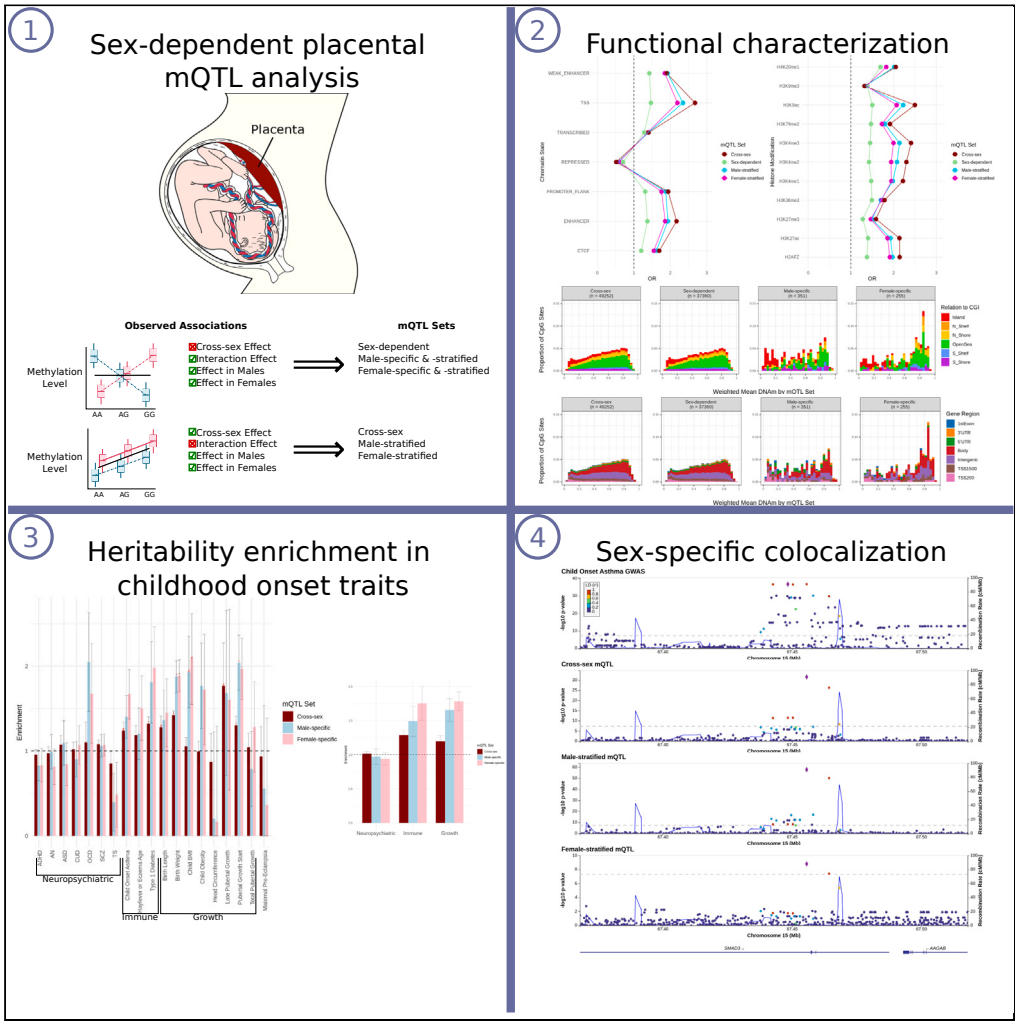
iScience

CellPress
OPEN ACCESS

Article

Sex-dependent placental methylation quantitative trait loci provide insight into the prenatal origins of childhood onset traits and conditions



William Casazza, Amy M. Inkster, Giulia F. Del Gobbo, ..., Wendy P. Robinson, Sara Mostafavi, Jessica K. Dennis

jessica.dennis@bcchr.ca

Highlights

Placental mQTL demonstrate sex-specificity

Placental mQTL are enriched in GWAS results for childhood-onset growth- and immune-related traits

Male- and female-specific mQTL showed higher trait enrichment than cross-sex mQTL

Nearly a fourth of mQTL that colocalized with GWAS loci were male- or female-specific

Casazza et al., iScience 27, 109047
February 16, 2024
https://doi.org/10.1016/j.isci.2024.109047

iScience

CellPress
OPEN ACCESS

Article

Sex-dependent placental methylation quantitative trait loci provide insight into the prenatal origins of childhood onset traits and conditions

William Casazza,[1,2,3] Amy M. Inkster,[3,4] Giulia F. Del Gobbo,[3,4,5] Victor Yuan,[3,4] Fabien Delahaye,[6] Carmen Marsit,[7] Yongjin P. Park,[8,9] Wendy P. Robinson,[3,4] Sara Mostafavi,[1,10] and Jessica K. Dennis[1,2,3,4,11,*]

SUMMARY

Molecular quantitative trait loci (QTLs) allow us to understand the biology captured in genome-wide association studies (GWASs). The placenta regulates fetal development and shows sex differences in DNA methylation. We therefore hypothesized that placental methylation QTL (mQTL) explain variation in genetic risk for childhood onset traits, and that effects differ by sex. We analyzed 411 term placentas from two studies and found 49,252 methylation (CpG) sites with mQTL and 2,489 CpG sites with sex-dependent mQTL. All mQTL were enriched in regions that typically affect gene expression in prenatal tissues. All mQTL were also enriched in GWAS results for growth- and immune-related traits, but male- and female-specific mQTL were more enriched than cross-sex mQTL. mQTL colocalized with trait loci at 777 CpG sites, with 216 (28%) specific to males or females. Overall, mQTL specific to male and female placenta capture otherwise overlooked variation in childhood traits.

INTRODUCTION

GWAS findings hold valuable clues about trait mechanisms.[1,2] Deciphering these clues, however, requires additional data on gene regulation, as over 90% of SNPs identified in GWAS lie in gene regulatory regions, as opposed to in the protein-coding gene region itself.[3,4] Molecular quantitative trait loci (molQTL) analysis is a powerful strategy to interpret the gene regulatory functions of GWAS SNPs. Under the umbrella of molQTL, expression quantitative trait loci (eQTL) are the most widely studied, and they are enriched for GWAS loci relative to other SNPs matched by minor allele frequency (MAF).[5] However, only 43% of eQTL share the same causal variant (colocalize) with a GWAS locus,[5] and up to 77% of eQTL in linkage disequilibrium (LD) with a trait-associated SNP are shared across more than one tissue.[6] As a result, the majority of GWAS loci have either no known effects on expression or their relationship with traits is clouded by their broad effects on gene expression across tissues.[5,7]

Therefore, to advance the functional interpretation of GWAS SNPs, we must extend molQTL discovery across molecular traits, tissues, and biological contexts. In this study, we identify DNA methylation QTL (mQTL) in placenta, and additionally focus on mQTL that have different effects in males vs. females. DNAm is an attractive molecular trait for the functional interpretation of GWAS results because it can provide insights on the precise molecular mechanism by which GWAS SNPs associate with traits and conditions: through biochemical modification to DNA sequence at a CpG site.[8] In addition, variation in DNAm is genetically influenced. In blood, the tissue in which DNAm is most widely studied, 21% of the variation of DNAm is explained by additive genetic variation in cis (i.e., via mQTL).[9] Importantly, mQTL provide information on gene regulation beyond what is provided by eQTL. For example, mQTL cover roughly twice as many genes as eQTL in blood.[5,9,10] In addition, a recent mapping of *cis*-mQTL across 8 GTEx tissues found that 79% of mQTL did not colocalize with an eQTL, and 55% of mQTL colocalized with a trait, of which only one-third also colocalized with an eQTL.[11] Thus, continuing to pursue mQTL in additional tissues could increase our understanding of gene regulatory mechanisms involved in complex traits and conditions.

The placenta is of the same genetic make up as the fetus and it is one of the first organs to form during gestation. Throughout pregnancy, it is responsible for the exchange of oxygen, nutrients, and hormones between mother and fetus. Despite its central role in human

[1]Centre for Molecular Medicine and Therapeutics, BC Children's Hospital, Vancouver, BC, Canada
[2]Bioinformatics Graduate Program, University of British Columbia, Vancouver, BC, Canada
[3]BC Children's Hospital Research Institute, Vancouver, BC, Canada
[4]Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada
[5]Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, ON, Canada
[6]Albert Einstein College of Medicine, The Bronx, NY, USA
[7]Rollins School of Public Health, Emory University, Atlanta, GA, USA
[8]Department of Statistics, University of British Columbia, Vancouver, BC, Canada
[9]Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada
[10]Paul Allen School of Computer Science and Engineering, University of Washington, Seattle, WA, USA
[11]Lead contact
*Correspondence: jessica.dennis@bcchr.ca
https://doi.org/10.1016/j.isci.2024.109047

development, however, placenta is under-studied and is not represented in large-scale molQTL resources such as GTEx. Several studies have previously investigated placental molQTL and their relationship with complex traits and conditions.[12–17] In the largest mQTL analysis to date, Delahaye et al.[13] analyzed 303 placental samples from the the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) study and characterized a small number (N = 4,342) of strongly associated (i.e., passing a stringent permutation test and quality threshold) mQTL, which were found to overlap two type 2 diabetes loci. Tekola-Ayele et al.[14] analyzed the same samples with a similar approach and using colocalization analysis, implicated four genes with placental DNAm and gene expression that shared genetic loci with birth weight. While these analyses of the NICHD study demonstrated the relevance of placental molecular traits to postnatal outcomes, the stringent thresholds used to map mQTL means that many mQTL remain unmapped, especially in larger sample sizes. In addition, both Delahaye and Tekola-Ayele analyzed at most two GWAS traits, and the relationship of placental mQTL to multiple postnatal outcomes has yet to be investigated. The Rhode Island Child Health Study (RICHS)[16,18] is another large study that has collected molecular data from 149 placental samples, but mQTL have not been mapped in the RICHS study, and overall, a more comprehensive study is needed to better understand how placental mQTL affect genome-wide risk of traits.

Additionally, neither the NICHD study nor RICHS have investigated sex differences in the genetic regulation of placental molecular traits. Oliva et al. recently analyzed sex differences in eQTL across 44 GTEx tissues and found that sex-dependent eQTL (i.e., cross-tissue mQTL with a genotype by sex interaction) were remarkably tissue specific: of 369 sex-dependent eQTL found in at least one tissue, only one was shared by two tissues.[19] Moreover, they found 74 eQTL in either males or females that colocalized with GWAS loci, 24 of which showed no evidence of colocalization in eQTL computed in all subjects. These results suggest that sex-dependent molQTL yield functional interpretations of GWAS loci beyond what is provided by cross-sex analysis. Analyzing sex-dependent molQTL could be especially important in the placenta, since sex is strongly associated with placental molecular traits such as DNAm, even after accounting for cell type proportions,[20,21] which can bias sex-dependent molQTL analysis.[19]

In this work, we identify placental mQTL with a shared effect in males and females (cross-sex mQTL) and mQTL that are modified by sex (sex-dependent mQTL) by meta-analyzing data from 411 term placentas from the NICHD study and RICHS (Tables S1 and S2). We then compare cross-sex and sex-dependent mQTL with regards to their genomic location. Finally, we quantify the relevance of cross-sex and sex-dependent mQTL to childhood onset traits and conditions using stratified linkage disequilibrium score regression (S-LDSC)[22–24] and co-localization analysis.

RESULTS

Placental methylation quantitative trait loci replicate across studies despite differences in genetic ancestry

Both the NICHD study and RICHS selected participants with low risk pregnancies and analyzed placentas from births ≥37 weeks gestation. However, participants in NICHD were selected from diverse race/ethnicity groups (25.5% White; Table S1), while those from RICHS were primarily of White race/ethnicity (72.8% White; Table S1). Ancestry-specific differences in allele frequencies can lead to ancestry-specific molQTL.[25,26] Therefore, before meta-analyzing mQTL effects across NICHD and RICHS, we first identified and compared ancestry-specific mQTL in NICHD using the three largest ancestry groups, as estimated by GRAF-pop (29.8% African American, or AFR_AM, 26.2% European, or EUR, and 27.6% Latin American, or LAT_AM; Table S2). We calculated the overlap of mQTL between ancestries using the π_1 statistic, which is the proportion of associations in one population that have similar effects in another (STAR Methods).[27,28] The π_1 estimates calculated between each pair of ancestries were high (π_1 = 0.89 EUR vs. AFR_AM, π_1 = 0.90 EUR vs. LAT_AM, and π_1 = 0.90 AFR_AM vs. LAT_AM, computed at all mQTL significant at FDR <0.05, Figures S1A–S1C). More CpG sites had at least one mQTL associated at a Bonferroni corrected $p < 0.05$ in AFR_AM and LAT_AM samples compared to EUR samples (19,268 in AFR_AM, 12,158 in LAT_AM, and 8,748 in EUR), which likely reflects the increased genetic variability in those populations relative to populations with EUR ancestry.[26,29,30] Nonetheless, 6,754 (77.2%) CpGs with a significant mQTL in EUR were shared by one or more ancestral group, and 4,733 (54.1%) CpGs had a significant mQTL in all populations (Figures S1D and S1E). While these ancestry-specific patterns warrant further investigation, our study was focused on sex differences. Therefore, reassured by the overlap across ancestries, we proceeded with an ancestry-pooled analysis in NICHD and RICHS, and accounted for ancestry using genotyping PCs in all analyses.

