Data 2). DNase I hotspots mark accessible chromatin while H3K4me1 broadPeaks appear in enhancer regions. Afterward, we annotated the mSites to their closest genes, performed a gene-set enrichment analysis with the Disease Ontology database³³, and obtained 219 enriched gene sets, including neuropathy (Benjamini–Hochberg $P = 1.88 \times 10^{-9}$), mood disorder (Benjamini–Hochberg $P = 1.22 \times 10^{-7}$), demyelinating disease (Benjamini–Hochberg $P = 2.48 \times 10^{-6}$), and BIP (Benjamini–Hochberg $P = 6.96 \times 10^{-6}$) (Supplementary Data 3 and Supplementary Fig. 2).

Apart from the permuted *cis*-mQTLs, in which correction for multiple correlated variants is performed via a beta approximation, nominal and conditional *cis*-mQTLs were also calculated in our 368 placenta DNA samples. Briefly, the nominal *cis*-mQTL approach reports all the SNP–CpG combinations while the conditional method uses the permuted QTLs to perform a stepwise regression procedure and map conditionally independent *cis*-QTLs (see “Methods” for more details). The three complete placental *cis*-mQTL databases are publicly available online at the following address: https://irlab.shinyapps.io/shiny_mqtl_placenta/.

Placental cell type-, gestational age- and sex-imQTLs

The most abundant cells in the fetal placenta are syncytiotrophoblasts (STB). During gestation, undifferentiated trophoblasts (TB) change into fully differentiated STB, a continuous, multinucleated, and specialized layer of epithelial cells³⁴. We estimated the cell type proportions of our samples with the reference-based method Planet²⁷. As

expected, the estimated content of STB in our placental samples was negatively correlated with the estimated proportion of TB ($P < 2.2 \times 10^{-16}$, $R = -0.89$) and positively correlated with gestational age (GA) ($P = 4.5 \times 10^{-5}$, $R = 0.21$) (Fig. 2a, b). We used the interaction modality of TensorQTL with the eigenMT^{35,36} correction to calculate STB-, TB- and GA-imQTLs, as proxies for STB-, TB- and GA-specific *cis*-mQTLs, and obtained 38 and 1 STB- and TB-imQTLs, respectively, with no significant GA-imQTL at a significance threshold of eigenMT FDR = 0.05 (Supplementary Data 4). The higher amount of STB-imQTLs revealed a higher statistical power for the most abundant cell type, as previously pointed out by Kim–Hellmuth et al.²⁵.

The only significant TB-imQTL was also a STB-imQTL. We also estimated sharing between STB- and TB- imQTLs with Storey’s pi1 method³⁷. Briefly, we took the eigenMT FDR < 0.05 STB-imQTLs and calculated pi1 in the nominal *P*-values of these SNP–CpG pairs in the unfiltered TB- and GA-imQTL sets. STB–TB sharing was estimated to be 0.68, while STB–GA sharing resulted in a pi1 = 0.51 (Supplementary Fig. 3). 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 $P_{\text{nominal}} < 5 \times 10^{-8}$. We found an overlap of 12 imSites between the two cell types, and no overlap between STB- and GA-imSites (Fig. 2c). 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 ($P = 4.1 \times 10^{-10}$ and $R = -0.99$), as in the example in Fig. 2d–f. Of note, the imQTL in Fig. 2 is the only eigenMT-significant TB-imQTL (FDR = 0.05). The moderate sharing between STB- and GA-imQTLs revealed by Storey’s pi1 suggests that the effect of GA on the genetic regulation of placental DNAm is at least partially 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 but obtained no significant imQTLs after eigenMT FDR adjustment.

Multi-omics approaches to unravel the placental origin of neuropsychiatric traits and disorders

We selected the largest GWAS available for each of the following traits: attention deficit and hyperactivity disorder (ADHD)³⁸, aggressive behavior in children (AGR)³⁹, autism spectrum disorder (ASD)⁴⁰, BIP⁴¹, internalizing problems (INT)⁴², MDD⁴³, obsessive-compulsive disorder (OCD)⁴⁴, panic disorder (PD)⁴⁵, suicidal attempt (SA)⁴⁶, and SCZ⁴⁷ (Table 2). Most of the studies were carried out in the context of either the Psychiatric Genomics Consortium (PGC)⁴⁸ or the Early Genetics and Lifecourse Epidemiology (EAGLE)⁴⁹ consortium. GWAS sample sizes ranged from 6183 in the case of SA and 500,199 samples in MDD (Table 2).

We searched for placenta DNAm sites pleiotropically associated with neuropsychiatric disorders using the SMR and the heterogeneity in dependent instruments (HEIDI) tests included in the SMR software⁵⁰. Briefly, SMR was originally developed to test for pleiotropic association between the expression level of a gene (or the methylation level of a CpG site) and a complex trait of interest using summary-level data from GWAS and QTL studies. We used the summary statistics of the aforementioned GWAS and the nominal placental *cis*-mQTL database filtered by $P_{\text{nominal}} < 5 \times 10^{-8}$, as suggested by the developers of SMR. Importantly, all 130,736 mSites considered for SMR were FDR significant in the permuted database. No significant hits were found for many of the traits, such as OCD or

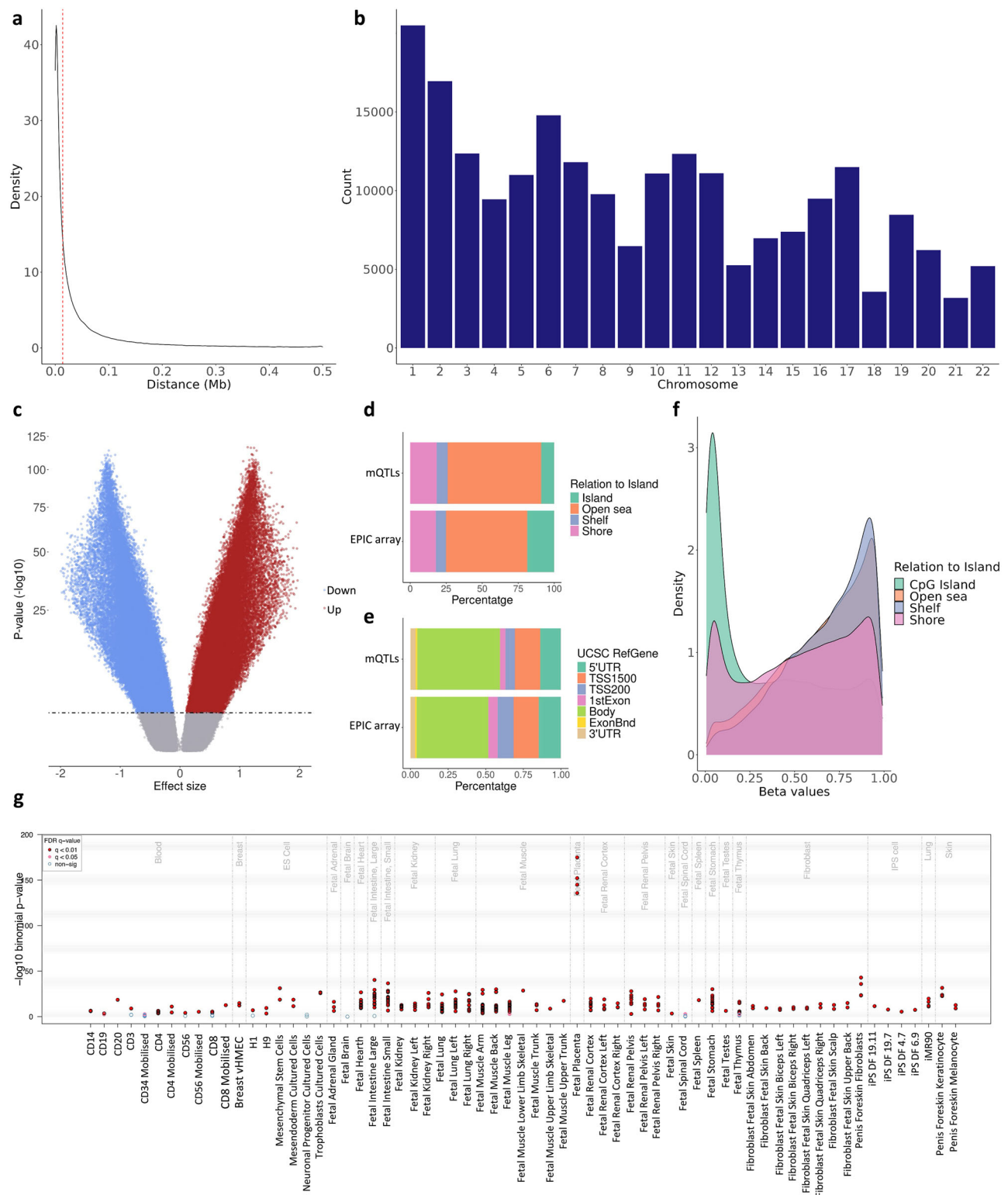


Fig. 1 | Characterization of the placental *cis*-mQTLs from the permuted database. The distance between the SNP-CpG pairs participating in the reported *cis*-mQTLs is displayed as (a) density plot, where the X-axis represents the absolute distance in Mb. The red line represents the absolute median SNP-CpG distance. The distribution of the reported *cis*-mQTLs along the chromosomes is shown in the (b) barplot, where the X-axis represents the autosomal chromosomes. The distribution of the effect sizes from the reported *cis*-mQTLs is pictured in the (c) volcano plot, where the Y-axis illustrates the $-\log_{10}$ beta *P*-value, and the X-axis the effect size. The blue and the red dots represent the significant mQTLs below a threshold of 5×10^{-8} in the nominal database with a negative and positive effect, respectively.

The distribution of the EPIC array probes and the nominal mSites considering the Relation To Island and the UCSC RefGene annotation is displayed in the (d, e) bar plots, respectively. The DNAm beta values of the participating mSites stratified by the Relation To Island annotation are shown in the (f) density plot, where the DNAm values are found on the X-axis. The eFORGE enrichment of DNase I hotspots considering the top 10,000 permuted mSites is shown in the (g) plot. The Y-axis represents the $-\log_{10}$ *P*-value of the binomial enrichment test as suggested by the eFORGE developers, and the X-axis the tissues studied. FDR-corrected *q*-values below 0.01 and 0.05 are represented by red and pink dots, respectively, while blue dots show *q*-values > 0.05.

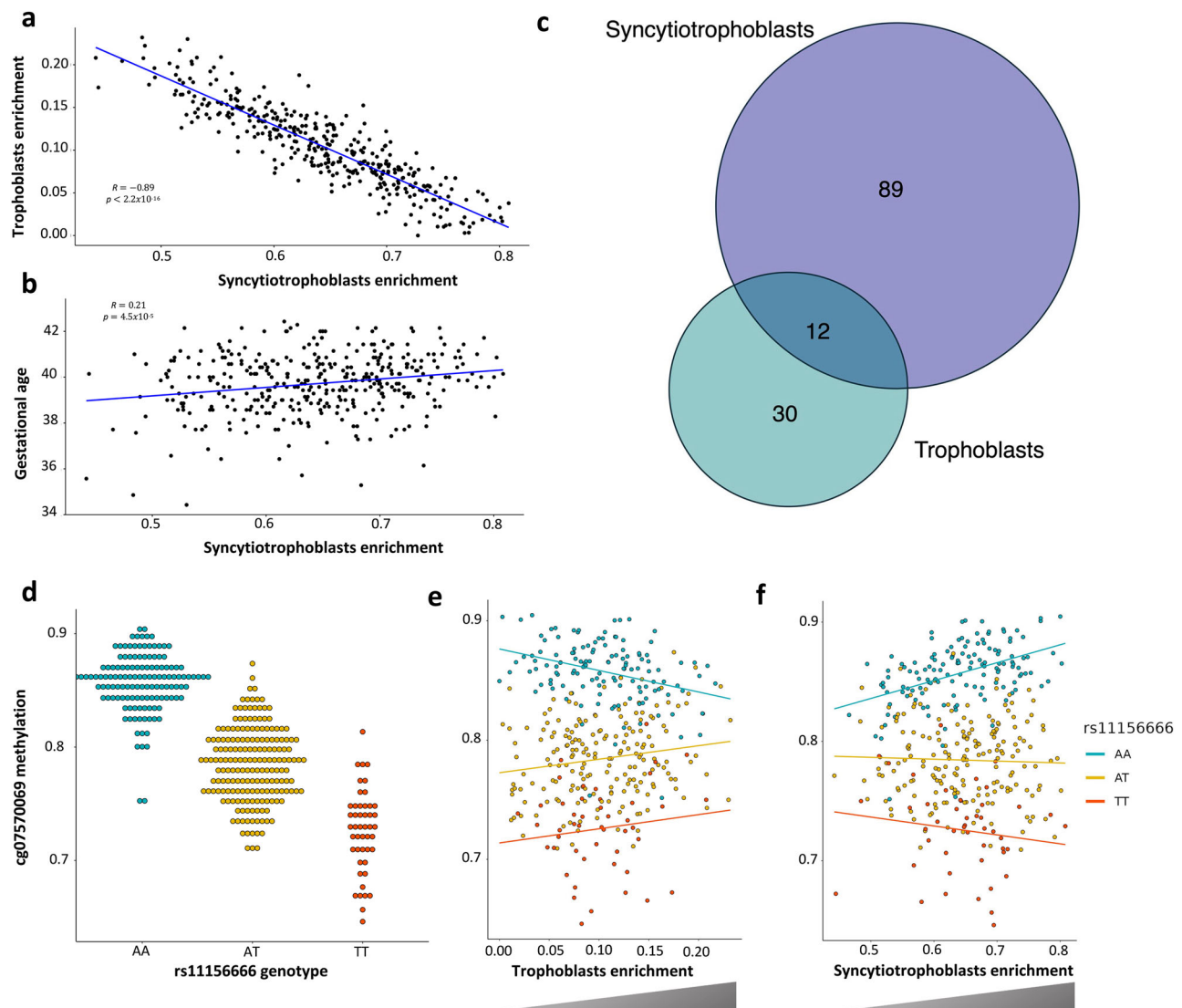


Fig. 2 | STB- and TB-imQTLs. The Pearson correlation between STB and TB cell types ($N = 368$) is shown in (a), the X-axis representing the STB proportion in the sample set, and the Y-axis showing the estimated TB proportion. The Pearson correlation between STB and GA ($N = 368$) is shown in (b), the X-axis representing the STB proportion in the sample set, and the Y-axis showing the GA. The intersection between STB- and TB-imQTLs is represented in the (c) Venn diagram. The standard cg07570069-rs7150262 mQTL, as well as the TB- and STB-imQTLs, are

displayed in the (d–f) dot plots, respectively. In all three, the Y-axes represent the cg07570069 DNAm beta values, ranging from 0 to 1. In d the X-axis displays the genotype of rs7150262, while in (e, f), the X-axes show the TB and STB proportions, respectively. The P -values of the (i)mQTLs were the following: $P_{\text{nominal}} = 1365 \times 10^{-97}$, eigenMT FDR = 2.362×10^{-2} , and eigenMT FDR = 1.579×10^{-4} for the standard mQTL, and the TB- and STB-imQTLs, respectively, obtained by linear regression calculations in TensorQTL.