As a final check, we compared mQTL computed in NICHD and RICHS and found a π_1 estimate of 0.74 (computed over 2,691,024 SNP-CpG pairs available in both studies showing association in NICHD at an FDR <0.05). Sex-dependent mQTL showed modest replication (π_1 = 0.28, computed over 80,363 SNP-CpG pairs available in both studies), which is in line with values previously reported for sex-dependent eQTL in GTEx tissues (π_1 = 0.28 in breast tissue, π_1 ranging from 0 to 0.12 in the remaining 43 GTEx V8 tissues analyzed by Oliva et al.)[19] Based on the high replication between NICHD and RICHS, in all subsequent analyses, we meta-analyzed mQTL across studies.

Sex-dependent placental methylation quantitative trait loci are distinct from cross-sex placental methylation quantitative trait loci

We meta-analyzed effects using MeCS[31] software to account for bias induced by correlation in mQTL effects between studies, and meta-analyzed mQTL effects (or genotype by sex interaction effects in the sex-dependent analysis) were called at a Bonferroni corrected $p < 0.05$. We defined 6 sets of mQTL (Figure 1; Table 1) which we used in all downstream analyses: (i) cross-sex, which have an effect independent of sex (typically just referred to as "mQTL"); (ii) sex-dependent, which have an effect modified by sex; (iii) male-stratified, which have an effect in male samples; (iv) female-stratified, which have an effect in female samples; (v) male-specific,which have an effect in males that also

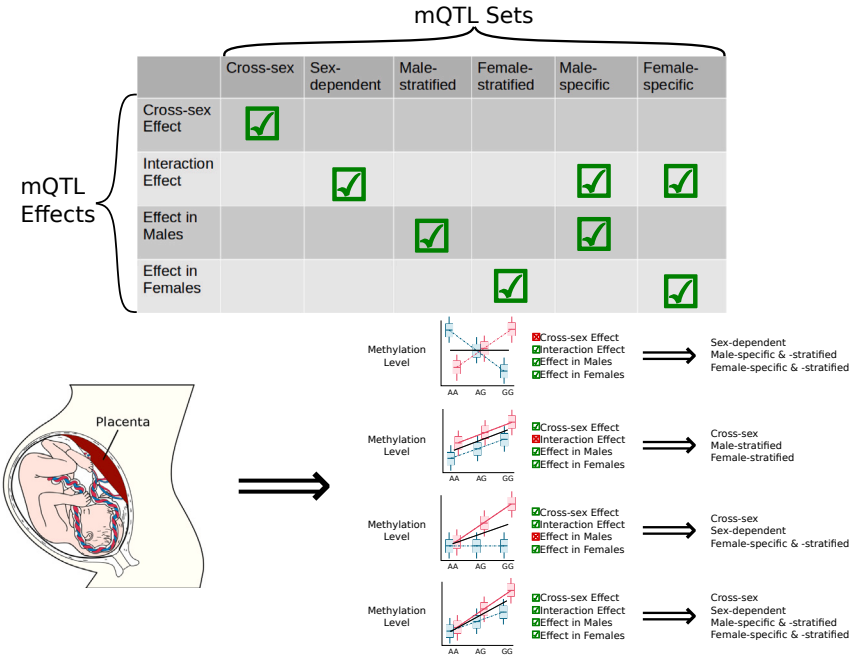


Figure 1. Defining placental mQTL sets from mQTL effects

We defined mQTL sets as: (i) cross-sex, which have an effect independent of sex; (ii) sex-dependent, which have an effect that differs between male and female samples (a genotype by sex interaction effect); (iii) male-stratified, which have an effect in males; (iv) female-stratified, which have an effect in females; (v) male-specific, which have an effect in males that also differs from the effect in females; and (vi) female-specific, which have an effect in females that also differs from the effect in males.

differs from the effect in females (intersection of sets (ii) and (iii)); and (vi) female-specific, which have an effect in females that differs from the effect in males (intersection of sets (ii) and (iv)).

We counted all CpG sites with at least one mQTL. Since multiple SNPs can be associated with a CpG site, we indirectly controlled for LD between SNPs by counting at the level of CpG sites,[9,11,32,33] and by using a stringent threshold for declaring statistically significant mQTL (Bonferroni-corrected $p < 0.05$).[27] Note that we intentionally did not report the number of independent mQTL, as there is no standard on how to compute independence,[9,11,32,34] and none of our downstream enrichment analyses required independent mQTL. In total, we found 49,252 CpG sites with a cross-sex mQTL and 2,489 CpG sites with a sex-dependent mQTL. These significant cross-sex mQTL were highly replicated in a recently released database of mQTL identified in 368 fetal placentas,[35] having a $\pi_1 = 0.987$ and a Spearman correlation of $\rho = 0.971$.

Of the CpG sites with sex-dependent mQTL, we found 351 CpG sites with a male-specific mQTL, and 255 CpG sites with a female-specific mQTL. Of the 351 CpG sites with a male-specific mQTL, 185 (53%) also had a cross-sex mQTL, with 68 sites sharing the same top-associated SNP. Of the 255 CpG sites with a female-specific mQTL, 153 (60%) also had a cross-sex mQTL, with 65 sites sharing the same top-associated SNP. Lastly, 75 CpG sites had both male- and female-specific mQTL effects (see Figure 1 for an example of these types of mQTL effects), of which 74 (99%) also had a cross-sex mQTL. Of these 75 sites, 67 had the same top-associated SNP. See Figures S1F–S1I for overlap between significant mQTL sets and for a comparison with male- and female-stratified mQTL.

To benchmark the number of mQTL we found against the number found in other studies, we called mQTL at a less-stringent FDR <0.05 and assessed inflation via Q-Q plots. We detected 211,862 CpG sites with at least one cross-sex mQTL at FDR<0.05, which is consistent with what is observed for mQTL in GTEx (108,844-206,802 depending on the tissue).[9,11] We generated Q-Q plots of all meta-analyzed mQTL p values (Figures S1J–S1M) and found little systematic inflation in the larger p values, whereas inflation for smaller p values in the upper ranges of these plots is expected and is consistent with deviation from the null observed in other mQTL studies.[13,36]

We further investigated the difference in male- vs. female-specific effect sizes (the strength of SNP-CpG associations). For computing correlation between effects, we considered all SNP-CpG pairs with either male- or female-specific mQTL, which includes multiple mQTL for certain CpG sites (8.2 SNPs per CpG site on average in male-specific mQTL, a median of 3 SNPs; 7.8 SNPs per CpG site on average in female-specific mQTL, a median of 2 SNPs). We found that effects were highly correlated between male- and female-specific mQTL (Spearman's $\rho = 0.87$, Figure S1N) and that male- and female-specific mQTL shared the same direction of effect at all CpG sites. In comparison, across adult tissues in GTEx V8 the correlation between male- and female-specific mQTL was 0.78 (Spearman's ρ), and 58% of associations were in the same direction in both males and females.[19] Whether sex-dependent mQTL share the same direction of effect in both males and females in other tissues would be worth exploring. The average effect in male-specific mQTL was smaller than the effect observed in

Table 1. Counts of CpG sites with at least one mQTL in each mQTL set and their corresponding gene

mQTL Set	No. CpG Sites	No. Corresponding Genes
Cross-sex	49,252	12,746
Sex-dependent	2,489	976
Male-specific	351	201
Female-specific	255	153
Male-stratified	31,384	9,507
Female-stratified	25,180	8,264

Corresponding genes were identified using Illumina's HumanMethylation450k array annotation. CpG sites are not necessarily exclusive to each set (see Figure 1).

female-specific mQTL (mean effect size 0.021 in males, 0.035 in females, Wilcoxon rank-sum test two-sided $p < 0.026$, Figure 2A; Table S3). This difference was larger when we considered the absolute value of the effect size (mean absolute beta 0.067 in males vs. 0.079 in females, Wilcoxon rank-sum test two-sided $p < 2.3e-11$, Figure S1O; Table S3). This effect size difference did not appear to be driven solely by differences in mQTL effect sizes on the X chromosome: female-specific mQTL had a larger mean signed effect size than male-specific mQTL in 6 autosomes, and a larger mean absolute effect size in 9 autosomes (Wilcoxon rank-sum test one-sided, Bonferroni-Holm corrected $p < 0.05$, Figure 2B; Table S3). This sex difference in effect sizes could be due to sex differences in responses to prenatal environmental stresses,[37] which would increase or decrease inter-individual variation in DNAm dependent on sex, ultimately affecting mQTL effect sizes.

Placental methylation quantitative trait loci are enriched in regions controlling gene expression and primarily occur at CpG sites with intermediate levels of DNAm

In order to establish whether sex-dependent mQTL occur in distinct genomic regions compared to cross-sex mQTL, we applied GARFIELD[38] to the minimum mQTL p value for each SNP (STAR Methods). Briefly, GARFIELD tests whether a given set of SNPs associated with a particular phenotype are enriched in a set of genomic regions defined from functional experiments from the ENCODE Project Consortium[39–41] and the NIH Roadmap Epigenomics Consortium.[42] Importantly, GARFIELD accounts for both LD between SNPs and redundancy in annotations. It does not overestimate enrichment due to many correlated mQTL in the same genomic region, and it penalizes annotations with similar enrichment across all mQTL in a set.

As expected, the enrichment of each mQTL set in each annotation was typically larger for sets with a larger number of mQTL, with cross-sex having the most mQTL, followed by male-stratified mQTL, female-stratified mQTL, and then sex-dependent mQTL (Table 1). Male- and female-stratified mQTL were most enriched within the same annotations as in cross-sex mQTL in the same order, albeit with smaller estimates. Male- and female-specific mQTL were excluded here, as GARFIELD uses only a single p value for each SNP and therefore could not handle the main and interaction term p values that define the male- and female-specific analyses. Across ChromHMM annotations in human cell lines (Figure 3A)[40,43] we observed that cross-sex placental mQTL were more enriched in transcription start sites (TSS) than in other chromatin states (TSS odds ratio (OR) of 2.68 vs. a mean OR = 1.60 in the remaining selected states), which was not observed for sex-dependent mQTL (TSS OR = 1.22 vs. a mean OR = 1.47 across other annotations). Within individual histone modifications averaged across experiments from ENCODE,[38,39] we found that cross-sex mQTL were most enriched in regions with H3K9 acetylation (mean OR = 2.50), followed by H3K4 trimethylation (mean OR = 2.41), both of which are indicative of active gene promoters (Figure 3B).[44] For sex-dependent mQTL, the highest enrichment was in H4K20 monomethylation (OR = 1.69), which is associated with transcription activation.[45] Overall, enrichment tracked strongly with the size of each mQTL set, and results suggested minor, but potentially meaningful, differences in the gene regulatory function of sex-dependent vs. cross-sex mQTL.

Next, we investigated the weighted-mean DNAm of CpG sites with at least one cross-sex, sex-dependent, male-specific, or female-specific mQTL. Both cross-sex and sex-dependent mQTL were primarily associated with CpG sites with intermediate levels of DNAm, and 72% of cross-sex and 71% of sex-dependent CpG sites had a weighted mean beta between 0.2 and 0.8 (Figure 3C). These proportions are much larger than what is observed in other human primary cell types (2% of sites with intermediate DNAm)[46] and is not surprising given that up to 40% of CpG sites in the placenta are intermediately methylated.[47,48] CpG sites corresponding to male- and female-specific mQTL were also primarily intermediately methylated (60% and 55%, respectively), but 31% of female-specific CpG sites had a weighted mean beta>0.8, compared to 21% of male-specific sites, and 16% of cross-sex and sex-dependent sites.