SA. In turn, SMR hits (Bonferroni $P_{\text{SMR}} < 0.05$ and $P_{\text{HEIDI}} > 0.05$) were identified for ADHD ($n = 1$), ASD ($n = 1$), BIP ($n = 30$), MDD ($n = 28$), and particularly SCZ ($n = 214$) (Fig. 3 and Supplementary Data 5). Five hits were common to BIP and SCZ, with the same effect direction. Additionally, most of these hits did not overlap with CpGs reported to change as a function of GA⁵¹, suggesting relative stability in the DNAm levels during pregnancy.

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 package⁵² 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 (PP) 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 the 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 ($PP_4 > 0.8$) (Fig. 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 ($PP_4 > 0.8$) (Fig. 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 ($PP_4 > 0.8$) (Fig. 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.

Table 2 | Overview of the neuropsychiatric traits and disorders included in this study

Trait	Abbreviation	PubMed ID	N	N cases	N controls	Original N of SNPs	GWAS significant loci	N SMR hits
Attention deficit and hyperactivity disorder	ADHD	30478444	55,374	20,183	35,191	8,094,094	12	1
Aggression	AGR	26087016	18,988			2,188,528	1	0
Autism spectrum disorder	ASD	30804558	46,351	18,382	27,969	9,112,386	3	1
Bipolar disorder	BIP	34002096	413,466	41,917	371,549	7,608,183	64	30
Internalizing problems	INT	35378236	64,561			5,445,594	0	0
Major depression disorder	MDD	30718901	500,199	170,756	329,443	8,483,301	101	28
Obsessive-compulsive disorder	OCD	28761083	9725	2688	7037	8,409,517	0	0
Panic disorder	PD	31712720	10,240	2,248	7992	10,151,624	0	0
Suicidal attempt	SA	34737426	6183	3288	2895	11,823,118	0	0
Schizophrenia	SCZ	35396580	320,404	76,755	243,649	7,583,660	287	214

To provide more insights into the functional relevance of our findings, we examined the association between placental DNAm and placental gene expression in the CpGs from the SMR and colocalization analyses. We interrogated expression quantitative trait methylation sites (eQTM) from an independent set of 195 fetal placenta samples from the Rhode Island Child Health (RICHS) study⁵³. Briefly, eQTM were calculated using linear models in MatrixEQTL⁵⁴, adjusted by fetal sex, 5 gene expression PCs, and Planet-estimated cell types, considering all the genes in a 0.5 Mb window up and downstream of each CpG. 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 the CpGs identified in the overlap between the two analyses (Supplementary Data 7 and Supplementary Fig. 4).

Interestingly, out of the 26 eQTM-genes related to SCZ, 10 mapped to the HLA region, and among these, 6 were correlated with DNAm levels of more than one CpG, with 91 significant combinations between 6 and 21 unique genes and CpG sites, respectively. In particular, the expression levels of both *TOB2P1* and *ZKSCAN4* were correlated with the DNAm of 21 and 18 CpG sites pleiotropically associated with SCZ, respectively, confirming the high complexity of the locus. Outside the HLA region, two eQTM-genes correlated with the DNAm levels of more than one CpG site, namely *SFMBT1* and *VPS37B* (Supplementary Fig. 5). Specifically, the expression of *SFMBT1* on chromosome 3 was correlated with DNAm at four different contiguous loci that were independently identified in the SMR approach, while *VPS37B* on chromosome 12 correlated with the DNAm levels of two different CpG sites, with the same associated top SNP in SMR. In conclusion, our results suggest some regional pleiotropy, also outside the complex HLA region, that could have a functional impact on the placental expression landscape.

Next, we identified the intersection between SMR, colocalization, and eQTM results for each trait (Supplementary Fig. 4). 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, 6 placental DNAm sites colocalized with 6 SCZ-associated loci on chromosomes 6, 7, 8, 11, 19, and 22, with DNAm levels that correlated with the placental expression of 6 nearby genes (*GLBIL3*, *LETM2*, *NAGA*, *PSMG3*, *SLC6A16*, and *ZSCAN12*) (Supplementary Fig. 6).

The overlapping hit in MDD was cg10318063 (Bonferroni $P_{\text{SMR}} = 0.003$) that colocalized with the MDD-associated locus on chr14:41940872-42476274 (PP4 = 0.976) (Fig. 4). The DNAm levels of this CpG correlated with the expression of *LRFN5* in placenta (FDR = 1.23×10^{-13} and $R = 0.539$). Remarkably, several SNPs in *LRFN5* have been found to be pleiotropically associated with both MDD and chronic pain, although the effector tissue or cell type has not been fully

established⁵⁵. On the other hand, it is also known that this gene is expressed in TB stem cells, and therefore could play a role in placenta⁵⁶. Finally, the CpG identified is located in the promoter region of *LRFN5*. Altogether, these results suggest that the reported association is potentially causal, rather than pleiotropic.

Among the CpGs found in SCZ, cg22959474 was pleiotropically associated with the disorder (Bonferroni $P_{\text{SMR}} = 8.741 \times 10^{-3}$), colocalized with the GWAS peak on chr22:41024556:42765301 (PP4 = 0.881), and DNAm levels of the CpG were correlated with the placental expression of *NAGA* (FDR = 9.501×10^{-3} and $R = -0.259$) (Fig. 5). Interestingly, this gene is known to regulate important neurodevelopmental processes, including the proliferation and differentiation of neural stem cells⁵⁷. In this context, it has been described that *NAGA* confers susceptibility to SCZ by modulating dendritic spine density⁵⁸. It is well known that these structures, located on the dendritic branches of neurons, are important for neural connectivity and plasticity, and that abnormalities in their formation and elimination can lead to neurological diseases such as Alzheimer's disease and SCZ⁵⁸. This would support the neurodevelopmental hypothesis of SCZ¹¹.

Finally, we performed a Reactome gene-set analysis⁵⁹ of the 26 genes at the intersection between the SMR and eQTM results in SCZ. We observed enrichment for 101 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 epigenetic modifiers. Additionally, we observed a remarkable enrichment in immune-related pathways such as human cytomegalovirus (HCMV) infection and overall viral infection, reinforcing the idea of MIA as a link between the placenta and SCZ risk.

Conditional SMR analysis of placental *cis*-mQTLs, BIP, MDD, and SCZ

The developers of the SMR software designed the HEIDI test assuming that a single causal variant in the *cis*-mQTL region affects both DNAm and the trait analyzed⁵⁰. Under the assumption of pleiotropy, when there are multiple causal variants in a region, the pleiotropic signal of one causal variant will be diluted by that of other non-pleiotropic causal variants. We therefore performed a GCTA conditional analysis⁶⁰ conditioning for the top associated *cis*-mQTL in both the GWAS and mQTL data sets. In those cases where a secondary signal was pinpointed by the presence of heterogeneity ($P_{\text{HEIDI}} < 0.05$) in the *cis*-mQTL region, in either the GWAS or the mQTL dataset, we performed another round of conditional analyses conditioning only on the secondary signal in both the GWAS and mQTL datasets, and then reran the SMR and HEIDI test at the top *cis*-mQTL using the estimates of SNP effects from the conditional analyses. We applied this approach to

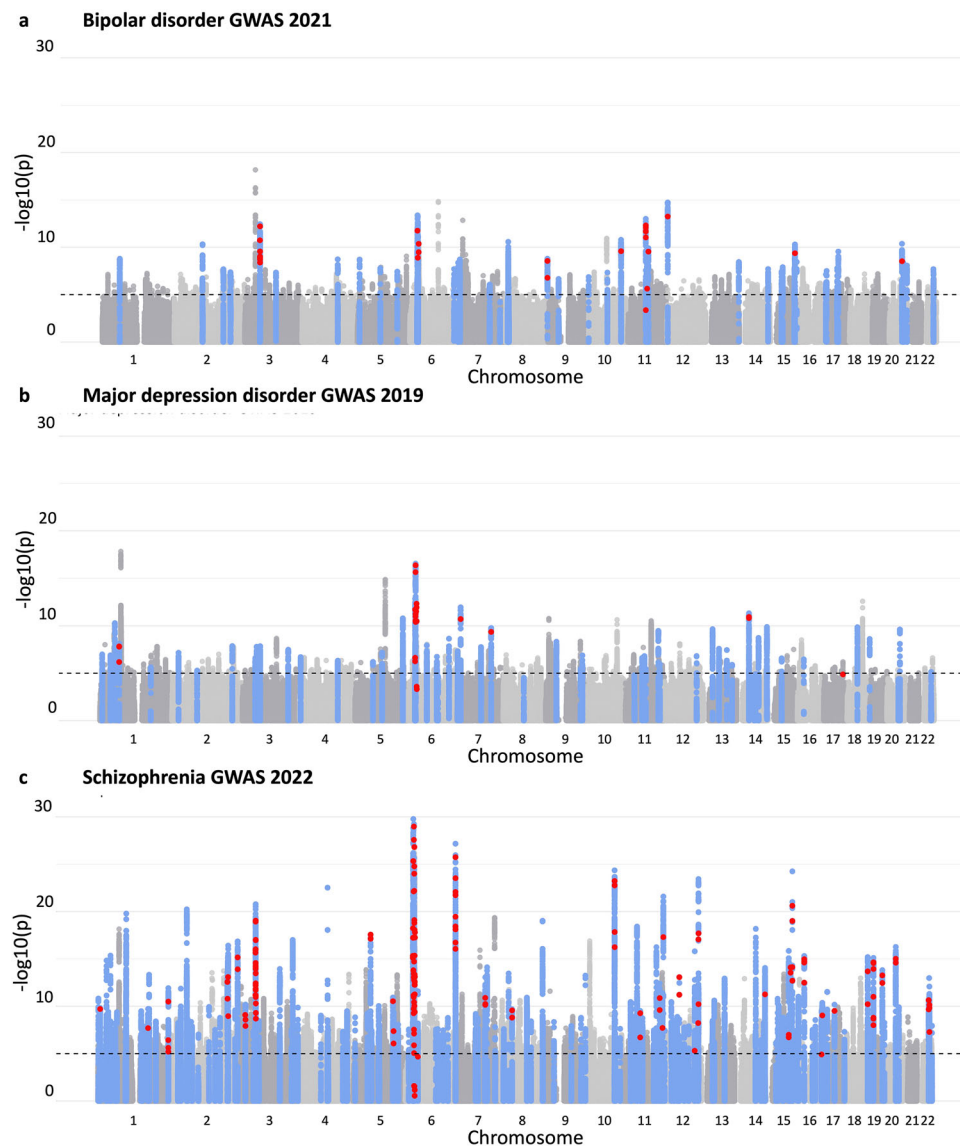


Fig. 3 | Manhattan plot of BIP, MDD, and SCZ GWAS highlighting the SMR and coloc results. The original GWAS from BIP, MDD, and SCZ were plotted in the (a–c) Manhattan plots, respectively. In the Y-axes the original $-\log_{10} P$ -values are displayed with the genome-wide suggestive association threshold of 1×10^{-5} depicted as a dashed line, while chromosomes are shown in the X-axis. The blue dots

represent genomic regions significantly colocizing ($PP4 > 0.8$) with our placental mQTLs, and the red dots are mVariants associated with CpGs that have been shown to pleiotropically associate with either BIP, MDD, or SCZ in the SMR approach (Bonferroni $P_{SMR} < 0.05$ and $P_{HEIDI} > 0.05$). Therefore, the blue dots represent the colocization results, and the red dots show the SMR results.