We next investigated how the level of DNAm of CpG sites with mQTL was spread across regions related to gene regulation. Specifically, we looked at the weighted-mean DNAm of these CpG sites relative to other CpG sites and gene regions, as included on Illumina's HumanMethylation450k array annotation (STAR Methods). DNAm occurs primarily at CpG dinucleotides, which cluster into CG repeats called CpG Islands (CGI) that span 300–3000 bp.[49] Illumina denotes shores as the region of DNA within 0-2kb of a CGI, shelves as the region of DNA within 2-4kb of a CGI, and denotes the remaining CpG sites that are greater than 4kb from a CGI as open sea regions. The terms north and south denote whether a shore or shelf comes before the 5′ end of a CGI, or after the 3′ end of a CGI. CGIs are a part of roughly 40% of gene promoters, and the position of individual CpG sites relative to these islands is thought to be related to their impact on transcription.[50] For example, high DNAm in CpG shores and shelves is associated with higher nearby gene expression.[51]

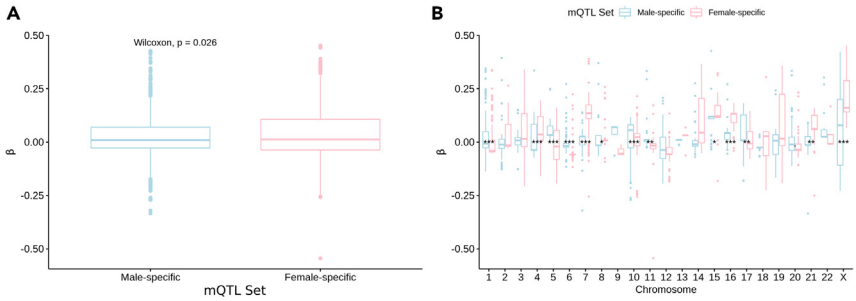


Figure 2. mQTL detected across chromosomes

(A) The difference in mean effect size in male- vs. female-specific mQTL across all chromosomes and (B) separately by chromosome, with Wilcoxon rank-sum test p values after Bonferroni-Holm correction. .: $p \leq 0.1$; *: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$.

We observed a similar distribution of DNAm levels per CGI region across our four mQTL sets of interest. The majority of highly methylated (beta >0.8) CpG sites were in open sea regions (51% of cross-sex, 52% of sex-dependent, 57% of male-specific and 52% of female-specific), followed by regions within shores or shelves (36% of cross-sex, 36% of sex-dependent, 35% of male-specific, and 38% of female-specific), and the remaining CpG sites within CGIs themselves (13% of cross-sex, 12% of sex-dependent, 10% of male-specific, and 10% of female-specific) (Figure 3D). Relative to gene regions, we also observed similar patterns of DNAm levels across all four mQTL sets. Highly methylated CpG sites were primarily located in gene bodies (51% of cross-sex, 51% of sex-dependent, 56% of male-specific, and 53% of female-specific), which typically indicates active transcription of those genes. Intermediately methylated CpG sites were evenly split between gene bodies and intergenic regions (37% and 33% for both cross-sex mQTL and sex-dependent mQTL, 37% and 30% for male-specific mQTL, and 38% and 30% for female-specific mQTL respectively). Finally, lowly methylated CpG sites were primarily within 1500bp of the TSS of genes (39% in cross-sex, 39% in sex-dependent, 41% of male-specific, and 43% of female-specific), indicating active transcription (Figure 3E).

Sex-dependent and cross-sex methylation quantitative trait loci show similar patterns of tissue specificity

To assess the tissue specificity of the placental mQTL we identified, we first calculated the replication (π_1) of placental mQTL across mQTL in two other prenatal tissues, umbilical cord blood and fetal brain, identified in independent datasets (Figure 4A; STAR Methods).[32,34] The umbilical cord is biologically inert and primarily functions in nutrient transport. The umbilical cord and placenta bear no developmental or functional relationship with each other, and we include the umbilical cord blood here as a negative control.[52,53] In contrast, we compare mQTL in the placenta and fetal brain to assess the developmental relevance of the placenta. While we would have liked to compare placental mQTL against mQTL from other fetal tissues too, no other datasets existed. We observed a relatively large proportion of placental mQTL overlapping cord blood mQTL ($\pi_1 = 0.76$) and an even larger proportion of placental mQTL overlapping fetal brain mQTL ($\pi_1 = 0.84$). Effect sizes of placental mQTL correlated poorly with those in cord blood (Spearman's $\rho = -0.31$) but largely shared the same direction of effect in the fetal brain (Spearman's $\rho = 0.65$), which suggests a higher degree of similarity between mQTL in placenta vs. in the fetal brain.

To assess the specificity of the overlap between placental and fetal brain mQTL, we next quantified the enrichment of our mQTL sets in DNAse1 hypersensitivity (DHS) sites in fetal tissues, as well as in adult tissues and cell lines studied in the ENCODE Project Consortium and the NIH Roadmap Epigenomics Consortium,[39–42] using GARFIELD. Enrichment was highest in the fetal membrane for both cross-sex mQTL (OR = 2.82) and sex-dependent mQTL (OR = 1.58), followed by the fetal placenta (Figure 4B, OR = 2.74 and OR = 1.43 in cross-sex vs. sex-dependent mQTL respectively). Of note, our placental mQTL were no more enriched in fetal brain DHS sites compared to other fetal tissue DHS sites. For adult tissues and cell lines (Figure 4C), cross-sex mQTL were most enriched in myometrium, which is tissue from the uterine wall (OR = 2.71), followed by bone (OR = 2.68), liver (OR = 2.64), and colon (OR = 2.64). Slight differences were found in sex-dependent mQTL, with enrichment being highest in bone (OR = 1.56), blastula (an early stage of embryonic development, OR = 1.51), colon (OR = 1.49), and liver (OR = 1.47). Overall, these results suggest that cross-sex and sex-dependent mQTL are enriched in regions that are active in the same sets of tissues.

Male- and female-specific placental methylation quantitative trait loci are more enriched for the heritability of immune-related and growth-related traits than are cross-sex placental methylation quantitative trait loci

In this study, we were particularly interested in how placental mQTL, including those specific to males or females, contributed to the genetic risk of childhood traits and conditions. We included GWAS of 18 childhood traits (Table S5; Figure 5A),[54,55] as well as a GWAS of maternal pre-eclampsia (the fetal genetic effect on maternal pre-eclampsia risk),[56] which is a placentally mediated condition. We then used S-LDSC to estimate the proportion of SNP-heritability (h^2_{SNP}) of these 19 complex traits that was explained by our placental mQTL sets and report the enrichment of each set: the proportion of h^2_{SNP} explained divided by the fraction of all SNPs present in the mQTL set (Table S6; STAR Methods). Enrichments larger than one mean that on average, an SNP from an mQTL set explains a larger proportion of h^2_{SNP} of a trait than expected by chance. In this way, we can compare the relative importance of SNPs in each mQTL set, despite each set containing different fractions of all SNPs included in the analysis. Lastly, in order to summarize enrichment across related traits, we also meta-analyzed across

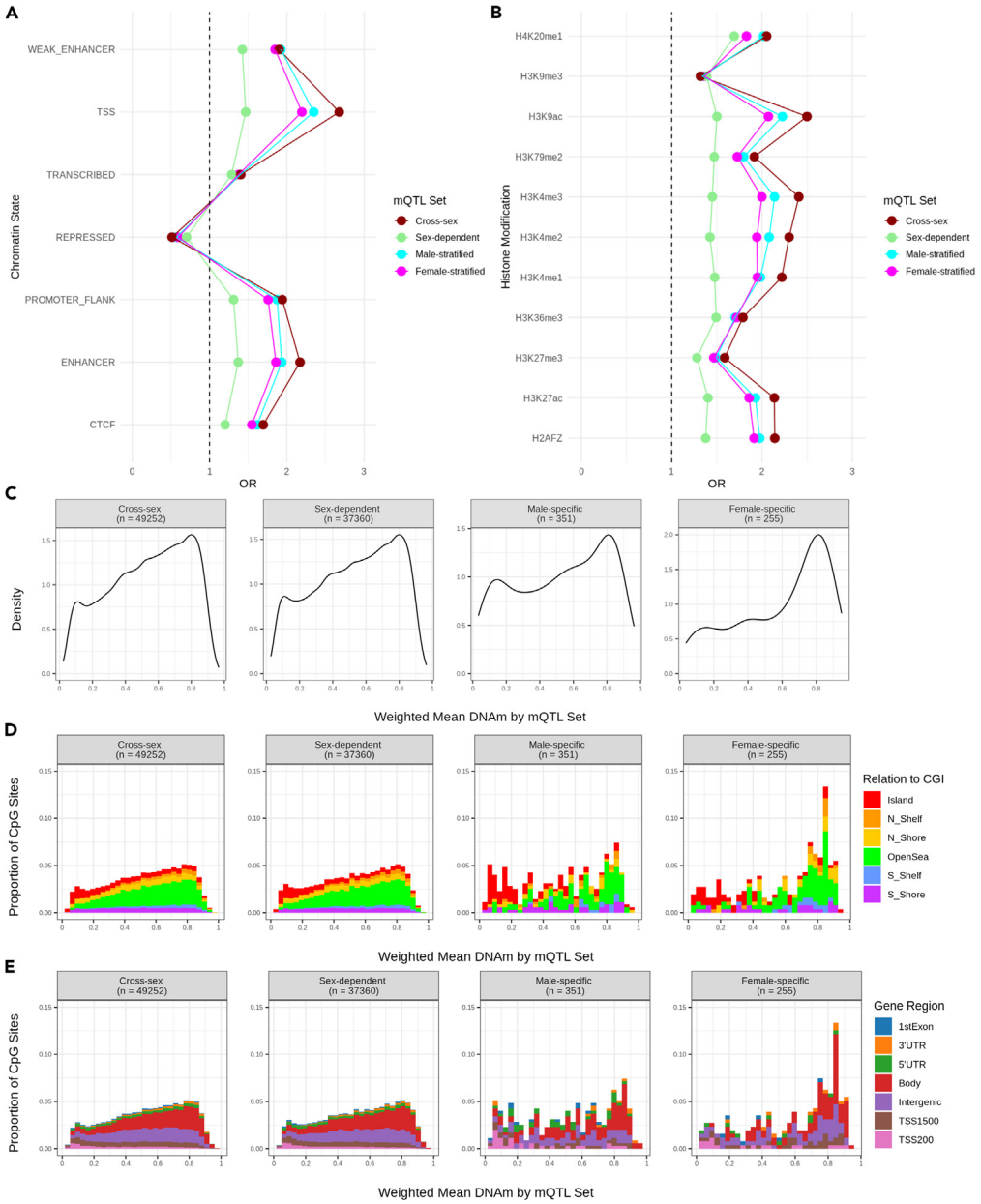


Figure 3. Functional role of placental mQTL

Enrichment of mQTL sets in chromatin states (A) averaged across human stem cell lines and (B) in histone modifications averaged across tissues in ENCODE. (C) Density distribution of weighted mean DNAm, stratified by mQTL set. Proportion of CpG sites by weighted mean DNAm values, stratified by mQTL set, and visualized according to (D) position relative to CpG island and (E) gene region.

neuropsychiatric (N = 7), immune-related (N = 3), and growth-related (N = 8) traits. These meta-analyzed enrichment estimates can be thought of as the average enrichment of mQTL sets across traits in each category (STAR Methods).

We found that placental mQTL were not enriched for h^2_{SNP} of any neuropsychiatric traits, either individually, or across all tratits in meta-analysis (see Table S6 for trait-level S-LDSC results; Table S4 for meta-analyzed results). Conversely, we found an enrichment (cross-trait FDR <0.05) of cross-sex placental mQTL in type 1 diabetes, birth weight, and pubertal growth start (Figure 5B). We also observed an enrichment for h^2_{SNP} of child onset asthma, birth length, and late pubertal growth at a within-trait FDR <0.05.

Male- and female-specific mQTL were enriched for h^2_{SNP} of birth weight and pubertal growth start (FDR <0.05), as well as of childhood BMI (within-trait FDR <0.05). In meta-analysis, female-specific mQTL were more enriched than cross-sex mQTL for h^2_{SNP} of immune-related traits (1.75, SE = 0.249, for female-specific mQTL vs. 1.29, SE = 0.06 for cross-sex mQTL) (Figure 5C). Additionally, both male- and female-specific

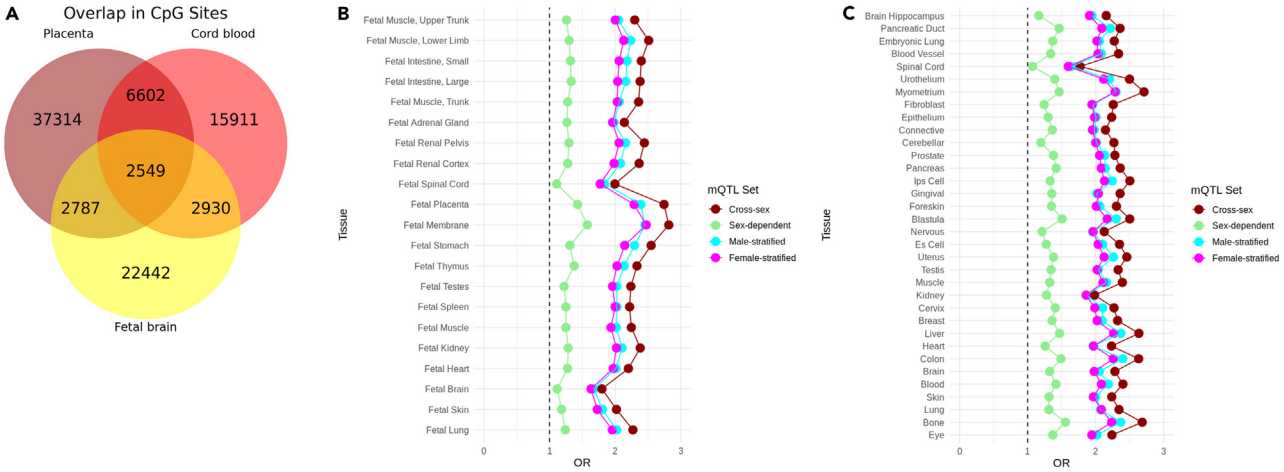


Figure 4. Enrichment of placental mQTL across tissues

(A) Overlap in CpG sites with at least one mQTL across prenatal tissues: umbilical cord blood from the ARIES study and whole fetal brain tissue from the HDBR study. Enrichment of placental mQTL in DNAse1 hypersensitivity sites in (B) fetal tissues from the NIH Roadmap Epigenomics Consortium and (C) adult tissues from the ENCODE Project Consortium and the NIH Roadmap Epigenomics Consortium. Estimates are averaged over samples from the same tissue and stem cell lines.

mQTL were more enriched than cross-sex mQTL for h^2_{SNP} of growth-related traits (1.66, SE = 0.167 for male-specific mQTL, 1.78, SE = 0.137 for female-specific mQTL, and 1.20, SE = 0.08 for cross-sex mQTL).

We also examined sex-stratified h^2_{SNP} estimates and the proportion of sex-stratified h^2_{SNP} explained by our mQTL sets. Since the sex-stratified GWAS sample sizes were smaller, fewer traits passed our quality control checks (STAR Methods): four neuropsychiatric traits (two with both sexes), one immune-related trait (with both sexes), and three growth-related traits (two with both sexes) (Figure S2A). Male- and female-specific mQTL were enriched for the sex-stratified h^2_{SNP} of pubertal growth start at an FDR <0.05 (Figure S2B, for meta-analyzed enrichment estimates see Figure S2C).

Overall, these results show that placental mQTL show larger enrichments for h^2_{SNP} of immune- and growth-related traits, with notably high enrichment in anthropometric traits such as birth weight. Moreover, male- and female-specific mQTL showed a larger enrichment for immune- and growth-related traits than either cross-sex, male- or female-stratified mQTL, which highlights the importance of considering sex-dependent mQTL in measuring the enrichment of h^2_{SNP} for these traits.

Placental methylation quantitative trait loci colocalize primarily with growth and immune-related traits, with additional CpG sites colocalizing with male- and female-specific placental methylation quantitative trait loci

We subsequently applied colocalization analyses using *coloc* (STAR Methods)[57] to assess whether any of the mQTL we identified were also likely to be associated with the GWAS loci of 18 childhood traits and maternal pre-eclampsia. We conducted this analysis for cross-sex and male- and female-stratified mQTL p values, for GWAS loci within a 150kb window centered on each CpG site. As in the GARFIELD analysis, we did not compute colocalization with male- and female-specific mQTL, since coloc uses only a single p value for each SNP and therefore could not handle both the main and interaction term p values that define the male- and female-specific analyses. Moreover, none of the GWAS we included report SNP*sex interaction term p values for use in coloc.

We found a considerable number of colocalized CpG sites in several of the childhood traits with moderate enrichment in S-LDSC, such as child onset asthma, type 1 diabetes, and birth weight (Figure 6A; Table S7). We also observed a high number of colocalized sites for schizophrenia (SCZ), which notably has a large number of GWAS loci.[58] Loci from each of these four traits showed colocalization with male- and female-stratified mQTL that was not detected with cross-sex mQTL, and we label these male- and female-specific colocalized mQTL (see Figures 6C–6H for examples of these distinctions). The top GWAS locus per each region assessed did not necessarily correspond to the top mQTL in this region, despite high confidence that the region contained an SNP responsible for both changes in DNAm and trait risk (STAR Methods).

In Figure 6B, we show the correlation between Z-scores for the SNP effect on DNAm and the SNP effect on the trait, for all colocalized mQTL/trait pairs and for each mQTL set. Larger correlations were observed for female-specific colocalized mQTL across traits, which could reflect sex biases in the heritability of the underlying GWAS traits, or the larger placental mQTL effect estimates we identified in females vs. males. Overall though, we observe little correlation between the Z-scores, suggesting that both increases and decreases in genetically mediated DNAm associate with complex trait risk.

Focusing on child-onset asthma, we assessed the biological relevance of the genes containing CpG sites that colocalized with cross-sex vs. male-specific vs. female-specific mQTL. Males have a higher prevalence of asthma up to age 13, and females have a higher prevalence of asthma in adulthood, often with greater severity of symptoms.[59] Genetic variants in the human leukocyte antigen (HLA) region are frequently

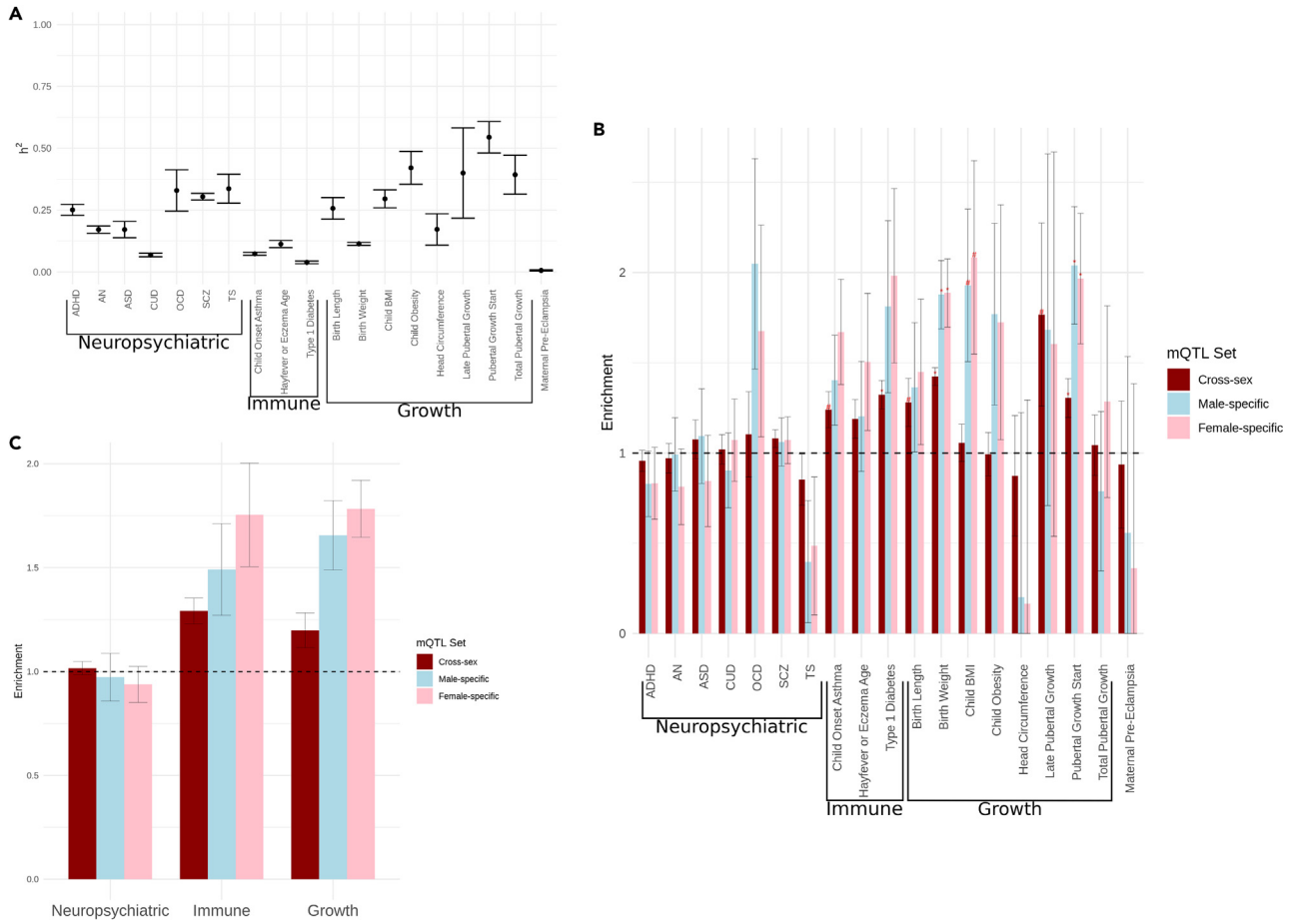


Figure 5. Placental mQTL enrichment in 19 complex traits

(A) SNP heritability (h^2_{SNP}) for each trait estimated by LD score regression.

(B) Enrichment of placental mQTL sets for h^2_{SNP} of each trait, accounting for 97 standard baseline regulatory effects. Enrichments with a within-trait FDR <0.05 are marked with #, whereas enrichments significant at an FDR <0.05, accounting for each mQTL annotation and GWAS assessed, are marked with an asterisk.

(C) Meta-analyzed enrichment estimates across traits in neuropsychiatric, immune-, and growth-related GWAS categories.

Error bars are the standard error of each estimate.

associated with both child- and adult-onset asthma, although in recent GWAS only 16 of 123 child-onset asthma loci (13%), and 10 of 56 adult-onset asthma loci (18%), were within the HLA region on chromosome 6.[60,61] We focus on genetic variants outside of chromosome 6 and the HLA region to avoid challenges in interpretation due to long-range LD.

A total of 192 CpG sites with placental mQTL colocalized with child onset asthma loci, of which 128 were cross-sex, 19 were male-specific, and 45 were female-specific. Of the 128 colocalizing cross-sex mQTL, 83 (65%) were associated with DNAm on CpG sites outside of chromosome 6. We mapped these to genes and then used Enrichr[62–64] to identify pathways enriched for these genes at an adjusted p value <0.05, at a minimum overlap of three genes across the three main genome ontology annotations. We found an enrichment for the regulation of type I interferon gamma production, a physiological response that contributes to inflammation and to asthma (adjusted p < 2.7e-4, GO:0032479, overlapping *TRAF3*, *IRF1*, *STAT6*, *POLR3H*, and *NFKB1*).[65] In comparison, of the 19 colocalized male-specific CpG sites, 11 (58%) were outside of chromosome 6, and of 45 colocalized female-specific CpG sites, 4 (9%) were outside of chromosome 6. The number of genes for which these CpG sites were annotated was too small for meaningful pathway enrichment analysis.