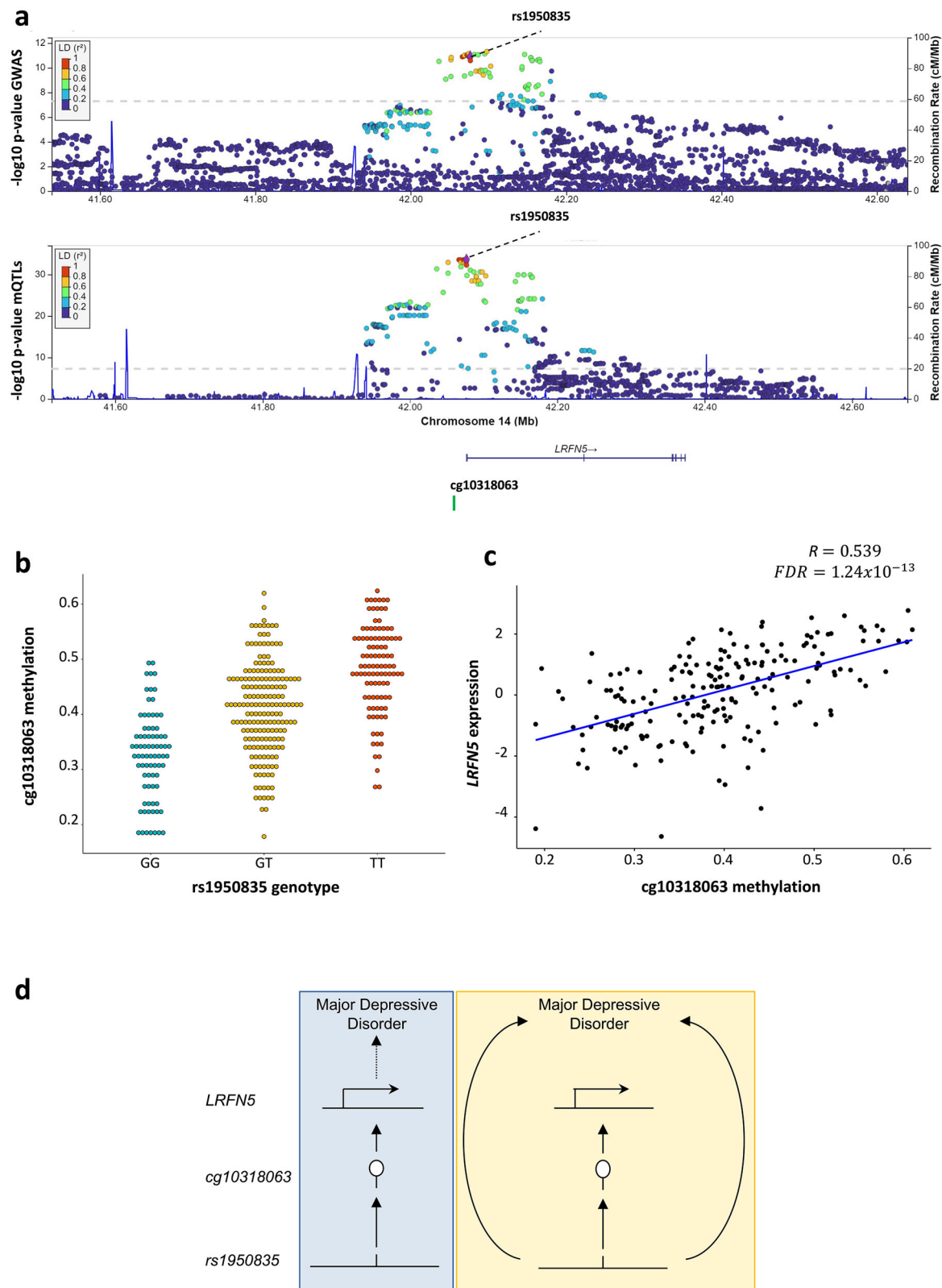
those CpGs that passed the SMR test but failed to pass the HEIDI test. Nevertheless, due to power constraints, only those CpGs with suggestive support for an association with both traits according to *coloc* ($PP3 + PP4 > 0.9$) were considered.

We discovered a secondary hit in SCZ, in which two independently associated SNPs, i.e., rs880638 (the primary signal) and rs1233572 (the secondary signal) had a pleiotropic effect on both cg27314558 and SCZ (Bonferroni $P_{SMR} = 0.015$ and $P_{SMR} = 0.024$, respectively) (Supplementary Fig. 7 and Supplementary Data 5 and 9). Even though we did not find any significant eQTM for this CpG, located in the HLA region, we have to take into account that small RNAs were not included in the eQTM analysis. Having a closer look at this genomic region using the UCSC web browser⁶¹, we found that cg27314558 is surrounded by multiple aminoacyl-transfer RNAs (tRNA), an alanine tRNA (*TRA-AGC6-I*) being the closest (Supplementary Fig. 7). In this context, it is well known that impaired placental amino acid transfer is associated with fetal growth restriction⁶². Additionally, a recent study revealed that placentas of preterm births show a significant negative correlation

with alanine concentration⁶³. Our results, namely the presence of two independent signals pointing to a single DNAm site in the placenta that, moreover, is close to multiple aminoacyl-tRNAs, suggest a causal association of the CpG site with SCZ.

SMR analysis of placental cell type-, GA- and sex-imQTLs in BIP, MDD, and SCZ

We filtered the imQTLs by $P_{nominal} < 5 \times 10^{-8}$, as suggested by the developers of SMR. Then we classified the cell type-imQTLs as positive or negative according to Kim-Hellmuth et al., with the aim of retrieving potentially cell type-specific mQTLs (that is, positive imQTLs) rather than those that more likely reflect interactions with another cell population that is decreasing (that is, negative imQTLs). In summary, positive imQTLs show an increase of the main genotype effect from low to high cell proportions, i.e., the mQTL effect size increases with the cell type abundance. Meanwhile, negative imQTLs show a decrease in the mQTL effect size as a function of the cell type proportion.



We used SMR to combine positive imQTLs with the BIP, MDD, and SCZ GWAS, and obtained one STB-imQTL (related to BIP), and two TB-imQTLs (one each in BIP and MDD) (Fig. 6, Supplementary Fig. 8, and Supplementary Data 10). We found an imQTL involving rs11637580 and cg27130493 (Fig. 6) in which cg27130493 appeared to be associated with BIP and changed its DNAm levels as a function of TB proportion, in an rs11637580 genotype-dependent manner (Bonferroni $P_{\text{SMR}} = 0.015$).

This CpG is located in the gene body of *SMAD3*. The overexpression of this gene in the placenta has been described to activate the ability of TB to form endothelial-like networks, while its defect has been associated with pre-eclampsia⁶⁴. Additionally, it is a well-known target for lithium treatment in BIP⁶⁵. Interestingly, rs11637580 is not directly associated with BIP in the original GWAS, suggesting a possible causal association between placental DNAm and BIP at this CpG site.

Fig. 4 | A CpG pleiotropically associated with MDD. The mVariant rs1950835, highlighted as a purple diamond, with the $-\log_{10}$ P -values of the original MDD GWAS and the placental mQTLs, is represented in the (a) locusZoom plot. The X-axis displays the involved genomic region of chromosome 14 in Mb, showing the distribution of the coding genes in the locus, as well as the location of the cg10318063 mSite ($PP_4 = 0.976$ in the colocalization test). The Y-axis shows the $-\log_{10}$ P -value from the original GWAS and placental mQTLs, and the SNPs are color-coded as a function of the LD with the highlighted SNP considering the 1000G reference panel. The mQTL, rs1950835-cg10318063 ($P_{\text{nominal}} = 2.446 \times 10^{-34}$ of the linear regression in TensorQTL) is plotted in the (b) dotplot, where the Y-axis

represents the beta DNAm values of the mSite, ranging from 0 to 1. The X-axis displays the genotype of the mVariant. The eQTM of cg10318063 is portrayed in the (c) dotplot, where the X-axis represents the DNAm values of the CpG involved, ranging from 0 to 1. The Y-axis displays the expression values of the *LRFN5* gene in the placenta. The hypothesis of the pleiotropical association between the SNP, the DNAm values of the CpG in the placenta, and the gene expression levels of *LRFN5* in the placenta and MDD are schematically represented in (d). The vertical pleiotropy (or causal association) hypothesis is represented with a blue background, and the horizontal pleiotropy hypothesis is highlighted with a yellow background.

Comparison with SMR results for brain mQTLs

In order to ascertain the tissue specificity of our findings, we also performed the same SMR analyses with fetal brain and brain *cis*-mQTLs in BIP, MDD, and SCZ. The fetal brain *cis*-mQTL database was published by Hannon et al. in ref. 66. Briefly, mQTLs were calculated in 166 human fetal brain samples (56–166 days post-conception). The DNAm data were obtained with the Illumina Infinium HumanMethylation 450 K array, and mQTL mapping was performed with the MatrixEQTL R package⁵⁴, considering a *cis*-window of ± 0.05 Mb. Only mQTLs with a $P < 1.5 \times 10^{-9}$ were made available, resulting in 556,513 fetal brain *cis*-mQTLs that were downloaded from the SMR portal. The second brain *cis*-mQTL database was originally from a 2018 publication by Qi et al.⁶⁷, in which *cis*-mQTLs from three different datasets were meta-analyzed: 468 brain cortical region samples from the ROSMAP study⁶⁸, 166 fetal brain samples from the study by Hannon et al.⁶⁶, and 526 frontal cortex region samples⁶⁹ were included, amounting to a final sample size of 1,160 brain samples and nearly 6 M *cis*-mQTLs. As in the case of the fetal brain mQTL database, this dataset was downloaded from the SMR data portal.

In BIP we obtained 13 and 35 significant SMR hits (Bonferroni $P_{\text{SMR}} < 0.05$ and $P_{\text{HEIDI}} > 0.05$) in the fetal brain and brain mQTL datasets, respectively (Supplementary Data 11), and with the same criteria, we obtained 10 and 23 SMR hits in MDD, and 50 and 188 in SCZ. As shown in Fig. 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, also at the level of mVariants (Supplementary Fig. 9). In the case of the fetal brain database, the smaller sample size limits the number of detected mQTLs and could lead to underestimating the proportion of the genetic risk that acts through the fetal brain. When we compared our most reliable hits in the placenta (intersection among SMR, *coloc*, and eQTM results) with the brain SMR results, neither MDD nor SCZ presented any overlapping DNAm sites.

Discussion

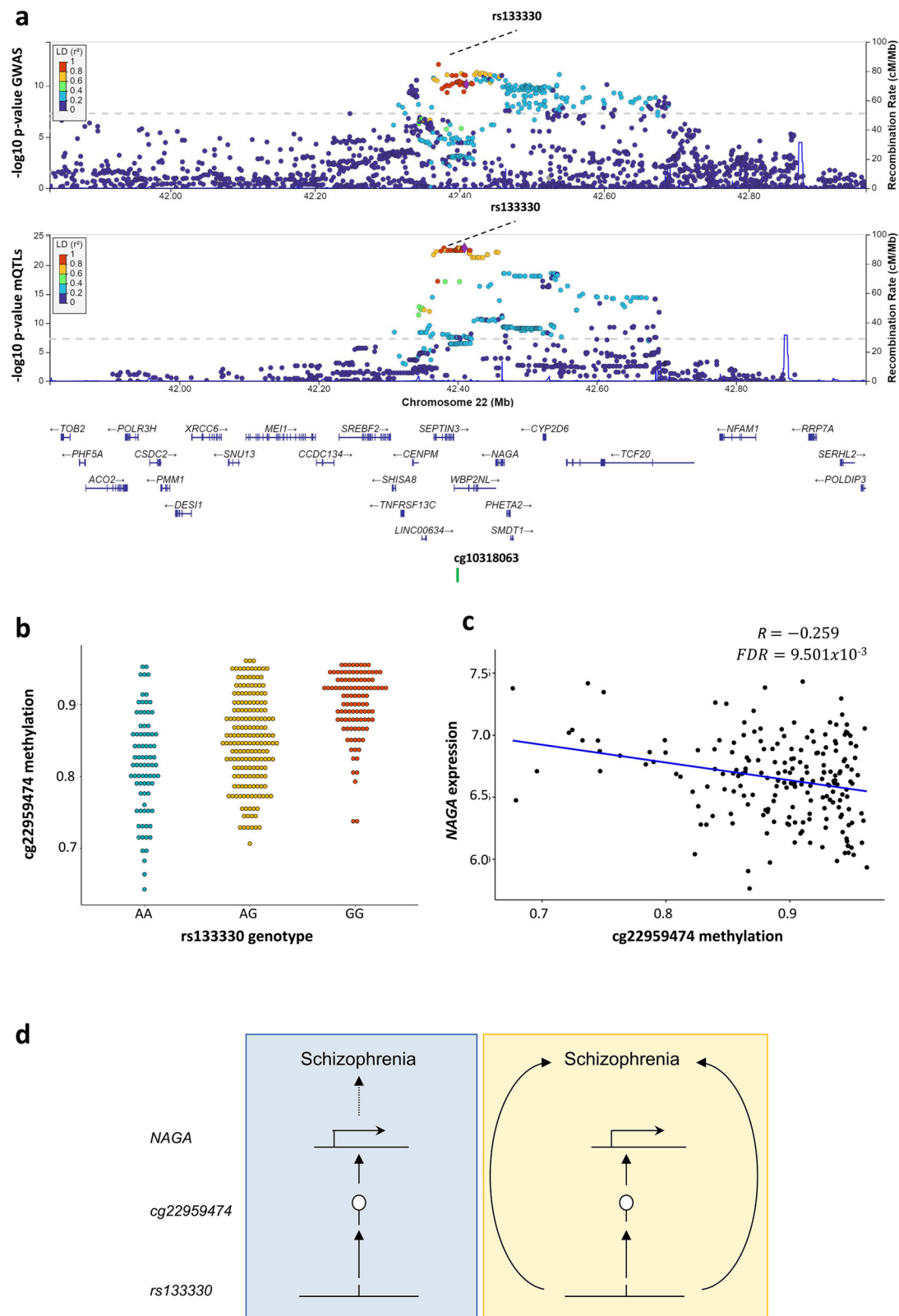
In this study, we have constructed different placental *cis*-mQTL databases with 368 placenta samples from the INMA project. Importantly, we have made all results publicly available both in their raw formats and by means of a user-friendly, shiny-based browser that enables to search for the mQTLs of interest by CpG, SNP, and/or genomic coordinates. We believe that this tool will be useful to the scientific community. In general, the placental *cis*-mQTLs were depleted in regions that are usually hypomethylated, such as promoters and CpG islands, very likely because lower and more stable DNAm values will hardly correlate with the genotype or any other variable. In turn, an enrichment of intermediate DNAm values was observed, usually present in CpG island shelves and shores, as well as in open sea regions. The genotype-regulated placental DNAm seems highly placenta-specific and interestingly, is enriched in placenta-specific active chromatin marks, and in pathways such as neuropathy, mood disorders, and BIP, among others. This, together with the fact that mQTLs have recently been highlighted as powerful instruments to reveal molecular links to traits otherwise missed by eQTL-GWAS colocalization approaches²⁴, encouraged us to conduct a multi-omics study to try to unveil the

placental contribution to the developmental origins of different neuropsychiatric disorders.