DISCUSSION

In this study, we conducted a meta-analysis of mQTL across two term placental studies (NICHD and RICHS) and a comprehensive sex-dependent mQTL analysis. We demonstrated that sex-dependent mQTL are located at partially distinct CpG sites from cross-sex mQTL. In contrast to blood mQTL, which are associated with mainly lowly or highly methylated CpG sites,[9] placental mQTL tended to be associated with intermediately methylated CpG sites.[47] Likewise, compared to previously reported mQTL in blood, placental mQTL were more enriched at the

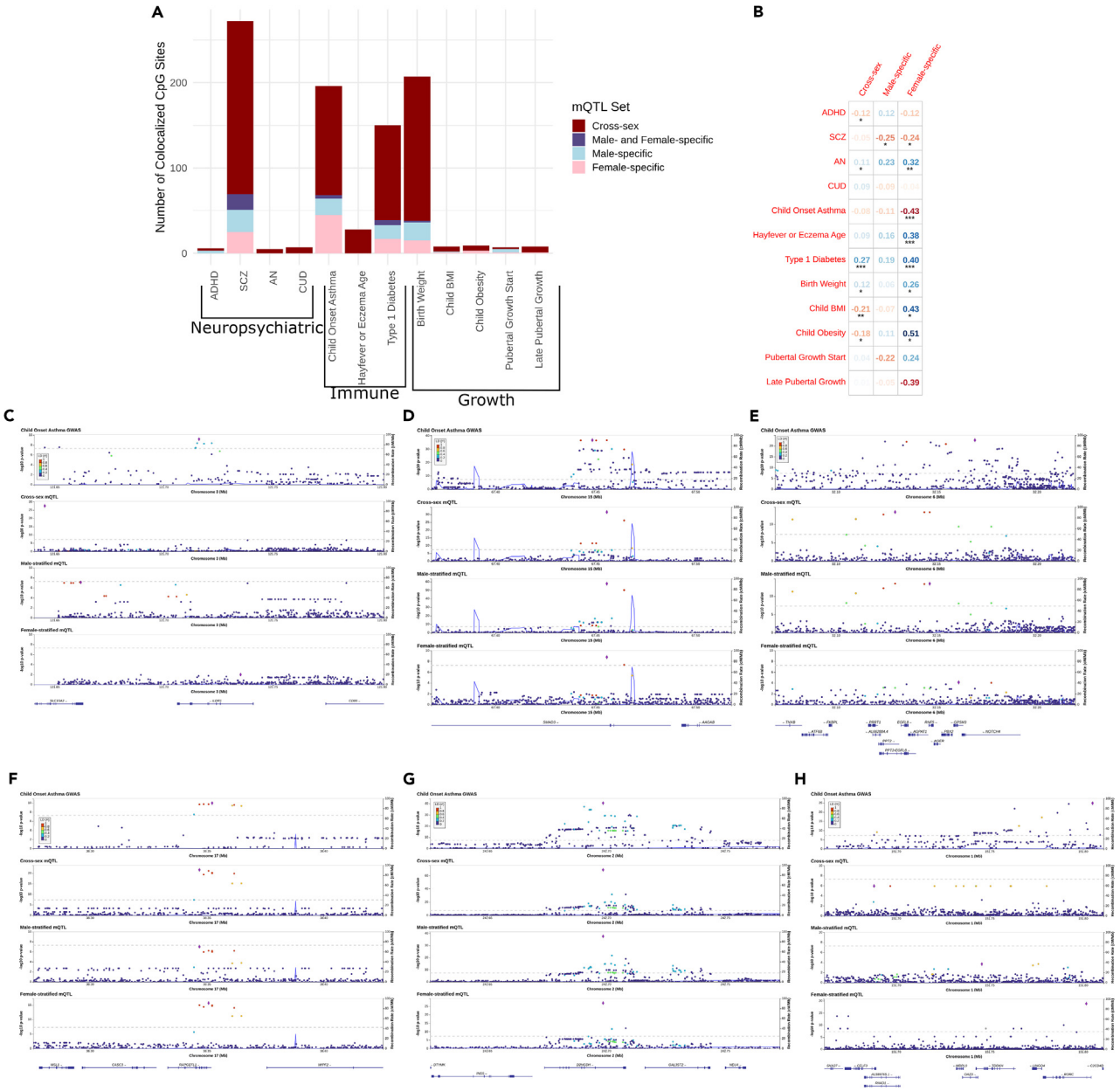


Figure 6. Colocalization of placental mQTL with GWAS loci of complex traits

(A) The number of CpG sites with an mQTL that colocalized with at least one GWAS locus at a posterior probability (H_4) > 0.9.

(B) The correlation between Z-scores for colocalized cross-sex, male- and female-specific mQTL and GWAS SNPs, with Spearman's ρ displayed for ρ > 0.01 and ρ < −0.01 and p values denoted as follows: *: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$.

(C–H) LocusZoom plots for different colocalization scenarios that occurred when testing for the colocalization of male- or female-stratified vs. cross-sex mQTL and loci from a GWAS of child onset asthma. We show colocalization in: (C) cross-sex mQTL only (cg16796354, within *IDLR1*); (D) cross-sex and male-stratified mQTL (cg24032190, within *SMAD3*); (E) male-stratified mQTL only (cg16689962, within *AGPAT1*); (F) cross-sex and female-stratified mQTL (cg19063856, within RAPGEFL1); (G) cross-sex, male- and female-stratified mQTL (cg02698622, within *D2HGDH*); (H) female-stratified mQTL only (cg01717973, within *OAZ3*).

TSS of genes and were more depleted in repressed regions, but otherwise showed similar enrichment to blood mQTL in chromatin states, intergenic regions, and regions with known histone modifications.[9,66]

Additionally, we showed that cross-sex mQTL were enriched for h^2_{SNP} of immune- and growth-related traits and that male- and female-specific mQTL were more enriched than cross-sex mQTL across traits in these GWAS categories. Perhaps most importantly, we found several CpG sites with only male- or female-specific colocalization with GWAS loci. Thus, sex-specific mQTL appear to capture some underlying aspect of heritable trait risk that is not being captured by cross-sex placental mQTL alone.

Motivated by the enrichment of placental mQTL for the h^2_{SNP} of immune-related traits, we homed in on asthma and mapped the colocalized asthma GWAS and mQTL to genes and pathways. Epidemiological studies have associated abnormal placental morphology, low birth length, and low birth weight to asthma risk.[67] Our study suggests that genetic risk for child onset asthma is meaningfully enriched for cross-sex and male- and female-specific mQTL, and we provide these associations as a resource for potentially understanding the biological consequences of child onset asthma loci.

Past eQTL[13,16,17] and integrative molQTL analyses in the placenta,[15,68] have identified genetically regulated placental gene expression that associates with several traits, including child BMI, birth weight,[17,68] asthma, and type 2 diabetes.[15] All of these traits have evidence of genetic sexual dimorphism, either having a male-female r_g lower than 1, or a difference in male vs. female effect size at risk loci.[59,69,70] Likewise, the prevalence and manifestation of these traits differ between males and females throughout the life course.[59,71,72] The placenta is a sexually dimorphic tissue. We and others have found as many as 2,745 CpG sites (annotated to 582 genes) associated with placental sex, which annotated to genes that are primarily related to immune function and growth factor signaling.[20,21] Given these molecular sex differences and their potential role in mediating complex traits, sex-dependent effects should remain a focus of future studies of the placenta.

Overall, this study demonstrates that the genetic regulation of placental DNAm is partially sex-dependent. Sex-dependent placental mQTL can occur in distinct functional genomic regions from cross-sex placental mQTL, suggesting they have a unique role in gene regulation. Placental mQTL explain a significant proportion of the h^2_{SNP} across conditions related to immune function and growth, and both male- and female-specific mQTL are more enriched than cross-sex mQTL across these conditions, despite being far fewer in number. Trends we observed in the enrichment of placental mQTL translated to higher instances of colocalization between mQTL and GWAS loci from childhood onset traits and conditions, and using male- and female-specific colocalization allowed us to detect 216 CpG sites (annotated to 98 genes) that otherwise did not show sufficient evidence of colocalization with cross-sex mQTL. Taken together, our findings demonstrate that the careful consideration of sex in mQTL analyses has the potential to provide additional information about the basis of complex traits, particularly when the tissue, molecular features, and traits queried are sexually dimorphic.

Limitations of the study

There are several limitations to this study. First, our understanding of the tissue-specificity of placental mQTL effects is limited, as efforts are still underway to characterize mQTL across tissues.[73] However, based on existing datasets, we showed that cross-sex placental mQTL are quite similar to mQTL in other prenatal tissues ($\pi_1 = 0.76$ in umbilical cord blood and $\pi_1 = 0.84$ fetal brain, see Results). By this metric, placental mQTL are more similar to prenatal tissue mQTL than placental eQTL are to eQTL in 44 adult tissues from GTEx (ranging from $\pi_1 = 0.32$ in cerebral hemisphere to $\pi_1 = 0.69$ in fibroblasts; previously computed using NICHD placental eQTL, N = 80).[13]

Second, given that this analysis was conducted in bulk-tissue samples of placenta, we were not able to fully address the role of cell type in our results beyond the use of principal components to account for interindividual differences in estimated cell type proportions (STAR Methods, Figures S3A–S3D). However, this only addresses the possibility that mQTL are due to systematic differences in cell-type proportion between samples, and would not correct for mQTL being biased to individual cell types. This exact scenario was observed in one cell-type specific eQTL analysis in immune cells, which found that 37% of genes with an eQTL were only detected in a single cell type.[74] Thus, molQTL derived from bulk tissue may be biased to certain cell types highly present in a given study, and should not necessarily be considered as ubiquitous throughout their respective tissue.

Third, sex-dependent mQTL are less numerous than cross-sex mQTL due to the increased power required to detect interaction effects.[75] As a result, in all of our experiments comparing sex-dependent mQTL (and by extension, male- and female-specific mQTL) to cross-sex mQTL, sex-dependent mQTL had larger standard errors, which hampered statistical inference and ultimately hampered our ability to discuss small differences in heritability captured by different mQTL sets. In particular, this made it difficult to draw conclusions from the τ^* metric, which can be more sensitive than the enrichment metric to small differences in mQTL annotations in single traits.[24] We computed τ^* here to make it easier for readers to detect large differences between S-LDSC enrichment of placenta mQTL vs. mQTL from other tissues. As we accrue larger sample sizes in molecular studies of placenta and other tissues, we will be better equipped to identify robust differences in the contribution of male- and female-specific mQTL to the h^2_{SNP} of complex traits. We were also limited in our ability to test male- and female-specific mQTL for enrichment in sex-stratified GWAS corresponding to their respective sex. Although sex-stratified GWAS results are increasingly being shared,[76] these GWAS are underpowered. Only the GWAS of ADHD, ASD, SCZ, hayfever or eczema age of onset, pubertal growth start, and total pubertal growth were sufficiently powered to meet our criteria for inclusion in the enrichment analyses in either sex. For these traits, we found higher enrichments of male- and female-specific mQTL as compared to enrichments found for sex-pooled GWAS, which emphasizes the need for large sex-stratified GWAS.

Fourth, although we compared mQTL across ancestries in NICHD before proceeding with an ancestry-pooled meta-analysis, and we replicated our mQTL results in an independent sample, residual population stratification might still have affected our results. Multiple strategies exist to account for diverse genetic ancestry in molQTL mapping, and we used the same approach (sample-wide genotyping PCs) as used in GTEx (15% of GTEx participants had significant non-European admixture).[11] An alternative approach would have been controlling for local ancestry using linear mixed-effect models, which in previously reported simulations provided a small but significant reduction in the false positive rate in instances where simulated eQTL were ancestry-specific.[77] Using this model to detect sex-specific effects would require new methodology, which was beyond the scope of our study. Another alternative approach would have been to first compute mQTL within each ancestry group, and then compute all downstream analysis within each ancestry group separately.[25] This ancestry-stratified approach might

have resulted in more instances of colocalization, but given our sample size, stratifying along each ancestry would have greatly reduced our power to detect sex-specific mQTL, which was our primary focus.

We also took care when integrating our ancestrally diverse placental mQTL with GWAS summary statistics, which were derived from samples of primarily European genetic ancestry. S-LDSC relies on, and has largely only been validated with, LD scores computed over 1000 Genomes European subjects.[22,23] In practice, this means that for any functional annotation, even one generated in an ancestrally diverse sample, we do not violate assumptions of S-LDSC as long as the LD scores for that annotation are computed with respect to a reference panel with an ancestry that matches the ancestry of each GWAS considered.[24] However, this does mean that our S-LDSC results do not take into account disease risk present in non-European populations. Meanwhile, in the case of colocalization analysis, recent work with eQTL in GTEx suggests that within populations with high non-European admixture, a handful of colocalizations are missed when not taking local, ancestry-specific genetic variation into account.[25] Although there is past evidence showing that mQTL are consistent across ancestry groups, and our work in NICHD shows moderate overlap,[78] this research bears repeating with larger datasets for investigating ancestry-specific genetic variation.