Kim-Hellmuth et al. showed that cell type-interacting eQTLs (ieQTLs) are enriched in GWAS signals and can improve GWAS-eQTL matching for the mechanistic understanding of those loci²⁵. In this context, we wondered whether this could also be true for placental mQTLs and calculated STB- and TB-imQTLs. STB were considered for several reasons, (1) they are the most abundant cell type in term placenta⁷⁰, (2) they cover the entire surface of the villous trees in the placenta, and thus are in direct contact with maternal blood³⁴, (3) they orchestrate the complex biomolecular interactions between the fetus and the mother, and (4) they act as an important endocrine organ, producing numerous growth factors and hormones that support and regulate placental and fetal development and growth^{71–74}. Given that TB are STB progenitor cells, and that there is a very strong negative correlation between the proportions of the two cell types in our samples, we decided to calculate TB-imQTLs as well. We observed higher statistical power for the most abundant cell type and therefore, a larger amount of STB-imQTLs compared to TB-imQTLs. However, there was a remarkable sharing and, as expected, the allelic effects in the overlapping imQTLs were negatively correlated between the two cell types. As TB is known to differentiate into STB throughout gestation³⁴, and STB content in term placenta is positively correlated with GA at birth, we wanted to know whether cell type-imQTLs are equivalent to GA-imQTLs. This was not the case, revealing that TB-to-STB differentiation has a notable effect on placental DNAm that is not explained by GA.

Remarkably, placental DNAm is pleiotropically associated with BIP, MDD, and in particular with SCZ, while it seems to barely associate with early onset conditions such as ADHD and ASD. These results could arise from the sample sizes and power constraints of the GWAS available, as well as from the higher heritability of certain disorders, including BIP and SCZ. However, it is important to note that a recent article found 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 annotable⁷⁵. 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 moderate 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. In addition, it is well known that BIP, MDD, and SCZ share a common genetic background and therefore, it is plausible that part of the genetic risk could act through common processes at the same developmental stages⁷⁶.

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 and Supplementary Fig. 10A). 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 ($P < 1.7 \times 10^{-6}$, $r = 0.97$) (Table 2 and Supplementary Fig. 10B, 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 worth mentioning that SCZ is by far the disorder showing the highest ratio between SMR hits and associated GWAS loci, suggesting that

Fig. 5 | A CpG pleiotropically associated with SCZ. The mVariant rs133330, highlighted as a purple diamond, with the $-\log_{10} P$ -values of the original SCZ GWAS and the placental mQTLs, is represented in the (a) locusZoom plot. The X-axis displays the involved genomic region of chromosome 22 in Mb, showing the distribution of the coding genes in the locus, as well as the location of the cg22950474 mSite ($PP_4 = 0.901$ in the colocalization test). The Y-axis shows the $-\log_{10} P$ -value from the original GWAS and placental mQTLs, and the SNPs are color-coded as a function of the LD with the highlighted SNP considering the 1000 G reference panel. The mQTL, rs133330-cg22950474 ($P_{\text{nominal}} = 1.871 \times 10^{-10}$ of the linear regression in TensorQTL) is plotted in the (b) dotplot, where the Y-axis represents

the beta DNAm values of the mSite, ranging from 0 to 1. The X-axis displays the genotype of the mVariant. The eQTM of cg22950474 is portrayed in the (c) dotplot, where the X-axis represents the DNAm values of the CpG involved, ranging from 0 to 1. The Y-axis displays the expression values of the *NAGA* gene in the placenta. The hypothesis of the pleiotropical association between the SNP, the DNAm values of the CpG in the placenta, the gene expression levels of *NAGA* in the placenta, and SCZ are schematically represented in (d). The vertical pleiotropy (or causal association) hypothesis is represented with a blue background, and the horizontal pleiotropy hypothesis is highlighted with a yellow background.

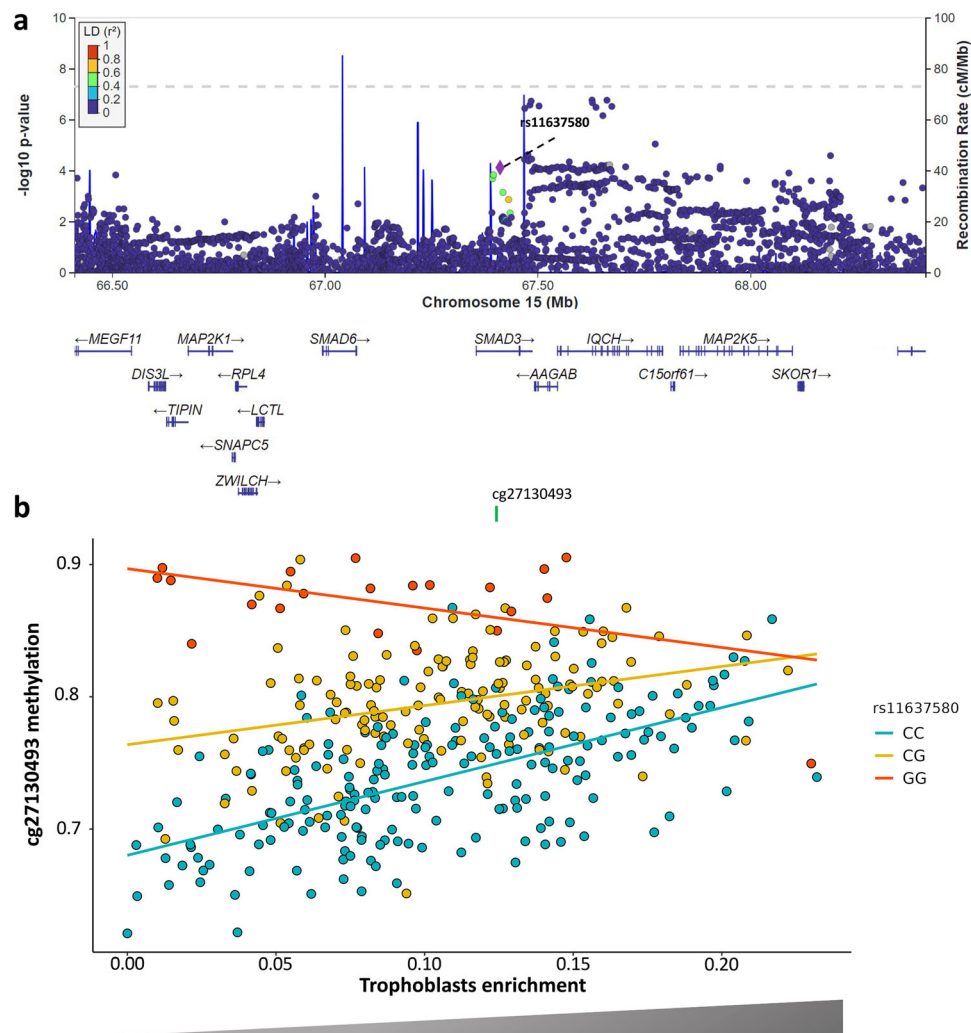


Fig. 6 | TB-imQTL pleiotropically associated with BIP. The imVariant rs11637580, highlighted as a purple diamond, is shown in the original BIP GWAS in the (a) locusZoom plot. The X-axis displays the involved genomic region in chromosome 15 in Mb, showing the distribution of the coding genes in the locus, as well as the location of the imSite cg27130493. The Y-axis shows the $-\log_{10} P$ -value from the original GWAS, and the SNPs are color-coded as a function of LD with the

highlighted SNP considering the 1000 G reference panel. The TB-imQTL ($P_{\text{nominal}} = 1.671 \times 10^{-8}$ of the linear regression in TensorQTL) is pictured in the (b) dotplot. The X-axis represents the cg27130493 DNAm beta values and the Y-axis the TB proportion, both ranging from 0 to 1. The genotype of the rs11637580 imVariant is color-coded as indicated in the legend.

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 Fig. 10D, together with the Venn diagrams showing the overlapping results (Supplementary Fig. 10E, F). In summary, our results reveal an at least moderate correlation between the number of SMR hits and the GWAS sample size, which 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.

It is possible to gain insight into the pathogenesis of complex disorders by defining the environmental, biological, and temporal context in which genes increase disease susceptibility. In 2018, Ursini et al. discovered that when early life complications are present, the polygenic risk score (PRS)-explained risk of SCZ is more than five times higher than when they are absent⁷⁷. SCZ loci that interact with early life complications are not only highly expressed in the placenta, but also differentially expressed between complicated and normal pregnancies

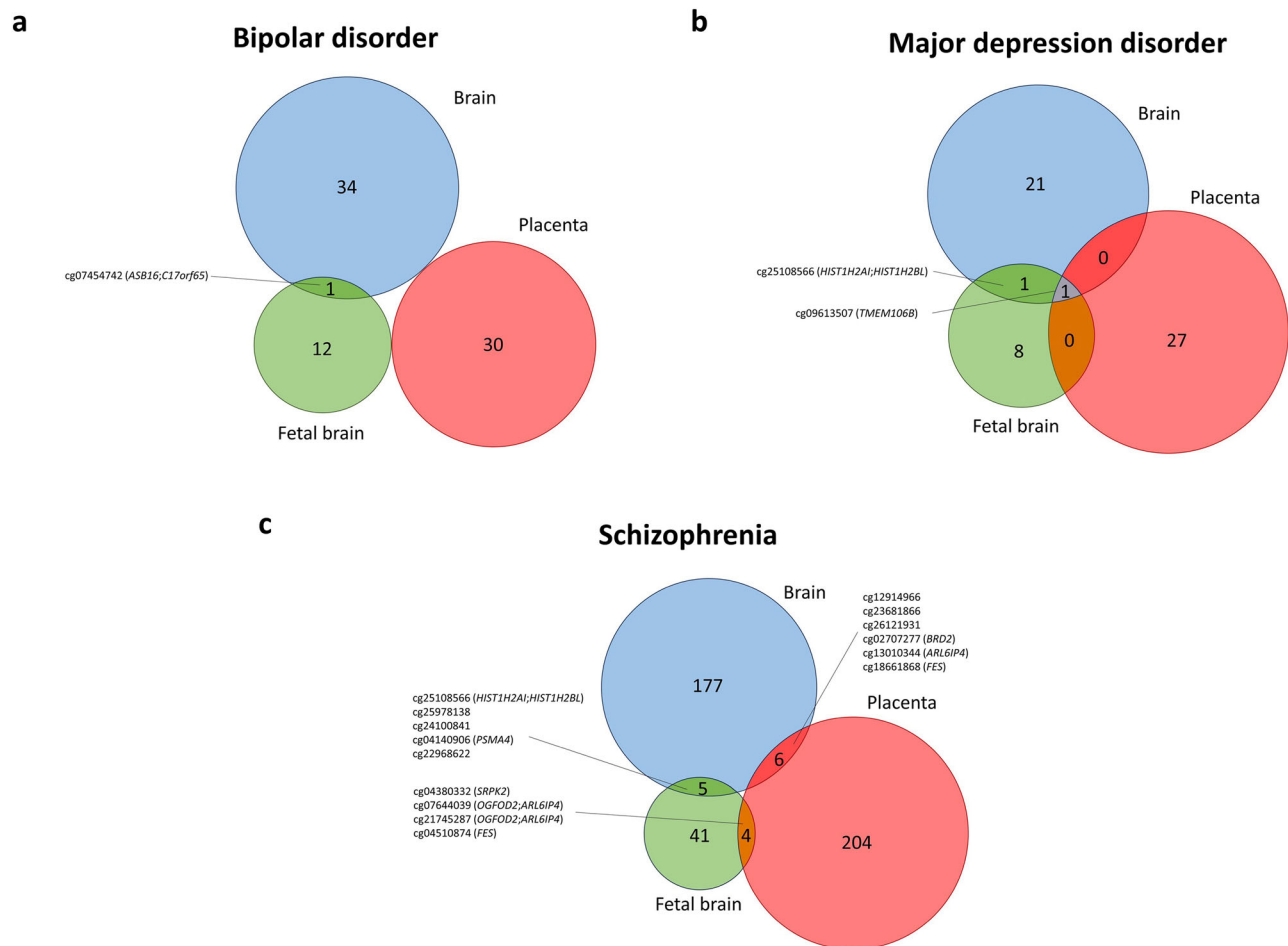


Fig. 7 | Overlap between the SMR results in brain, placenta, and fetal brain, in BIP, MDD, and SCZ. The overlap between the mSites pleiotropically associated between the three tissues and BIP, MDD, and SCZ are represented in the (a–c) Venn

diagrams, respectively. Overlapping mSites are shown, and also the closest gene from the Illumina annotation file (between brackets).

and enriched in response-to-stress pathways. Remarkably, three out of the genes whose expression correlates with DNAm sites associated with SCZ according to the different approaches presented here, were among the 248 placenta-enriched SCZ genes described by Ursini (i.e., *NAGA*, *PSMG3*, and *SFMBT1*). Of note, among these, *NAGA* has been described to confer SCZ risk by modulating the dendritic spine density in neurodevelopment⁵⁸. Altogether, we propose that these genes could not only mediate SCZ risk through placental expression, but rather through DNAm-regulated placental expression patterns.