Lastly, our mQTL mapping strategy focused on single CpG sites. DNAm can be correlated across multiple neighboring CpG sites, as in differentially methylated or co-methylated regions,[79–83] and these correlated sites could be of interest in future mQTL studies.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Genotype data processing and quality control
  - DNA methylation array processing and quality control
  - Criteria for GWAS summary statistics
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Mapping ancestry-stratified mQTL in NICHD
  - Mapping cross-sex, sex-dependent, male-stratified, and female-stratified mQTL
  - Assessing overlap of placental mQTL with mQTL in other prenatal tissues
  - Assessing enrichment of placental mQTL in regulatory regions across tissues using GARFIELD
  - Linkage disequilibrium score regression analyses
  - Colocalization of placental mQTL and GWAS loci

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.109047.

ACKNOWLEDGMENTS

This project was supported by a British Columbia Children's Hospital Research Institute Catalyst Grant. JKD is supported by a NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation, and is a Michael Smith Health Research BC Scholar.

phs001586: Rhode Island Child Health Study (RICHS): This work was supported by the National Institutes of Health [NIH-NIMH R01MH094609, NIH-NIEHS R01ES022223, NIH-NIEHS PO1ES022832, NIH-NIEHS R24ES028507, NIH-NIEHS R21ES028226, and NIH-NIEHS R01ES025145]. phs001717: This research was supported by the Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health (Contract Numbers: HHSN275200800013C; HHSN27500006; HHSN275200800003I; HHSN275200800014C; HHSN275200800012C; HHSN275200800028C; HHSN275201000009C). We also want to thank the Genomics Shared Resource of Roswell Park Cancer Institute, supported by National Cancer Institute (NCI) grant P30CA016056.

This research was supported in part through computational resources and services provided by Advanced Research Computing at the University of British Columbia (https://doi.org/10.14288/SOCKEYE).

AUTHOR CONTRIBUTIONS

Conceptualization, WC, JKD, and SM; methodology, WC, YPP, and JKD.; software, WC, AI, JD, and VY; validation WC, AI, JD, and VY; formal analysis, WC; resources, JKD, WR, and SM; data curation, WC, CM, and FD; writing - original draft, WC; writing - review and editing, JKD, WR, and AI; visualization, WC; supervision, JKD and SM.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 9, 2022
Revised: June 19, 2023
Accepted: January 23, 2024
Published: January 26, 2024

REFERENCES

1. King, E.A., Davis, J.W., and Degner, J.F. (2019). Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. PLoS Genet. *15*, e1008489. https://doi.org/10.1371/journal.pgen.1008489.
2. Nelson, M.R., Tipney, H., Painter, J.L., Shen, J., Nicoletti, P., Shen, Y., Floratos, A., Sham, P.C., Li, M.J., Wang, J., et al. (2015). The support of human genetic evidence for approved drug indications. Nat. Genet. *47*, 856–860. https://doi.org/10.1038/ng.3314.
3. Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., Reynolds, A.P., Sandstrom, R., Qu, H., Brody, J., et al. (2012). Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. Science *337*, 1190–1195. https://doi.org/10.1126/science.1222794.
4. Schork, A.J., Thompson, W.K., Pham, P., Torkamani, A., Roddey, J.C., Sullivan, P.F., Kelsoe, J.R., O'Donovan, M.C., Furberg, H., Tobacco and Genetics Consortium, et al.. (2013). All SNPs Are Not Created Equal: Genome-Wide Association Studies Reveal a Consistent Pattern of Enrichment among Functionally Annotated SNPs. PLoS Genet. *9*, e1003449. https://doi.org/10.1371/journal.pgen.1003449.
5. GTEx Consortium (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science *369*, 1318–1330. https://doi.org/10.1126/science.aaz1776.
6. Gamazon, E.R., Segrè, A.V., van de Bunt, M., Wen, X., Xi, H.S., Hormozdiari, F., Ongen, H., Konkashbaev, A., Derks, E.M., Aguet, F., et al. (2018). Using an atlas of gene regulation across 44 human tissues to inform complex disease- and trait-associated variation. Nat. Genet. *50*, 956–967. https://doi.org/10.1038/s41588-018-0154-4.
7. Umans, B.D., Battle, A., and Gilad, Y. (2021). Where Are the Disease-Associated eQTLs? Trends Genet. *37*, 109–124. https://doi.org/10.1016/j.tig.2020.08.009.
8. Do, C., Shearer, A., Suzuki, M., Terry, M.B., Gelernter, J., Greally, J.M., and Tycko, B. (2017). Genetic–epigenetic interactions in cis: a major focus in the post-GWAS era. Genome Biol. *18*, 120. https://doi.org/10.1186/s13059-017-1250-y.
9. Min, J.L., Hemani, G., Hannon, E., Dekkers, K.F., Castillo-Fernandez, J., Luijk, R., Carnero-Montoro, E., Lawson, D.J., Burrows, K., Suderman, M., et al. (2021). Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. Nat. Genet. *53*, 1311–1321. https://doi.org/10.1038/s41588-021-00923-x.
10. Pierce, B.L., Tong, L., Argos, M., Demanelis, K., Jasmine, F., Rakibuz-Zaman, M., Sarwar, G., Islam, M.T., Shahriar, H., Islam, T., et al. (2018). Co-occurring expression and methylation QTLs allow detection of common causal variants and shared biological mechanisms. Nat. Commun. *9*, 804. https://doi.org/10.1038/s41467-018-03209-9.
11. Oliva, M., Demanelis, K., Lu, Y., Chernoff, M., Jasmine, F., Ahsan, H., Kibriya, M.G., Chen, L.S., and Pierce, B.L. (2023). DNA methylation QTL mapping across diverse human tissues provides molecular links between genetic variation and complex traits. Nat. Genet. *55*, 112–122. https://doi.org/10.1038/s41588-022-01248-z.
12. Kikas, T., Rull, K., Beaumont, R.N., Freathy, R.M., and Laan, M. (2019). The Effect of Genetic Variation on the Placental Transcriptome in Humans. Front. Genet. *10*, 550. https://doi.org/10.3389/fgene.2019.00550.
13. Delahaye, F., Do, C., Kong, Y., Ashkar, R., Salas, M., Tycko, B., Wapner, R., and Hughes, F. (2018). Genetic variants influence on the placenta regulatory landscape. PLoS Genet. *14*, e1007785. https://doi.org/10.1371/journal.pgen.1007785.
14. Tekola-Ayele, F., Zeng, X., Chatterjee, S., Ouidir, M., Lesseur, C., Hao, K., Chen, J., Tesfaye, M., Marsit, C.J., Workalemahu, T., and Wapner, R. (2022). Placental multi-omics integration identifies candidate functional genes for birthweight. Nat. Commun. *13*, 2384. https://doi.org/10.1038/s41467-022-30007-1.
15. Bhattacharya, A., Li, Y., and Love, M.I. (2021). MOSTWAS: Multi-Omic Strategies for Transcriptome-Wide Association Studies. PLoS Genet. *17*, e1009398. https://doi.org/10.1371/journal.pgen.1009398.
16. Peng, S., Deyssenroth, M.A., Di Narzo, A.F., Lambertini, L., Marsit, C.J., Chen, J., and Hao, K. (2017). Expression quantitative trait loci (eQTLs) in human placentas suggest developmental origins of complex diseases. Hum. Mol. Genet. *26*, 3432–3441. https://doi.org/10.1093/hmg/ddx265.
17. Peng, S., Deyssenroth, M.A., Di Narzo, A.F., Cheng, H., Zhang, Z., Lambertini, L., Ruusalepp, A., Kovacic, J.C., Bjorkegren, J.L.M., Marsit, C.J., et al. (2018). Genetic regulation of the placental transcriptome underlies birth weight and risk of childhood obesity. PLoS Genet. *14*, e1007799. https://doi.org/10.1371/journal.pgen.1007799.
18. Appleton, A.A., Murphy, M.A., Koestler, D.C., Lesseur, C., Paquette, A.G., Padbury, J.F., Lester, B.M., and Marsit, C.J. (2016). Prenatal Programming of Infant Neurobehaviour in a Healthy Population. Paediatr. Perinat. Epidemiol. *30*, 367–375. https://doi.org/10.1111/ppe.12294.
19. Oliva, M., Muñoz-Aguirre, M., Kim-Hellmuth, S., Wucher, V., Gewirtz, A.D.H., Cotter, D.J., Parsana, P., Kasela, S., Balliu, B., Viñuela, A., et al. (2020). The impact of sex on gene expression across human tissues. Science *369*, eaba3066. https://doi.org/10.1126/science.aba3066.
20. Martin, E., Smeester, L., Bommarito, P.A., Grace, M.R., Boggess, K., Kuban, K., Karagas, M.R., Marsit, C.J., O'Shea, T.M., and Fry, R.C. (2017). Sexual epigenetic dimorphism in the human placenta: implications for susceptibility during the prenatal period. Epigenomics *9*, 267–278. https://doi.org/10.2217/epi-2016-0132.
21. Inkster, A.M., Yuan, V., Konwar, C., Matthews, A.M., Brown, C.J., and Robinson, W.P. (2021). A cross-cohort analysis of autosomal DNA methylation sex differences in the term placenta. Biol. Sex Differ. *12*, 38. https://doi.org/10.1186/s13293-021-00381-4.
22. Bulik-Sullivan, B.K., Loh, P.-R., Finucane, H.K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics Consortium, Patterson, N., Daly, M.J., Neale, B.M., and Neale, B.M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat. Genet. *47*, 291–295. https://doi.org/10.1038/ng.3211.
23. Finucane, H.K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P.-R., Anttila, V., Xu, H., Zang, C., Farh, K., et al. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat. Genet. *47*, 1228–1235. https://doi.org/10.1038/ng.3404.
24. Hormozdiari, F., Gazal, S., van de Geijn, B., Finucane, H.K., Ju, C.J.-T., Loh, P.-R., Schoech, A., Reshef, Y., Liu, X., O'Connor, L., et al. (2018). Leveraging molecular quantitative trait loci to understand the genetic architecture of diseases and complex traits. Nat. Genet. *50*, 1041–1047. https://doi.org/10.1038/s41588-018-0148-2.
25. Gay, N.R., Gloudemans, M., Antonio, M.L., Abell, N.S., Balliu, B., Park, Y., Martin, A.R., Musharoff, S., Rao, A.S., Aguet, F., et al. (2020). Impact of admixture and ancestry on eQTL analysis and GWAS colocalization in GTEx. Genome Biol. *21*, 233. https://doi.org/10.1186/s13059-020-02113-0.
26. Kachuri, L., Mak, A.C.Y., Hu, D., Eng, C., Huntsman, S., Elhawary, J.R., Gupta, N., Gabriel, S., Xiao, S., Keys, K.L., et al. (2023). Gene expression in African Americans, Puerto Ricans and Mexican Americans reveals ancestry-specific patterns of genetic architecture. Nat. Genet. *55*, 952–963. https://doi.org/10.1038/s41588-023-01377-z.
27. Storey, J.D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. USA *100*,