More recently, the placenta-specific SCZ-PRS has also been shown to interact with both brain volume and cognitive function, suggesting particular neurodevelopmental trajectories in the path toward SCZ⁷⁸. Especially in this work, Ursini and colleagues also studied the enrichment in the placenta of genes involved in the risk of 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 the placenta, although its placenta-specific PRS did not interact with either brain volume nor cognition. The placental enrichment observed in BIP makes sense since SCZ and BIP present the highest genetic correlation compared to any other pair of psychiatric traits⁷⁹. This is in line with the convergence between the CpGs associated with both BIP and SCZ found in this study. Nevertheless, there are remarkable differences in outcome between the two disorders, and SCZ patients have been described to be more prone to suffer from severe cognitive impairment than BIP patients⁸⁰. In conclusion,

placental DNAm could be important in BIP although through trajectories that are different from those characterized by Ursini and colleagues in SCZ. In any case, we believe that our findings support a hypothesis according to which placental DNAm could translate both the genetic basis and the environmental milieu into fetal genetic programs that could eventually result in impaired neurodevelopmental trajectories leading to SCZ, and maybe, also to other neuropsychiatric disorders. This hypothesis will need to be tested with additional research.

Regarding the genes affected by placental DNAm pleiotropically associated with SCZ according to SMR, we want to highlight that they were enriched mostly in epigenetic regulation pathways and immune-related terms, including HCMV infection and late events, as well as assembly of the human immunodeficiency virus (HIV) virion, among others. On the one hand, the multiple epigenetic pathways involved suggest that SCZ risk alleles could cause deep epigenetic changes in the placenta as a result of DNAm modifications in close proximity to the histone-coding genes *H2BC6* and *H3C4*. These genes have been previously related to recurrent spontaneous abortion and neurulation, respectively^{57,58}. On the other hand, those histones, together with *VPS37B*, were some of the most relevant genes contributing to the enrichment of 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 genome⁸¹. 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 others⁸². 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 to the activation of the pathogenesis of both Coronavirus and influenza, further reinforcing the idea of the implication of viral infections⁸³.

Altogether, environmental insults that occur in pregnancy and early life, including maternal infections, are hypothesized to program the immune and developmental epigenetic code in the fetus, thereby influencing the risk of suffering from neurodevelopmental disorders later in life^{84–86}. The MIA hypothesis proposes that exposure to a dysregulated maternal immune milieu in utero affects fetal neurodevelopment^{87,88}. Moreover, in humans, maternal factors implicated in MIA are associated with epigenetic modifications in the placenta^{89,90}. The placenta plays a pivotal role in maintaining immune homeostasis in the maternal-fetal interface. However, when a sustained placental inflammatory response occurs due to maternal environmental factors, the offspring can suffer from developmental abnormalities⁹¹.

Apart from the genes leading the aforementioned pathway enrichments, several others are also worthy of mention, as noted by their correlation with multiple CpG sites associated with SCZ. On the one hand, among the genes in the HLA region and despite the high complexity of the locus, *TOB2P1* is the one locally correlated with more SCZ-associated DNAm sites. Additionally, its RNA expression level in the placenta is regulated by the genotype of nearby SNPs⁹². On the other hand, outside the HLA region, we would like to mention *SFMBT1* and *VPS37B*. The former is an epigenetic regulator that is crucial for TB maintenance and placental development^{93,94}. Moreover, it was one of the 139 placenta and SCZ-specific risk genes prioritized in the TWAS by Ursini and colleagues⁸³, and it is near a CpG identified by our group to be associated with pre-pregnancy maternal body mass index¹⁹. The relevance of *VPS37B* is beyond any doubt given its immune function and its involvement in viral infection. In conclusion, all those genes deserve further investigation.

Taken individually, the genes that are most likely to be causally involved in the different neuropsychiatric disorders studied were *LRFN5* in MDD, and *NAGA* in SCZ. In the case of *LRFN5*, several variants in this gene have been found to be pleiotropically associated with both MDD and chronic pain⁵⁵. Additionally, it is well known that *LRFN5* is involved in brain cell communication and is located in a large and complex genomic niche that is highly conserved in mammals⁹⁵. The specific locus structure of this region increases ASD susceptibility in males. 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 memory⁵⁸. 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 patients⁹⁶. Interestingly, in the context of pregnancy, gestational stress leads to significant changes in 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 manner⁹⁷.

The placental DNAm sites that resulted from the SMR analyses confronting the BIP, MDD, and SCZ GWAS, and the STB- and TB-imQTLs are also worthy of mention. For example, we found that cg27130493 is pleiotropically associated with BIP while it changes its

methylation levels as a function of TB proportion, in a rs11637580 genotype-dependent manner. This CpG is located in the gene body of *SMAD3*, a gene with singular functions in the placenta and also in BIP itself^{64,65}. Additionally, this SNP is not directly associated with BIP, increasing the likelihood of causality rather than pleiotropy. The most important limitation of this part of the study is the lack of placental cell type-specific expression data to ascertain whether cg27130493 and other DNAm sites identified in this approach are correlated with the expression of nearby genes in the placenta, in a cell type-specific manner. However, the fact that the CpG sites change not only depending on the genotype of adjacent SNPs, but also as a function of estimated placental cell proportions, increases the probability of the placenta being the effector organ of those genetic associations.

Lastly, we tried to ascertain the tissue-specificity of our findings by comparing placenta SMR hits to those from two brain *cis*-mQTL databases. We found a limited overlap among the tissues, suggesting that the majority of our pleiotropically associated DNAm sites could be relatively specific to the placenta (or at least, not related to brain DNAm). However, it is important to consider that although the SMR approach was executed exactly in the same manner for the three databases, these were quite different from each other. In particular, the fetal brain database was calculated in a limited number of fetal brain samples, in SNP-CpG windows ten times smaller than the ones we used (0.05 Mb vs 0.5 Mb), and DNAm was measured with the Infinium HumanMethylation 450 K array (with approximately half the probes in the Infinium HumanMethylation EPIC array used in the present study), therefore resulting in half a million *cis*-mQTLs compared to the more than 9 million mQTLs that were included in the SMR analysis in our case. This suggests a more limited mapping potential with the fetal brain mQTL database and thus, a possible underestimation of the fetal brain DNAm that is really involved in the disorders studied.

The main limitation of our study is that, as pointed out by Ursini et al. in relation to their own work⁷⁸, the considerable overlap in cell biology between the brain and placenta does not allow to exclude the possibility that part of the pleiotropy observed here is related to a more direct effect in the brain exerted by the same DNAm sites as observed in the placenta. However, we believe that the different pieces of evidence, including the intersection with the placental eQTLs and even the presence of secondary signals, support that a part of the genetic risk of suffering from BIP, MDD and very especially SCZ, could act through placental DNAm at specific genomic loci. Another limitation is that different methylation beadchips have been employed in the INMA and RICHS cohorts, and thus, we are probably underestimating the effects of DNAm over placental gene expression. Finally, at this point, we lack direct experimental support for our results. Further research, especially regarding in vitro validation of the most likely causal candidates, would enormously help supplement our findings.

In conclusion, we find placental *cis*-mQTLs to be highly placenta-specific, with a remarkable enrichment in genomic regions active in placenta and neurodevelopment- and mental health-related pathways. We prove that they are useful to map the etiologic window of neuropsychiatric disorders to prenatal stages and conclude that part of the genetic risk of BIP, MDD, and in particular SCZ, act through placental DNAm at specific genomic loci. In fact, some of the observed associations might be causal rather than pleiotropic due to the presence of secondary association signals in conditional analyses, involvement of cell type-specific imQTLs, and association with the expression levels of relevant genes in the placenta. It is of particular interest that SCZ-associated placental DNAm correlates with the expression of immune- and neurodevelopment-related genes in the placenta, providing further support to the hypothesis of the neurodevelopmental origins of SCZ. 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.

Methods

Ethics

All cohorts obtained ethics approval and informed consent from participants prior to data collection through their Institutional Ethics Boards, in accordance to the criteria set by the Declaration of Helsinki. This particular study did not involve the acquisition of new data and therefore, did not require additional approval by the Ethics Board.

Placental biopsies and DNA extraction

In the INMA²⁶ project, 2506 mother-fetus pairs were followed until birth, and a selection of 397 placentas were collected, representing the three geographical areas involved in the study. Collected placentas were stored at -80°C in a central biobank until processing. Biopsies of approximately 5 cm^3 were obtained from the inner region of the placenta, approximately 1.0–1.5 cm below the fetal membranes, corresponding to the villous parenchyma, and at a distance of $\sim 5\text{ cm}$ from the site of cord insertion. 25 mg of placental tissue was used for DNA extraction, previously rinsed twice for 5 min in 0.8 mL of 0.5X PBS in order to remove traces of maternal blood. Genomic DNA from the placenta was isolated using the DNAeasy[®] Blood and Tissue Kit (Qia-gen, CA, USA). DNA quality was evaluated on a NanoDrop spectrophotometer (Thermo Scientific, Waltham, MA, USA) and additionally, 100 ng of DNA was run on 1.3% agarose gels to confirm that samples did not present visual signs of degradation. Isolated genomic DNA was stored at -20°C until further processing.

Genotype data

Genome-wide genotyping was performed using the Illumina GSA Beadchip at the Human Genotyping Facility (HuGeF), Dept Internal Medicine, Erasmus MC, Rotterdam, the Netherlands, and the Spanish National Genotyping Center, CEGEN, Madrid, Spain. Genotype calling was done using the GeneTrain2.0 algorithm based on HapMap clusters implemented in the GenomeStudio software. Samples were genotyped in four batches.

The quality control of the genotype data from 397 INMA samples and 509,450 genetic variants was performed using the PLINK 1.9 software following the standard recommendations^{98–100}. All plink files were initially processed with Will Rayner's preparation Perl script available from Mark McCarthy's Group as recommended in the documentation from the Michigan Imputation Server¹⁰¹, using the HRC r1.1 2016 reference panel²⁹. Variants with a call rate below 95%, minor allele frequency (MAF) below 1%, or a P -value from the Hardy-Weinberg Equilibrium (HWE) exact test below 1×10^{-6} were removed. Samples with discordant sex, those with average heterozygosity values above or below 4 standard deviations, or with more than 3% missing genotype were filtered out. Identity-by-descent values were calculated with PLINK, and from those sample pairs that showed PI-HAT estimates above 0.18, the sample with a higher proportion of missing genotypes was removed.

The final dataset was imputed with the Michigan Imputation Server using the HRC reference panel, Version r1.1 2016. Before imputation, data was converted into VCF format. Phasing of haplotypes was done with Eagle v2.4¹⁰² and genotype imputation with Minimac4¹⁰³, both implemented in the code by the Michigan Imputation Server. Finally, we removed variants with an imputation quality r^2 below 0.9, a MAF lower than 5%, a HWE P -value below 0.05, and more than two alleles, to avoid SNPs with few or no individuals bearing the minor allele homozygous genotype in our sample set. Only those samples with paired DNAm data were considered in this analysis. The final dataset consisted of 368 samples and 4,171,035 SNPs.

DNA methylation data

DNAm was assessed with the Infinium Methylation EPIC BeadChip from Illumina, following the manufacturer's protocol in the Erasmus Medical Center core facility. Briefly, 750 ng of DNA from 397 placental samples were bisulfite-converted using the EZ 96-DNA methylation kit from Zymo Research, following the manufacturer's standard protocol, and DNAm was measured using the Infinium protocol. Three technical duplicates were included. Samples were randomized taking into account region-of-origin and sex. As the number of samples in each condition was different, a perfect randomization was not possible. However, all the plates had samples from all three geographical areas involved, and an equilibrated number of male and female samples.

The quality control of the DNAm data, including 865,859 DNAm probes, was performed using the PACEAnalysis R package (v.0.1.7)¹⁰⁴. Before the quality control with PACEAnalysis, one sample was discarded because of too many missing values in relevant variables. With the R package, we discarded those samples with a call rate below 95%, sex inconsistencies, intentioned or non-intentioned duplicates, and those contaminated with DNA from another subject or the mother. Only those samples with paired genotype data were considered in this study. Probes with a call rate lower than 95%, in the sex chromosomes, with SNPs (European MAF $\geq 5\%$) and cross-hybridizing potential were excluded from the analysis¹⁰⁵.

The methylation beta values were normalized in different steps. Dye-bias and Noob background correction, implemented in the minfi R package, were applied^{106,107}, followed by normalization of the data with the functional normalization method¹⁰⁸. Then, to correct for the bias of type-2 probe values, the beta-mixture quantile (BMIQ) normalization was applied¹⁰⁹. After that, we explored the clustering of the data through Principal Component Analysis (PCA) and tested the association of the 12 first PCs with the main and the technical variables. Array batch effect was controlled with the ComBat method¹¹⁰. To correct for the possible outliers, we Winsorized the extreme values to the 1% percentile (0.5% on each side), where percentiles were estimated with the empirical beta distribution. Finally, the rank-based inverse normal transformation (RNT) was applied to the beta values, and these estimations were the DNAm values considered for mQTL mapping. The final dataset consisted of 368 samples and 747,486 DNAm probes (CpGs).

Cell type proportions of six populations (STB, TB, nucleated red blood cells, Hofbauer cells, endothelial cells, and stromal cells) were estimated from DNAm using the placenta reference panel from the 3rd trimester implemented in the Planet R package²⁷.

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 TensorQTL¹¹¹. TensorQTL nominal modality performs linear regressions between the genotype and the normalized DNAm RNT values, as implemented in FastQTL¹¹². 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 calculated 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., we kept all DNAm PCs that cumulatively explained 80% of the variance with a maximum number of 20 PCs for subsequent steps²⁸. 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.

The TensorQTL *cis*-region was specified as ± 0.5 Mb from each tested CpG position, consistent with the results of previous studies where the distance between the majority of the *cis*-mVariants and the mSites was ≤ 0.5 Mb^{66,113–115}. In the final nominal *cis*-mQTL database, we included only those probes with at least one *cis*-mQTL at $P_{\text{nominal}} < 5 \times 10^{-8}$. This same regression model was used to build two additional *cis*-mQTL databases; permuted and conditional. The permuted *cis*-mQTL database consists of correcting for multiple correlated variants tested via a beta approximation which models the permutation outcome with a beta distribution as described in Ongen et al. 2016¹¹². This database was presented as the main database and characterized accordingly, but the nominal was used for some downstream analyses. In turn, the conditional database uses the permuted QTLs to perform a stepwise regression procedure to map conditionally independent *cis*-QTLs as described by the GTEx Consortium¹¹⁶.

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 each mSite and mVariant pair, and the mQTLs along the chromosomes, followed by a volcano plot to check the distribution of the negative and positive effect sizes.

Second, using the annotation from IlluminaHumanMethylationEPICanno.ilm10b4.hg19 R package¹¹⁷, we performed several chi-square tests to assess the enrichment and depletion of the UCSC RefGene and Relation to CpG Island annotations. We also depicted density plots of the DNAm values according to the Relation to CpG Island annotation. Moreover, for the top 10,000 mSites, we assessed enrichment and depletion of overlap with tissue-specific and cell type-specific regulatory features including DNase I hypersensitivity sites (DHS), all 15-state chromatin marks, and all five H3 histone marks (i.e., H3K27me3, H3K4me1, H3K4me3, H3K36me3, H3K9me3) from consolidated ROADMAP Epigenomics Mapping Consortium¹¹⁸ using eFORGE v2.0^{30–32}. The enrichment and depletion with each of the three putative functional elements were tested separately and compared to the respective data from the consolidated ROADMAP epigenomics reference panel. Briefly, we selected only the top 10,000 probes following the advice of the developers of eFORGE 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.

Lastly, over-representation analyses of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) gene sets were conducted with MissMethyl R package^{119–122}. MissMethyl performs a hypergeometric test taking into account the bias derived from the differing number of probes per gene and the multiple genes annotated per CpG¹²². Gene set enrichment analyses with the Disease Ontology (DO) database were conducted with the DOSE R package¹²³. With DOSE, we performed the gene set enrichment analysis considering the genes annotated in the IlluminaHumanMethylationEPICanno.ilm10b4.hg19 R package¹¹⁷ as background.

Placental *cis*-imQTL analysis

The cell type-imQTLs were computed with the interaction modality in TensorQTL, consisting of nominal associations for a linear model that includes a genotype per interaction term¹¹¹. 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 ones³⁵, 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. 2021²⁵, 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 *pi1* method³⁶. 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* package¹²⁴ to estimate *pi1* (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 retrieved the imSites with a $P_{\text{nominal}} < 5 \times 10^{-8}$ in STB- and TB-imQTLs to draw a Venn diagram. Finally, with this 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.

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.

Genome-wide association studies

The GWAS summary-statistics used in this analysis were the public largest to date association studies of ADHD³⁸, AGR³⁹, ASD⁴⁰, BIP⁴¹, INT⁴², MDD⁴³, OCD⁴⁴, PD⁴⁵, SA⁴⁶, and SCZ⁴⁷ from EAGLE, Indonesia Schizophrenia Consortium, International Obsessive Compulsive Disorder Foundation Genetics Collaborative (IOCDF-GC), OCD Collaborative Genetics Association Studies (OC GAS), PsychENCODE, Psychiatric Genomics Consortium (PGC), Psychosis Endophenotypes International Consortium, UK Biobank (UKB), SynGO Consortium, and 23andMe, among others.

GWAS summary statistics were harmonized to dbSNP build 155 with the INMA genotype data as a reference. The harmonization steps included: changing rsID to chromosome:position nomenclatures, correcting the effect allele, the effect size, and the allele frequency according to the reference genotypes if applicable, and creating a .ma format file with the summary-statistics as indicated in the SMR pipeline⁵⁰. All locusZoom plots were constructed as suggested on their web page (<http://locuszoom.org/>), using the LD panel of 1000G¹²⁵.

Multi-SNP-based Mendelian randomization analysis

Mendelian Randomization analysis was carried out considering *cis*-mQTL SNPs as the instrumental variables (IVs), CpG methylation as the exposure (X), and the neuropsychiatric traits as the outcome (Y). Multi-SNP-based MR (SMR-multi)¹²⁶ analysis was performed by the SMR software⁵⁰. SMR integrates GWAS and QTL summary statistics to test for pleiotropic associations between quantitative traits, such as methylation or expression, and a complex trait, for instance, a disease. More precisely, SMR-multi includes multiple SNPs at a *cis*-mQTL locus in the SMR test to calculate the causative effect of an exposure on an outcome (b_{xy}). First, SMR-multi selected all SNPs with $P_{\text{nominal}} < 5 \times 10^{-8}$ in the *cis* region (± 0.5 Mb of the CpG). Second, it removed SNPs in very high LD with the top associated SNP (LD $r^2 > 0.9$). Then, the causative effect (b_{xy}) from the exposure on the outcome was estimated at each of the SNPs and combined in a single test using an approximate set-based test accounting for LD among SNPs¹²⁶. Additionally, the HEIDI test was performed. HEIDI uses multiple SNPs in a *cis*-mQTL region to distinguish pleiotropy from linkage. As it is described by Zhu et al.⁵⁰, under the hypothesis of pleiotropy, where DNAm and a trait share the same causal variant, the b_{xy} values calculated for any SNPs in LD with the causal variant are identical. Therefore, testing against the null

hypothesis that there is a single causal variant is equivalent to testing whether there is heterogeneity in the b_{xy} values estimated for the SNPs in the *cis*-mQTL region. For each DNAm probe that passed the genome-wide significance ($P_{\text{nominal}} < 5 \times 10^{-8}$) threshold for the SMR test, HEIDI tested the heterogeneity in the b_{xy} values estimated for multiple SNPs in the *cis*-mQTL region. In this analysis, significant pleiotropic associations between DNAm and neuropsychiatric diseases were selected as those with P_{SMR} -corrected Bonferroni < 0.05 and $P_{\text{HEIDI}} > 0.05$ (not showing heterogeneity).

In the two interaction *cis*-mQTL databases used for SMR, we included probes with at least one *cis*-imQTL at $P_{\text{nominal}} < 5 \times 10^{-8}$. The imQTLs obtained from each model were categorized as positively and negatively correlated with cell type estimates, or as uncertain if this was ambiguous, following Kim Hellmuth et al. 2021²⁵. To categorize the imQTLs into these groups, genotype main effects at low (<25th percentile) vs high (>75th percentile) cell type proportions were compared. imQTLs with positive cell type correlation showed an increase of the main genotype effect from low to high cell proportions (β_{mQTL} low $< \beta_{\text{mQTL}}$ high). imQTLs with negative cell type correlation showed a decrease (β_{mQTL} low $> \beta_{\text{mQTL}}$ high) and the uncertain group contained imQTLs for which the sign flipped between low and high cell type proportions (β_{mQTL} low $\neq \beta_{\text{mQTL}}$ high).

Colocalization analyses

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 et al. 2014, with the R *coloc* package⁵². 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 PP_i for hypothesis H_i .

- H_0 : there exist no causal variants for either trait;
- H_1 : there exists a causal variant for one trait only, GWAS;
- H_2 : there exists a causal variant for one trait only, DNAm;
- H_3 : there exist two distinct causal variants, one for each trait;
- H_4 : there exists a single causal variant common to both traits.

Conditional analyses

We performed a GCTA conditional analysis^{60,127} conditioning on the top *cis*-mQTL of those probes that passed SMR (P_{SMR} Bonferroni < 0.05) but failed to pass the HEIDI test ($P_{\text{HEIDI}} < 0.05$), due to the fact that heterogeneity may sometimes be driven by real secondary signals. We also performed the conditional analysis using GWAS summary data of the same set of SNPs (SNPs in the *cis*-mQTL region conditioning on the top *cis*-mQTL) for each of the three phenotypes (BIP, MDD, SCZ). For any of these regions where there was evidence of a secondary signal ($P_{\text{conditional}} < 5 \times 10^{-8}$), in either mQTL or GWAS data, we reran the conditional analyses in both mQTL and GWAS data conditioning on the secondary signal and then used the conditional results for SMR and HEIDI tests. In this step, significant secondary pleiotropic associations between DNAm and the neuropsychiatric diseases were selected as those with Bonferroni corrected $P_{\text{SMR}} < 0.05$ and $P_{\text{HEIDI}} > 0.05$ (not showing heterogeneity).

RICHs eQTLs

RICHs recruited mother and infant pairs from March 2009 to May 2013 following delivery at the Women and Infants Hospital of Rhode Island. RICHs selected infants both small for GA, and large for GA and controls born appropriate for GA matched on sex, GA (± 3 days), and maternal age (± 2 years). The study protocol was approved by the Institutional Review Boards of Brown University and Women and Infants Hospital of Rhode Island. Placental RNA-seq data from a subset of samples ($n = 200$) were obtained using the Illumina Hi-Seq 2500 platform, aligned to the human reference genome and RNA transcript abundance was quantified using Salmon¹²⁸. About 20 million single-end RNA-seq reads were generated on each sample¹²⁹. Placental DNAm data ($n = 220$) were obtained using the Illumina Infinium HumanMethylation450 BeadChip, preprocessed, and normalized with the minfi R package¹⁰⁶. eQTLs were calculated by implementing linear models in MatrixEQTL, considering *cis*-windows of 0.5 Mb up and downstream of each CpG in a total of 195 placenta samples. Linear regressions were adjusted by sex, five PCs of expression, and the Planet estimated cell types. Results were corrected with FDR.

The gene set enrichment analysis of the eQTL-genes associated to SCZ was performed using the Reactome gene list analysis⁵⁹, with the default settings recommended by the developers in the Reactome database release 83 and the pathway browser version 3.7.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The full datasets generated in the present study are publicly available online at the following address: https://irlab.shinyapps.io/shiny_mqtl_placenta/, including all the mQTLs used for downstream analyses. Individual genotype and DNAm raw data are not publicly available because neither children nor their parents have given consent for open publication of individual-level and sensitive data. However, these data are accessible upon request, through external collaboration with the INMA project. Detailed guidelines and application forms are available at <https://www.proyecto-inma.org/proyecto-inma/politica-de-colaboracion/>. Once completed, application forms should be sent to inma@proyecto-inma.org. In summary, data access is granted to experienced researchers affiliated with research institutions to address specific scientific questions within a limited timeframe. If the scientific question is already under study or of interest to multiple researchers, collaboration is encouraged. The INMA executive committee makes the final decision on all external applications. The GWAS summary statistics files reused in this study are available from the Psychiatric Genomics Consortium (PGC) downloads page: <https://pgc.unc.edu/for-researchers/download-results/> (ADHD, ASD, BIP, MDD, OCD, PD, SCZ), the EArly Genetics and Lifecourse Epidemiology (EAGLE) Consortium downloads page (upon registration): <https://www.eagle-consortium.org/downloads/> (AGR, INT), and the EBI GWAS catalog: <https://www.ebi.ac.uk/gwas/studies/GCST90042799> (SA). Source data for Figures and Supplementary Figs. have been uploaded and can be accessed online. Source data are provided with this paper.

Code availability

The code for the genotype and methylation QC, as well as the TensorQTL nominal mapping, is available in this GitHub repository link: <https://github.com/ariadnacilleros/Cis-mQTL-mapping-protocol-for-methylome>, and is also available on Zenodo [<https://doi.org/10.5281/zenodo.14198427>]¹³⁰. The rest of the analysis is available in this GitHub repository link: <https://github.com/ariadnacilleros/Cilleros-PortetA>. et al and also on Zenodo [<https://doi.org/10.5281/zenodo.14198405>]¹³⁰.