9440–9445. https://doi.org/10.1073/pnas.1530509100.
28. Storey, J.D., Bass, A.J., Dabney, A., and Robinson, D. (2017). Qvalue: Q-Value Estimation for False Discovery Rate Control.
29. 1000 Genomes Project Consortium, Abecasis, G.R., Auton, A., Durbin, R.M., Abecasis, G.R., Kang, H.M., Garrison, E.P., Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global reference for human genetic variation. Nature *526*, 68–74. https://doi.org/10.1038/nature15393.
30. Jin, Y., Schaffer, A.A., Feolo, M., Holmes, J.B., and Kattman, B.L. (2019). GRAF-pop: A Fast Distance-Based Method To Infer Subject Ancestry from Multiple Genotype Datasets Without Principal Components Analysis. G3 (Bethesda) *9*, 2447–2461. https://doi.org/10.1534/g3.118.200925.
31. Qi, T., Wu, Y., Zeng, J., Zhang, F., Xue, A., Jiang, L., Zhu, Z., Kemper, K., Yengo, L., Zheng, Z., et al. (2018). Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. Nat. Commun. *9*, 2282. https://doi.org/10.1038/s41467-018-04558-1.
32. Gaunt, T.R., Shihab, H.A., Hemani, G., Min, J.L., Woodward, G., Lyttleton, O., Zheng, J., Duggirala, A., McArdle, W.L., Ho, K., et al. (2016). Systematic identification of genetic influences on methylation across the human life course. Genome Biol. *17*, 61. https://doi.org/10.1186/s13059-016-0926-z.
33. McRae, A.F., Marioni, R.E., Shah, S., Yang, J., Powell, J.E., Harris, S.E., Gibson, J., Henders, A.K., Bowdler, L., Painter, J.N., et al. (2018). Identification of 55,000 Replicated DNA Methylation QTL. Sci. Rep. *8*, 17605. https://doi.org/10.1038/s41598-018-35871-w.
34. Hannon, E., Spiers, H., Viana, J., Pidsley, R., Burrage, J., Murphy, T.M., Troakes, C., Turecki, G., O'Donovan, M.C., Schalkwyk, L.C., et al. (2016). Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. Nat. Neurosci. *19*, 48–54. https://doi.org/10.1038/nn.4182.
35. Cilleros-Portet, A., Lesseur, C., Marí, S., Cosin-Tomas, M., Lozano, M., Irizar, A., Burt, A., García-Santisteban, I., Martín, D.G., Escaramís, G., et al. (2023). Potentially causal associations between placental DNA methylation and schizophrenia and other neuropsychiatric disorders. Preprint at medRxiv. https://doi.org/10.1101/2023.03.07.23286905.
36. McClay, J.L., Shabalin, A.A., Dozmorov, M.G., Adkins, D.E., Kumar, G., Nerella, S., Clark, S.L., Bergen, S.E.; Swedish Schizophrenia Consortium, and Hultman, C.M., et al. (2015). High density methylation QTL analysis in human blood via next-generation sequencing of the methylated genomic DNA fraction. Genome Biol. *16*, 291. https://doi.org/10.1186/s13059-015-0842-7.
37. Sandman, C.A., Glynn, L.M., and Davis, E.P. (2013). Is there a viability–vulnerability tradeoff? Sex differences in fetal programming. J. Psychosom. Res. *75*, 327–335. https://doi.org/10.1016/j.jpsychores.2013.07.009.
38. Iotchkova, V., Ritchie, G.R.S., Geihs, M., Morganella, S., Min, J.L., Walter, K., Timpson, N.J., UK10K Consortium, Dunham, I., Birney, E., and Soranzo, N. (2019). GARFIELD classifies disease-relevant genomic features through integration of functional annotations with association signals. Nat. Genet. *51*, 343–353. https://doi.org/10.1038/s41588-018-0322-6.
39. Davis, C.A., Hitz, B.C., Sloan, C.A., Chan, E.T., Davidson, J.M., Gabdank, I., Hilton, J.A., Jain, K., Baymuradov, U.K., Narayanan, A.K., et al. (2018). The Encyclopedia of DNA elements (ENCODE): data portal update. Nucleic Acids Res. *46*, D794–D801. https://doi.org/10.1093/nar/gkx1081.
40. Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317–330. https://doi.org/10.1038/nature14248.
41. ENCODE Project Consortium (2012). An Integrated Encyclopedia of DNA Elements in the Human Genome. Nature *489*, 57–74. https://doi.org/10.1038/nature11247.
42. Bernstein, B.E., Stamatoyannopoulos, J.A., Costello, J.F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M.A., Beaudet, A.L., Ecker, J.R., et al. (2010). The NIH Roadmap Epigenomics Mapping Consortium. Nat. Biotechnol. *28*, 1045–1048. https://doi.org/10.1038/nbt1010-1045.
43. Ernst, J., and Kellis, M. (2017). Chromatin-state discovery and genome annotation with ChromHMM. Nat. Protoc. *12*, 2478–2492. https://doi.org/10.1038/nprot.2017.124.
44. Karmodiya, K., Krebs, A.R., Oulad-Abdelghani, M., Kimura, H., and Tora, L. (2012). H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. BMC Genom. *13*, 424. https://doi.org/10.1186/1471-2164-13-424.
45. Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Peng, W., Zhang, M.Q., and Zhao, K. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nat. Genet. *40*, 897–903. https://doi.org/10.1038/ng.154.
46. Elliott, G., Hong, C., Xing, X., Zhou, X., Li, D., Coarfa, C., Bell, R.J.A., Maire, C.L., Ligon, K.L., Sigaroudinia, M., et al. (2015). Intermediate DNA methylation is a conserved signature of genome regulation. Nat. Commun. *6*, 6363. https://doi.org/10.1038/ncomms7363.
47. Del Gobbo, G.F., Konwar, C., and Robinson, W.P. (2020). The significance of the placental genome and methylome in fetal and maternal health. Hum. Genet. *139*, 1183–1196. https://doi.org/10.1007/s00439-019-02058-w.
48. Konwar, C., Del Gobbo, G., Yuan, V., and Robinson, W.P. (2019). Considerations when processing and interpreting genomics data of the placenta. Placenta *84*, 57–62. https://doi.org/10.1016/j.placenta.2019.01.006.
49. Bommarito, P.A., and Fry, R.C. (2019). Chapter 2-1 - The Role of DNA Methylation in Gene Regulation. In Toxicoepigenetics, S.D. McCullough and D.C. Dolinoy, eds. (Academic Press), pp. 127–151. https://doi.org/10.1016/B978-0-12-812433-8.00005-8.
50. Cheng, X., Hashimoto, H., Horton, J.R., and Zhang, X. (2011). Chapter 2 - Mechanisms of DNA Methylation, Methyl-CpG Recognition, and Demethylation in Mammals. In Handbook of Epigenetics, T. Tollefsbol, ed. (Academic Press), pp. 9–24. https://doi.org/10.1016/B978-0-12-375709-8.00002-2.
51. Edgar, R., Tan, P.P.C., Portales-Casamar, E., and Pavlidis, P. (2014). Meta-analysis of human methylomes reveals stably methylated sequences surrounding CpG islands associated with high gene expression. Epigenet. Chromatin *7*, 28. https://doi.org/10.1186/1756-8935-7-28.
52. Plant, T.M., and Zeleznik, A.J. (2014). Knobil and Neill's Physiology of Reproduction (Academic Press).
53. Burton, G.J., and Fowden, A.L. (2015). The placenta: a multifaceted, transient organ. Philos. Trans. R. Soc. Lond. B Biol. Sci. *370*, 20140066. https://doi.org/10.1098/rstb.2014.0066.
54. Solmi, M., Radua, J., Olivola, M., Croce, E., Soardo, L., Salazar de Pablo, G., Il Shin, J., Kirkbride, J.B., Jones, P., Kim, J.H., et al. (2022). Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. Mol. Psychiatry *27*, 281–295. https://doi.org/10.1038/s41380-021-01161-7.
55. Arain, M., Haque, M., Johal, L., Mathur, P., Nel, W., Rais, A., Sandhu, R., and Sharma, S. (2013). Maturation of the adolescent brain. Neuropsychiatr. Dis. Treat. *9*, 449–461. https://doi.org/10.2147/NDT.S39776.
56. Steinthorsdottir, V., McGinnis, R., Williams, N.O., Stefansdottir, L., Thorleifsson, G., Shooter, S., Fadista, J., Sigurdsson, J.K., Auro, K.M., Berezina, G., et al. (2020). Genetic predisposition to hypertension is associated with preeclampsia in European and Central Asian women. Nat. Commun. *11*, 5976. https://doi.org/10.1038/s41467-020-19733-6.
57. Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace, C., and Plagnol, V. (2014). Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. PLoS Genet. *10*, e1004383. https://doi.org/10.1371/journal.pgen.1004383.
58. Pardiñas, A.F., Holmans, P., Pocklington, A.J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S.E., Bishop, S., Cameron, D., Hamshere, M.L., et al. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat. Genet. *50*, 381–389. https://doi.org/10.1038/s41588-018-0059-2.
59. Chowdhury, N.U., Guntur, V.P., Newcomb, D.C., and Wechsler, M.E. (2021). Sex and gender in asthma. Eur. Respir. Rev. *30*, 210067. https://doi.org/10.1183/16000617.0067-2021.
60. Ferreira, M.A.R., Mathur, R., Vonk, J.M., Szwajda, A., Brumpton, B., Granell, R., Brew, B.K., Ullemar, V., Lu, Y., Jiang, Y., et al. (2019). Genetic Architectures of Childhood- and Adult-Onset Asthma Are Partly Distinct. Am. J. Hum. Genet. *104*, 665–684. https://doi.org/10.1016/j.ajhg.2019.02.022.
61. Pividori, M., Schoettler, N., Nicolae, D.L., Ober, C., and Im, H.K. (2019). Shared and Distinct Genetic Risk Factors for Childhood Onset and Adult Onset Asthma: Genome-and Transcriptome-wide Studies. Lancet Respir. Med. *7*, 509–522. https://doi.org/10.1016/S2213-2600(19)30055-4.
62. Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a

comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. *44*, W90–W97. https://doi.org/10.1093/nar/gkw377.