References

1. Strauss, R. S. Effects of the intrauterine environment on childhood growth. *Br. Med. Bull.* **53**, 81–95 (1997).
2. Barker, D. J. The fetal and infant origins of adult disease. *BMJ* **301**, 1111 (1990).
3. O'Donnell, K., O'Connor, T. G. & Glover, V. Prenatal stress and neurodevelopment of the child: focus on the HPA axis and role of the placenta. *Dev. Neurosci.* **31**, 285–292 (2009).
4. Lapehn, S. & Paquette, A. G. The placental epigenome as a molecular link between prenatal exposures and fetal health outcomes through the DOHaD hypothesis. *Curr. Environ. Health. Rep.* **9**, 490–501 (2022).
5. Bhattacharya, A. et al. Placental genomics mediates genetic associations with complex health traits and disease. *Nat. Commun.* **13**, 706 (2022).
6. Lester, B. M. & Marsit, C. J. Epigenetic mechanisms in the placenta related to infant neurodevelopment. *Epigenomics* **10**, 321–333 (2018).
7. Yen, S. S. The placenta as the third brain. *J. Reprod. Med.* **39**, 277–280 (1994).
8. Rapoport, J. L., Addington, A. M., Frangou, S. & Psych, M. R. The neurodevelopmental model of schizophrenia: update 2005. *Mol. Psychiatry* **10**, 434–449 (2005).
9. Murray, R. M. & Lewis, S. W. Is schizophrenia a neurodevelopmental disorder? *Br. Med. J.* **295**, 681–682 (1987).
10. Owen, M. J. & O'Donovan, M. C. Schizophrenia and the neurodevelopmental continuum: evidence from genomics. *World Psychiatry* **16**, 227–235 (2017).
11. Weinberger, D. R. The neurodevelopmental origins of schizophrenia in the penumbra of genomic medicine. *World Psychiatry* **16**, 225–226 (2017).
12. Canetta, S. E. & Brown, A. S. PRENATAL INFECTION, MATERNAL IMMUNE ACTIVATION, AND RISK FOR SCHIZOPHRENIA. *Transl. Neurosci.* **3**, 320–327 (2012).
13. Cannon, M., Jones, P. B. & Murray, R. M. Obstetric complications and schizophrenia: historical and meta-analytic review. *Am. J. Psychiatry* **159**, 1080–1092 (2002).
14. Carnac, T. Schizophrenia hypothesis: autonomic nervous system dysregulation of fetal and adult immune tolerance. *Front. Syst. Neurosci.* **16**, 844383 (2022).
15. Novakovic, B. & Saffery, R. The ever growing complexity of placental epigenetics—role in adverse pregnancy outcomes and fetal programming. *Placenta* **33**, 959–970 (2012).
16. Konwar, C., Del Gobbo, G., Yuan, V. & Robinson, W. P. Considerations when processing and interpreting genomics data of the placenta. *Placenta* **84**, 57–62 (2019).
17. Loyfer, N. et al. A DNA methylation atlas of normal human cell types. *Nature* **613**, 355–364 (2023).
18. Everson, T. M. et al. Placental DNA methylation signatures of maternal smoking during pregnancy and potential impacts on fetal growth. *Nat. Commun.* **12**, 5095 (2021).
19. Fernandez-Jimenez, N. et al. A meta-analysis of pre-pregnancy maternal body mass index and placental DNA methylation identifies 27 CpG sites with implications for mother-child health. *Commun. Biol.* **5**, 1313 (2022).
20. Do, C. et al. Mechanisms and disease associations of haplotype-dependent allele-specific DNA methylation. *Am. J. Hum. Genet.* **98**, 934–955 (2016).
21. Delahaye, F. et al. Genetic variants influence on the placenta regulatory landscape. *PLoS Genet.* **15**, e1008118 (2019).
22. Tekola-Ayele, F. et al. Placental multi-omics integration identifies candidate functional genes for birthweight. *Nat. Commun.* **13**, 2384 (2022).
23. Casazza, W. et al. Sex-dependent placental methylation quantitative trait loci provide insight into the prenatal origins of childhood onset traits and conditions. *iScience* **27**, 109047 (2024).
24. Oliva, M. et al. DNA methylation QTL mapping across diverse human tissues provides molecular links between genetic variation and complex traits. *Nat. Genet.* **55**, 112–122 (2023).
25. Kim-Hellmuth, S. et al. Cell type-specific genetic regulation of gene expression across human tissues. *Science* **369**, eaaz8528 (2020).
26. Guxens, M. et al. Cohort profile: the INMA—Infancia y medio ambiente—(environment and childhood). *Proj. Int. J. Epidemiol.* **41**, 930–940 (2012).
27. Yuan, V. et al. Cell-specific characterization of the placental methylome. *BMC Genomics* **22**, 6 (2021).
28. Min, J. L. et al. Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. *Nat. Genet.* **53**, 1311–1321 (2021).
29. McCarthy, S. et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
30. Breeze, C. E. et al. eFORGE v2.0: updated analysis of cell type-specific signal in epigenomic data. *Bioinformatics* **35**, 4767–4769 (2019).
31. Breeze, C. E. et al. eFORGE: a tool for identifying cell type-specific signal in epigenomic data. *Cell. Rep.* **17**, 2137–2150 (2016).
32. Breeze, C. E. Cell type-specific signal analysis in epigenome-wide association studies. *Methods Mol. Biol.* **2432**, 57–71 (2022).
33. Schriml, L. M. et al. Human Disease Ontology 2018 update: classification, content and workflow expansion. *Nucleic Acids Res.* **47**, D955–D962 (2019).
34. Ellery, P. M., Cindrova-Davies, T., Jauniaux, E., Ferguson-Smith, A. C. & Burton, G. J. Evidence for transcriptional activity in the syncytiotrophoblast of the human placenta. *Placenta* **30**, 329–334 (2009).
35. Davis, J. R. et al. An efficient multiple-testing adjustment for eQTL studies that accounts for linkage disequilibrium between variants. *Am. J. Hum. Genet.* **98**, 216–224 (2016).
36. Garrido-Martin, D. et al. A fast non-parametric test of association for multiple traits. *Genome Biol.* **24**, 230 (2023).
37. Storey, J. D. & Tibshirani, R. Statistical significance for genome-wide studies. *Proc. Natl. Acad. Sci. USA* **100**, 9440–9445 (2003).
38. Demontis, D. et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* **51**, 63–75 (2019).
39. Pappa, I. et al. A genome-wide approach to children's aggressive behavior: The EAGLE consortium. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **171**, 562–572 (2016).
40. Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium. Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia. *Mol. Autism* **8**, 21 (2017).
41. Mullins, N. et al. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat. Genet.* **53**, 817–829 (2021).
42. Jami, E. S. et al. Genome-wide Association Meta-analysis of Childhood and Adolescent Internalizing Symptoms. *J. Am. Acad. Child Adolesc. Psychiatry* **61**, 934–945 (2022).
43. Howard, D. M. et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat. Neurosci.* **22**, 343–352 (2019).
44. International Obsessive Compulsive Disorder Foundation Genetics Collaborative (IOCDF-GC) and OCD Collaborative Genetics Association Studies (OCAS). Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Mol. Psychiatry* **23**, 1181–1188 (2018).
45. Forstner, A. J. et al. Genome-wide association study of panic disorder reveals genetic overlap with neuroticism and depression. *Mol. Psychiatry* **26**, 4179–4190 (2021).

46. Jiang, L., Zheng, Z., Fang, H. & Yang, J. A generalized linear mixed model association tool for biobank-scale data. *Nat. Genet.* **53**, 1616–1621 (2021).
47. Trubetskoy, V. et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).
48. Psychiatric Genomics Consortium (PGC, Accessed 23 Jan 2023); <https://pgc.unc.edu/>
49. EArly Genetics and Lifecourse Epidemiology (EAGLE, Accessed 23 Jan 2023); <https://eagle-consortium.org/>
50. Zhu, Z. et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
51. Lee, Y. et al. Placental epigenetic clocks: estimating gestational age using placental DNA methylation levels. *Aging* **11**, 4238–4253 (2019).
52. Giambartolomei, C. et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).
53. Appleton, A. A. et al. Prenatal Programming of Infant Neurobehaviour in a Healthy Population. *Paediatr. Perinat. Epidemiol.* **30**, 367–375 (2016).
54. Shabalin, A. A. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* **28**, 1353–1358 (2012).
55. Johnston, K. J. A. et al. Identification of novel common variants associated with chronic pain using conditional false discovery rate analysis with major depressive disorder and assessment of pleiotropic effects of LRFN5. *Transl. Psychiatry* **9**, 310 (2019).
56. Mischler, A. et al. Two distinct trophoblast lineage stem cells from human pluripotent stem cells. *J. Biol. Chem.* **296**, 100386 (2021).
57. Li, Y. et al. A missense variant in NDUFA6 confers schizophrenia risk by affecting YY1 binding and NAGA expression. *Mol. Psychiatry* **26**, 6896–6911 (2021a).
58. Li, Y., Li, S., Liu, J., Huo, Y. & Luo, X. J. The schizophrenia susceptibility gene NAGA regulates dendritic spine density: further evidence for the dendritic spine pathology of schizophrenia. *Mol. Psychiatry* **26**, 7102–7104 (2021b).
59. Gillespie, M. et al. The reactome pathway knowledgebase 2022. *Nucleic Acids Res.* **50**, D687–D692 (2022).
60. Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).
61. Kent, W. J. et al. The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
62. Cleal, J. K., Lofthouse, E. M., Sengers, B. G. & Lewis, R. M. A systems perspective on placental amino acid transport. *J. Physiol.* **596**, 5511–5522 (2018).
63. Mercer, G. V. et al. Identifying placental metabolic biomarkers of preterm birth using nuclear magnetic resonance of intact tissue samples. *Placenta* **143**, 80–86 (2023).
64. Brkić, J. et al. Differential role of Smad2 and Smad3 in the acquisition of an endovascular trophoblast-like phenotype and pre-eclampsia. *Front. Endocrinol.* **11**, 436 (2020).
65. Liang, M. H., Wendland, J. R. & Chuang, D. M. Lithium inhibits Smad3/4 transactivation via increased CREB activity induced by enhanced PKA and AKT signaling. *Mol. Cell. Neurosci.* **37**, 440–453 (2008).
66. Hannon, E. et al. Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. *Nat. Neurosci.* **19**, 48–54 (2016).
67. Qi, T. et al. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. *Nat. Commun.* **9**, 2282 (2018).
68. Ng, B. et al. An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. *Nat. Neurosci.* **20**, 1418–1426 (2017).
69. Jaffe, A. E. et al. Mapping DNA methylation across development, genotype and schizophrenia in the human frontal cortex. *Nat. Neurosci.* **19**, 40–47 (2016).
70. Mayhew, T. M. & Barker, B. L. Villous trophoblast: morphometric perspectives on growth, differentiation, turnover and deposition of fibrin-type fibrinoid during gestation. *Placenta* **22**, 628–638 (2001).
71. White, R., Jobling, S., Hoare, S. A., Sumpter, J. P. & Parker, M. G. Environmentally persistent alkylphenolic compounds are estrogenic. *Endocrinology* **135**, 175–182 (1994).
72. McLaren, A. J. *Endocrinol.* 42,453 (1968); McLaren, A. In: *Ovo-Implantation: human gonadotropins and prolactin* (Hubinout, P. O. et al., Eds.) 18–37 (Karger, New York, 1970).
73. Finn, C. A. & Bredl, J. C. Studies on the development of the implantation reaction in the mouse uterus: influence of actinomycin D. *J. Reprod. Fert.* **34**, 247–253 (1973).
74. Braga, V. M. & Gendler, S. J. Modulation of Muc-1 mucin expression in the mouse uterus during the estrus cycle, early pregnancy and placentation. *J. Cell Sci.* **105**, 397–405 (1993).
75. Rao, S., Yin, L., Xiang, Y. & So, H. C. Correction: analysis of genetic differences between psychiatric disorders: exploring pathways and cell types/tissues involved and ability to differentiate the disorders by polygenic scores. *Transl. Psychiatry* **12**, 237 (2022).
76. Romero, C. et al. Exploring the genetic overlap between twelve psychiatric disorders. *Nat. Genet.* **54**, 1795–1802 (2022).
77. Ursini, G. et al. Convergence of placenta biology and genetic risk for schizophrenia. *Nat. Med.* **24**, 792–801 (2018).
78. Ursini, G. et al. Placental genomic risk scores and early neurodevelopmental outcomes. *Proc. Natl. Acad. Sci. USA* **118**, e2019789118 (2021).
79. Grotzinger, A. D. et al. Genetic architecture of 11 major psychiatric disorders at biobehavioral, functional genomic and molecular genetic levels of analysis. *Nat. Genet.* **54**, 548–559 (2022).
80. Bora, E. Differences in cognitive impairment between schizophrenia and bipolar disorder: Considering the role of heterogeneity. *Psychiatry Clin. Neurosci.* **70**, 424–433 (2016).
81. 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).
82. 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).
83. Ursini, G. et al. Prioritization of potential causative genes for schizophrenia in placenta. *Nat. Commun.* **14**, 2613 (2023).
84. Han, V. X., Patel, S., Jones, H. F. & Dale, R. C. Maternal immune activation and neuroinflammation in human neurodevelopmental disorders. *Nat. Rev. Neurol.* **17**, 564–579 (2021).
85. Weber-Stadlbauer, U. Epigenetic and transgenerational mechanisms in infection-mediated neurodevelopmental disorders. *Transl. Psychiatry* **7**, e1113 (2017).
86. Richetto, J. et al. Genome-wide DNA methylation changes in a mouse model of infection-mediated neurodevelopmental disorders. *Biol. Psychiatry* **81**, 265–276 (2017).
87. Estes, M. L. & McAllister, A. K. Maternal immune activation: implications for neuropsychiatric disorders. *Science* **353**, 772–777 (2016).
88. Knuesel, I. et al. Maternal immune activation and abnormal brain development across CNS disorders. *Nat. Rev. Neurol.* **10**, 643–660 (2014).

89. Banik, A. et al. Maternal factors that induce epigenetic changes contribute to neurological disorders in offspring. *Genes* **8**, 150 (2017).
90. Hodyl, N. A., Roberts, C. T. & Bianco-Miotto, T. Cord blood DNA methylation biomarkers for predicting neurodevelopmental outcomes. *Genes* **7**, 117 (2016).
91. Howerton, C. L. & Bale, T. L. Prenatal programming: at the intersection of maternal stress and immune activation. *Horm. Behav.* **62**, 237–242 (2012).
92. Apicella, C. et al. Pan-genomic regulation of gene expression in normal and pathological human placentas. *Cells* **12**, 578 (2023).
93. Zhang, J. et al. SFBMT1 functions with LSD1 to regulate expression of canonical histone genes and chromatin-related factors. *Genes Dev.* **27**, 749–766 (2013).
94. Miri, K. et al. The imprinted polycomb group gene Sfbmt2 is required for trophoblast maintenance and placenta development. *Development* **140**, 4480–4489 (2013).
95. Lybaek, H. et al. LRFN5 locus structure is associated with autism and influenced by the sex of the individual and locus conversions. *Autism Res.* **15**, 421–433 (2022).
96. Glausier, J. R. & Lewis, D. A. Dendritic spine pathology in schizophrenia. *Neuroscience* **251**, 90–107 (2013).
97. Murmu, M. S., et al (2006). Changes of spine density and dendritic complexity in the prefrontal cortex in offspring of mothers exposed to stress during pregnancy. *Eur. J. Neurosci.* **24**, 1477–1487 (2006).
98. Anderson, C. A. et al. Data quality control in genetic case-control association studies. *Nat. Protoc.* **5**, 1564–1573 (2010).
99. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
100. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
101. Das, S. et al. Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).
102. Loh, P. R. et al. Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* **48**, 1443–1448 (2016).
103. Fuchsberger, C., Abecasis, G. R. & Hinds, D. A. minimac2: faster genotype imputation. *Bioinformatics* **31**, 782–784 (2015).
104. Binder, A.M. QC steps: epi-centered research. <https://www.epicenteredresearch.com/pace/qcsteps/> (2021).
105. 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).
106. Fortin, J. P., Triche, T. J. Jr & Hansen, K. D. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics* **33**, 558–560 (2017).
107. Triche, T. J. Jr, Weisenberger, D. J., Van Den Berg, D., Laird, P. W. & Siegmund, K. D. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res.* **41**, e90 (2013).
108. Fortin, J. P. et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol.* **15**, 503 (2014).
109. Teschendorff, A. E. et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* **29**, 189–196 (2013).
110. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).
111. Taylor-Weiner, A. et al. Scaling computational genomics to millions of individuals with GPUs. *Genome Biol.* **20**, 228 (2019).
112. Ongen, H., Buil, A., Brown, A. A., Dermizakis, E. T. & Delaneau, O. Fast and efficient QTL mapper for thousands of molecular phenotypes. *Bioinformatics* **32**, 1479–1485 (2016).
113. Gibbs, J. R. et al. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* **6**, e1000952 (2010).
114. Drong, A. W. et al. The presence of methylation quantitative trait loci indicates a direct genetic influence on the level of DNA methylation in adipose tissue. *PLoS One* **8**, e55923 (2013).
115. Olsson, A. H. et al. Genome-wide associations between genetic and epigenetic variation influence mRNA expression and insulin secretion in human pancreatic islets. *PLoS Genet.* **10**, e1004735 (2014).
116. GTEx Consortium; Laboratory, Data Analysis & Coordinating Center (LDACC)—Analysis Working Group; Statistical Methods groups—Analysis Working Group. Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017).
117. Hansen, K. D. IlluminaHumanMethylationEPICanno.ilm10b4.hg19: annotation for Illumina’s EPIC methylation arrays. R package version 0.6.0, https://bitbucket.com/kasperdanielhansen/Illumina_EPIC (2017).
118. Roadmap Epigenomics Consortium. et al. Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
119. Maksimovic, J., Gordon, L. & Oshlack, A. SWAN: subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol.* **13**, R44 (2012).
120. Phipson, B. & Oshlack, A. DiffVar: a new method for detecting differential variability with application to methylation in cancer and aging. *Genome Biol.* **15**, 465 (2014).
121. Maksimovic, J., Gagnon-Bartsch, J. A., Speed, T. P. & Oshlack, A. Removing unwanted variation in a differential methylation analysis of Illumina HumanMethylation450 array data. *Nucleic Acids Res.* **43**, e106 (2015).
122. Phipson, B., Maksimovic, J. & Oshlack, A. missMethyl: an R package for analyzing data from Illumina’s HumanMethylation450 platform. *Bioinformatics* **32**, 286–288 (2016).
123. Yu, G., Wang, L. G., Yan, G. R. & He, Q. Y. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics* **31**, 608–609 (2015).
124. Storey, J.D., Bass, A.J., Dabney, A., Robinson, D. (2023). qvalue: Q-value estimation for false discovery rate control. <https://doi.org/10.18129/B9.bioc.qvalue>, R package version 2.34.0 <https://bioconductor.org/packages/qvalue>
125. Boughton, A. P. et al. LocusZoom.js: interactive and embeddable visualization of genetic association study results. *Bioinformatics* **37**, 3017–3018 (2021).
126. Wu, Y. et al. Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nat. Commun.* **9**, 918 (2018).
127. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
128. Love, M. I., Hogenesch, J. B. & Irizarry, R. A. Modeling of RNA-seq fragment sequence bias reduces systematic errors in transcript abundance estimation. *Nat. Biotechnol.* **34**, 1287–1291 (2016).
129. Peng, S. et al. Expression quantitative trait loci (eQTLs) in human placentas suggest developmental origins of complex diseases. *Hum. Mol. Genet.* **26**, 3432–3441 (2017).
130. Cilleros-Portet, A. et al. Potentially causal associations between placental DNA methylation and schizophrenia and other neuropsychiatric disorders. Zenodo <https://doi.org/10.5281/zenodo.14198427> (2024).

Acknowledgements

We would like to acknowledge all the INMA and RICHS participants and researchers, for their kind collaboration and support. INMA-Gipuzkoa is funded by grants from Instituto de Salud Carlos III (PI06/O867 and PI09/

00090, incl. FEDER funds), Department of Health of the Basque Government (2005111093), Provincial Government of Gipuzkoa (DFG06/002), and annual agreements with the municipalities of the study area (Zumarraga, Urretxu, Legazpi, Azkoitia, Azpeitia and Beasain). INMA-Sabadell was funded by grants from Instituto de Salud Carlos III (Red INMA G03/176, PS09/00432, PI17/01225, PI17/01935, and CP18/00018), Fundació La Marató de TV3 (090430), and Generalitat de Catalunya-CIRIT (1999SGR 00241), the European Community's Seventh Framework Program (FP7/2007-206) under grant agreement no 308333 (HELIX project), and from the European Joint Programming Initiative "A Healthy Diet for a Healthy Life" (JPI HDHL and Instituto de Salud Carlos III) under the grant agreement no AC18/00006 (NutriPROGRAM project). ISGlobal acknowledges support from the Spanish Ministry of Science and Innovation and the State Research Agency through the "Centro de Excelencia Severo Ochoa 2019-2023" Program (CEX2018-000806-S), and support from the Generalitat de Catalunya through the CERCA Program. INMA-Valencia is funded by Grants from UE (FP7-ENV-2011 cod 282957, HEALTH.2010.2.4.5-1, and H2020 No 874583, the ATHLETE project), the Ministry of Universities (CAS21/00008, Margarita Salas Grant MS21-133 and NextGeneration EU), Instituto de Salud Carlos III (FIS-FEDER: 13/1944, 16/1288, 17/00663, and 19/1338; FIS-FSE: 17/00260; Miguel Servet-FSE: MSII20/0006), CIBERESP, Generalitat Valenciana (BEST/2020/059, AICO/2020/285 and CIAICO/2021/132). The RICHs cohort is supported by the US National Institute of Environmental Health Sciences (U24 ES02507). M.C.-T. is funded by a Beatriz de Pinós Postdoctoral Contract awarded by Generalitat de Catalunya-AGAUR and European Commission- Horizon 2020 (2019 BP 00107). I.G.-S. is funded by the Basque Department of Health (SAN2020111043) and the Spanish Ministry of Equity (12-4-ID22) and the UPV/EHU Collaborative Projects (COLAB22/01). A.H.-L. is a predoctoral fellow supported by grant PRE-C-2020-0091 from the MCIN/AEI/10.13039/501100011033 and by ESF Investing in your future. B.P.G.-G. is supported by the Mexican National Council for Science and Technology grant 2021-000007-01EXTF-00209. M.F.F. is funded by the EU Commission (QLK4-1999-01422, QLK4-2002-00603, and CONTAMED FP7-ENV-212502) and the Consejería de Salud de la Junta de Andalucía (Grant number 0675/10). J.R.B. is funded by Research Grant PID2019-106382RB-I00 funded by MCIN/AEI/10.13039/501100011033. N.F.-J. is funded by research grants 2019/111085 from the Basque Department of Health, and PI21/01491 from the Instituto de Salud Carlos III (ISCIII), co-funded by the European Union.

Author contributions

N.F.-J., J.R.B., and A.C.-P. conceived and designed the study. A.C.-P., C.L., S.M., C.E.B., and N.F.J. carried out data analyses. I.G.S., D.G.M., A.H.-L., B.P.G.-G., and M.B. helped with the data analysis and provided valuable analytical advice. M.C.T., M.L., A.I., A.B., G.E., R.S.-B., L.S.M., J.C., S.L., M.F.F., M.V., J.I., M.G., C.M., M.B., and J.R.B. actively

participated in sample recruitment and/or data acquisition. A.C.-P., N.F.-J., and J.R.B. interpreted the results. N.F.J. and A.C.-P. wrote the first draft of the manuscript. N.F.J. directed the project. All authors (A.C.-P., C.L., S.M., M.C.T., M.L., A.I., A.B., I.G.-S., D.G.M., G.E., A.H.-L., R.S.B., C.E.B., B.P.G.-G., L.S.M., J.C., S.L., M.F.F., M.V., J.I., M.G., C.M., M.B., J.R.B., and N.F.-J.) made substantial contributions to the acquisition, analysis, or interpretation of data, and read and critically revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-57760-3>.

Correspondence and requests for materials should be addressed to Nora Fernandez-Jimenez.

Peer review information *Nature Communications* thanks Meritxell Oliva, Gianluca Ursini and the other, anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025, corrected publication 2025

Ariadna Cilleros-Portet ¹, **Corina Lesseur** ², **Sergi Mari** ¹, **Marta Cosin-Tomas**^{3,4,5}, **Manuel Lozano** ^{6,7}, **Amaia Irizar**^{4,8,9}, **Amber Burt** ¹⁰, **Iraia García-Santisteban**¹, **Diego Garrido-Martín**¹¹, **Geòrgia Escaramís**^{4,12}, **Alba Hernangomez-Laderas** ¹, **Raquel Soler-Blasco**^{4,6,13}, **Charles E. Breeze**¹⁴, **Bárbara P. Gonzalez-Garcia**¹, **Loreto Santa-Marina** ^{4,9,15}, **Jia Chen**², **Sabrina Llop**^{4,6}, **Mariana F. Fernández** ^{4,16,17}, **Martine Vrijheid** ^{3,4,5}, **Jesús Ibarluzea** ^{4,9,15}, **Mònica Guxens** ^{3,4,5,18,19}, **Carmen Marsit** ¹⁰, **Mariona Bustamante** ^{3,4,5}, **Jose Ramon Bilbao** ^{1,20} & **Nora Fernandez-Jimenez** ¹✉

¹Department of Genetics, Physical Anthropology and Animal Physiology, Biobizkaia Health Research Institute and University of the Basque Country (UPV/EHU), Leioa, Spain. ²Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ³ISGlobal, Barcelona, Spain. ⁴Spanish Consortium for Research on Epidemiology and Public Health (CIBERESP), Instituto de Salud Carlos III, Madrid, Spain. ⁵Universitat Pompeu Fabra, Barcelona, Spain. ⁶Epidemiology and Environmental Health Joint Research Unit, FISABIO-Universitat Jaume I-Universitat de València, Valencia, Spain. ⁷Preventive Medicine and Public Health, Food Sciences, Toxicology and Forensic Medicine Department, Universitat de València, Valencia, Spain. ⁸Department of Preventive Medicine and Public Health, University of the Basque Country (UPV/EHU), Leioa, Spain. ⁹Biogipuzkoa Health Research Institute, San Sebastian, Spain. ¹⁰Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta,

GA, USA. ¹¹Department of Genetics, Microbiology and Statistics, Faculty of Biology, Universitat de Barcelona (UB), Barcelona, Spain. ¹²Departament de Biomedicina, Facultat de Medicina i Ciències de la Salut, Institut de Neurociències, Universitat de Barcelona, Barcelona, Spain. ¹³Department of Nursing, Universitat de València, Valencia, Spain. ¹⁴UCL Cancer Institute, University College London, London, UK. ¹⁵Department of Health of the Basque Government, Subdirectorate of Public Health of Gipuzkoa, San Sebastian, Spain. ¹⁶Department of Radiology and Physical Medicine, Biomedical Research Center (CIBM), School of Medicine, University of Granada, Granada, Spain. ¹⁷Instituto de Investigación Biosanitaria de Granada (ibs.GRANADA), Granada, Spain. ¹⁸Department of Child and Adolescent Psychiatry/Psychology, Erasmus MC, University Medical Centre, Rotterdam, The Netherlands. ¹⁹ICREA, Barcelona, Spain. ²⁰CIBER de Diabetes y Enfermedades Metabólicas Asociadas (CIBERDEM), Madrid, Spain. ✉e-mail: nora.fernandez@ehu.eus