63. Xie, Z., Bailey, A., Kuleshov, M.V., Clarke, D.J.B., Evangelista, J.E., Jenkins, S.L., Lachmann, A., Wojciechowicz, M.L., Kropiwnicki, E., Jagodnik, K.M., et al. (2021). Gene Set Knowledge Discovery with Enrichr. Curr. Protoc. *1*, e90. https://doi.org/10.1002/cpz1.90.
64. Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., and Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinf. *14*, 128. https://doi.org/10.1186/1471-2105-14-128.
65. Mitchell, C., Provost, K., Niu, N., Homer, R., and Cohn, L. (2011). Interferon-gamma acts on the airway epithelium to inhibit local and systemic pathology in allergic airway disease. J. Immunol. *187*, 3815–3820. https://doi.org/10.4049/jimmunol.1100436.
66. Bonder, M.J., Luijk, R., Zhernakova, D.V., Moed, M., Deelen, P., Vermaat, M., van Iterson, M., van Dijk, F., van Galen, M., Bot, J., et al. (2017). Disease variants alter transcription factor levels and methylation of their binding sites. Nat. Genet. *49*, 131–138. https://doi.org/10.1038/ng.3721.
67. Barker, D.J.P., and Thornburg, K.L. (2013). The Obstetric Origins of Health for a Lifetime. Clin. Obstet. Gynecol. *56*, 511–519. https://doi.org/10.1097/GRF.0b013e31829cb9ca.
68. Bhattacharya, A., Freedman, A.N., Avula, V., Harris, R., Liu, W., Pan, C., Lusis, A.J., Joseph, R.M., Smeester, L., Hartwell, H.J., et al. (2022). Placental genomics mediates genetic associations with complex health traits and disease. Nat. Commun. *13*, 706. https://doi.org/10.1038/s41467-022-28365-x.
69. Randall, J.C., Winkler, T.W., Kutalik, Z., Berndt, S.I., Jackson, A.U., Monda, K.L., Kilpeläinen, T.O., Esko, T., Mägi, R., Li, S., et al. (2013). Sex-stratified Genome-wide Association Studies Including 270,000 Individuals Show Sexual Dimorphism in Genetic Loci for Anthropometric Traits. PLoS Genet. *9*, e1003500. https://doi.org/10.1371/journal.pgen.1003500.
70. Bernabeu, E., Canela-Xandri, O., Rawlik, K., Talenti, A., Prendergast, J., and Tenesa, A. (2021). Sex differences in genetic architecture in the UK Biobank. Nat. Genet. *53*, 1283–1289. https://doi.org/10.1038/s41588-021-00912-0.
71. Kautzky-Willer, A., Harreiter, J., and Pacini, G. (2016). Sex and Gender Differences in Risk, Pathophysiology and Complications of Type 2 Diabetes Mellitus. Endocr. Rev. *37*, 278–316. https://doi.org/10.1210/er.2015-1137.
72. Alur, P. (2019). Sex Differences in Nutrition, Growth, and Metabolism in Preterm Infants. Front. Pediatr. *7*, 22. https://doi.org/10.3389/fped.2019.00022.
73. Oliva, M., Demanelis, K., Jasmine, F., Lu, Y., Hahsan, H., Kibriya, M.G., Chen, L.S., and Pierce, B.L. (2021). Genetic regulation of DNA methylation across tissues reveals thousands of molecular links to complex traits. https://doi.org/10.21203/rs.3.rs-492596/v1.
74. Yazar, S., Alquicira-Hernandez, J., Wing, K., Senabouth, A., Gordon, M.G., Andersen, S., Lu, Q., Rowson, A., Taylor, T.R.P., Clarke, L., et al. (2022). Single-cell eQTL mapping identifies cell type–specific genetic control of autoimmune disease. Science *376*, eabf3041. https://doi.org/10.1126/science.abf3041.
75. Zhang, P., Lewinger, J.P., Conti, D., Morrison, J.L., and Gauderman, W.J. (2016). Detecting Gene-Environment Interactions for a Quantitative Trait in a Genome-Wide Association Study. Genet. Epidemiol. *40*, 394–403. https://doi.org/10.1002/gepi.21977.
76. Martin, J., Khramtsova, E.A., Goleva, S.B., Blokland, G.A.M., Traglia, M., Walters, R.K., Hübel, C., Coleman, J.R.I., Breen, G., Børglum, A.D., et al. (2021). Examining Sex-Differentiated Genetic Effects Across Neuropsychiatric and Behavioral Traits. Biol. Psychiatry *89*, 1127–1137. https://doi.org/10.1016/j.biopsych.2020.12.024.
77. Zhong, Y., Perera, M.A., and Gamazon, E.R. (2019). On Using Local Ancestry to Characterize the Genetic Architecture of Human Traits: Genetic Regulation of Gene Expression in Multiethnic or Admixed Populations. Am. J. Hum. Genet. *104*, 1097–1115. https://doi.org/10.1016/j.ajhg.2019.04.009.
78. Smith, A.K., Kilaru, V., Kocak, M., Almli, L.M., Mercer, K.B., Ressler, K.J., Tylavsky, F.A., and Conneely, K.N. (2014). Methylation quantitative trait loci (meQTLs) are consistently detected across ancestry, developmental stage, and tissue type. BMC Genom. *15*, 145. https://doi.org/10.1186/1471-2164-15-145.
79. Chen, D.-P., Lin, Y.-C., and Fann, C.S.J. (2016). Methods for identifying differentially methylated regions for sequence- and array-based data. Brief. Funct. Genomics *15*, 485–490. https://doi.org/10.1093/bfgp/elw018.
80. Rakyan, V.K., Down, T.A., Thorne, N.P., Flicek, P., Kulesha, E., Gräf, S., Tomazou, E.M., Bäckdahl, L., Johnson, N., Herberth, M., et al. (2008). An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). Genome Res. *18*, 1518–1529. https://doi.org/10.1101/gr.077479.108.
81. Paquette, A.G., Houseman, E.A., Green, B.B., Lesseur, C., Armstrong, D.A., Lester, B., and Marsit, C.J. (2016). Regions of variable DNA methylation in human placenta associated with newborn neurobehavior. Epigenetics *11*, 603–613. https://doi.org/10.1080/15592294.2016.1195534.
82. Gatev, E., Gladish, N., Mostafavi, S., and Kobor, M.S. (2020). CoMeBack: DNA methylation array data analysis for co-methylated regions. Bioinformatics *36*, 2675–2683. https://doi.org/10.1093/bioinformatics/btaa049.
83. Gatev, E., Inkster, A.M., Negri, G.L., Konwar, C., Lussier, A.A., Skakkebaek, A., Sokolowski, M.B., Gravholt, C.H., Dunn, E.C., Kobor, M.S., and Aristizabal, M.J. (2021). Autosomal sex-associated co-methylated regions predict biological sex from DNA methylation. Nucleic Acids Res. *49*, 9097–9116. https://doi.org/10.1093/nar/gkab682.
84. Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.-R.; ReproGen Consortium; Psychiatric Genomics Consortium; Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium 3, and Duncan, L., et al. (2015). An atlas of genetic correlations across human diseases and traits. Nat. Genet. *47*, 1236–1241. https://doi.org/10.1038/ng.3406.
85. Gazal, S., Finucane, H.K., Furlotte, N.A., Loh, P.-R., Palamara, P.F., Liu, X., Schoech, A., Bulik-Sullivan, B., Neale, B.M., Gusev, A., and Price, A.L. (2017). Linkage disequilibrium–dependent architecture of human complex traits shows action of negative selection. Nat. Genet. *49*, 1421–1427. https://doi.org/10.1038/ng.3954.
86. Iotchkova, V., Ritchie, G.R.S., Geihs, M., Morganella, S., Min, J.L., Walter, K., Timpson, N., UK10K Consortium, Dunham, I., Birney, E., et al. (2016). GARFIELD - GWAS Analysis of Regulatory or Functional Information Enrichment with LD correction. Genomics. https://doi.org/10.1101/085738.
87. Hormozdiari, F., van de Bunt, M., Segrè, A.V., Li, X., Joo, J.W.J., Bilow, M., Sul, J.H., Sankararaman, S., Pasaniuc, B., and Eskin, E. (2016). Colocalization of GWAS and eQTL Signals Detects Target Genes. Am. J. Hum. Genet. *99*, 1245–1260. https://doi.org/10.1016/j.ajhg.2016.10.003.
88. Aryee, M.J., Jaffe, A.E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A.P., Hansen, K.D., and Irizarry, R.A. (2014). Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA Methylation microarrays. Bioinformatics *30*, 1363–1369. https://doi.org/10.1093/bioinformatics/btu049.
89. Fenton, T.R., and Kim, J.H. (2013). A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. BMC Pediatr. *13*, 59. https://doi.org/10.1186/1471-2431-13-59.
90. Lam, M., Awasthi, S., Watson, H.J., Goldstein, J., Panagiotaropoulou, G., Trubetskoy, V., Karlsson, R., Frei, O., Fan, C.-C., De Witte, W., et al. (2020). RICOPILI: Rapid Imputation for COnsortias PIpeLIne. Bioinformatics *36*, 930–933. https://doi.org/10.1093/bioinformatics/btz633.
91. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience *4*, 7. https://doi.org/10.1186/s13742-015-0047-8.
92. Anderson, C.A., Pettersson, F.H., Clarke, G.M., Cardon, L.R., Morris, A.P., and Zondervan, K.T. (2010). Data quality control in genetic case-control association studies. Nat. Protoc. *5*, 1564–1573. https://doi.org/10.1038/nprot.2010.116.
93. Galinsky, K.J., Bhatia, G., Loh, P.-R., Georgiev, S., Mukherjee, S., Patterson, N.J., and Price, A.L. (2016). Fast Principal-Component Analysis Reveals Convergent Evolution of ADH1B in Europe and East Asia. Am. J. Hum. Genet. *98*, 456–472. https://doi.org/10.1016/j.ajhg.2015.12.022.
94. Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I., Chew, E.Y., Levy, S., McGue, M., et al. (2016). Next-generation genotype imputation service and methods. Nat. Genet. *48*, 1284–1287. https://doi.org/10.1038/ng.3656.
95. Verlouw, J.A.M., Clemens, E., de Vries, J.H., Zolk, O., Verkerk, A.J.M.H., am Zehnhoff-Dinnesen, A., Medina-Gomez, C., Lanvers-Kaminsky, C., Rivadeneira, F., Langer, T., et al. (2021). A comparison of genotyping arrays. Eur. J. Hum. Genet. *29*, 1611–1624. https://doi.org/10.1038/s41431-021-00917-7.

96. Fortin, J.-P., Labbe, A., Lemire, M., Zanke, B.W., Hudson, T.J., Fertig, E.J., Greenwood, C.M., and Hansen, K.D. (2014). Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biol. *15*, 503. https://doi.org/10.1186/s13059-014-0503-2.
97. Triche, T.J., Weisenberger, D.J., Van Den Berg, D., Laird, P.W., and Siegmund, K.D. (2013). Low-level processing of Illumina Infinium DNA Methylation BeadArrays. Nucleic Acids Res. *41*, e90. https://doi.org/10.1093/nar/gkt090.
98. Fortin, J.-P., Triche, T.J., and Hansen, K.D. (2017). Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. Bioinformatics *33*, 558–560. https://doi.org/10.1093/bioinformatics/btw691.
99. Chen, Y.a., Lemire, M., Choufani, S., Butcher, D.T., Grafodatskaya, D., Zanke, B.W., Gallinger, S., Hudson, T.J., and Weksberg, R. (2013). Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics *8*, 203–209. https://doi.org/10.4161/epi.23470.
100. Price, E.M., and Robinson, W.P. (2018). Adjusting for Batch Effects in DNA Methylation Microarray Data, a Lesson Learned. Front. Genet. *9*, 83. https://doi.org/10.3389/fgene.2018.00083.
101. Touleimat, N., and Tost, J. (2012). Complete pipeline for Infinium® Human Methylation 450K BeadChip data processing using subset quantile normalization for accurate DNA methylation estimation. Epigenomics *4*, 325–341. https://doi.org/10.2217/epi.12.21.
102. Shabalin, A.A. (2012). Matrix eQTL: ultra fast eQTL analysis via large matrix operations. Bioinformatics *28*, 1353–1358. https://doi.org/10.1093/bioinformatics/bts163.
103. Sul, J.H., Martin, L.S., and Eskin, E. (2018). Population structure in genetic studies: Confounding factors and mixed models. PLoS Genet. *14*, e1007309. https://doi.org/10.1371/journal.pgen.1007309.
104. Bookman, E.B., McAllister, K., Gillanders, E., Wanke, K., Balshaw, D., Rutter, J., Reedy, J., Shaughnessy, D., Agurs-Collins, T., Paltoo, D., et al. (2011). Gene-Environment Interplay in Common Complex Diseases: Forging an Integrative Model—Recommendations From an NIH Workshop. Genet. Epidemiol. *35*, 217–225. https://doi.org/10.1002/gepi.20571.
105. Assary, E., Vincent, J.P., Keers, R., and Pluess, M. (2018). Gene-environment interaction and psychiatric disorders: Review and future directions. Semin. Cell Dev. Biol. *77*, 133–143. https://doi.org/10.1016/j.semcdb.2017.10.016.
106. Do, C., Lang, C.F., Lin, J., Darbary, H., Krupska, I., Gaba, A., Petukhova, L., Vonsattel, J.-P., Gallagher, M.P., Goland, R.S., et al. (2016). Mechanisms and Disease Associations of Haplotype-Dependent Allele-Specific DNA Methylation. Am. J. Hum. Genet. *98*, 934–955. https://doi.org/10.1016/j.ajhg.2016.03.027.
107. Plongthongkum, N., van Eijk, K.R., de Jong, S., Wang, T., Sul, J.H., Boks, M.P.M., Kahn, R.S., Fung, H.-L., Ophoff, R.A., and Zhang, K. (2014). Characterization of Genome-Methylome Interactions in 22 Nuclear Pedigrees. PLoS One *9*, e99313. https://doi.org/10.1371/journal.pone.0099313.
108. Ng, B., White, C.C., Klein, H.-U., Sieberts, S.K., McCabe, C., Patrick, E., Xu, J., Yu, L., Gaiteri, C., Bennett, D.A., et al. (2017). An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. Nat. Neurosci. *20*, 1418–1426. https://doi.org/10.1038/nn.4632.
109. Banovich, N.E., Lan, X., McVicker, G., van de Geijn, B., Degner, J.F., Blischak, J.D., Roux, J., Pritchard, J.K., and Gilad, Y. (2014). Methylation QTLs Are Associated with Coordinated Changes in Transcription Factor Binding, Histone Modifications, and Gene Expression Levels. PLoS Genet. *10*, e1004663. https://doi.org/10.1371/journal.pgen.1004663.
110. Mostafavi, S., Battle, A., Zhu, X., Urban, A.E., Levinson, D., Montgomery, S.B., and Koller, D. (2013). Normalizing RNA-Sequencing Data by Modeling Hidden Covariates with Prior Knowledge. PLoS One *8*, e68141. https://doi.org/10.1371/journal.pone.0068141.
111. Zhou, H.J., Li, L., Li, Y., Li, W., and Li, J.J. (2022). PCA outperforms popular hidden variable inference methods for molecular QTL mapping. Genome Biol. *23*, 210. https://doi.org/10.1186/s13059-022-02761-4.
112. Yuan, V., Hui, D., Yin, Y., Peñaherrera, M., Beristain, A., and Robinson, W. (2020). Cell-specific Characterization of the Placental Methylome. In Review. https://doi.org/10.21203/rs.3.rs-38223/v3.
113. Hannon, E., Spiers, H., Viana, J., Pidsley, R., Burrage, J., Murphy, T.M., Troakes, C., Turecki, G., O'Donovan, M.C., Schalkwyk, L.C., et al. (2016). Methylation quantitative trait loci in the developing brain and their enrichment in schizophrenia-associated genomic regions. Nat. Neurosci. *19*, 48–54. https://doi.org/10.1038/nn.4182.
114. Lumley, T. (2018). rmeta: Meta-Analysis.
115. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The Human Genome Browser at UCSC. Genome Res. *12*, 996–1006. https://doi.org/10.1101/gr.229102.
116. Wallace, C. (2021). A more accurate method for colocalisation analysis allowing for multiple causal variants. PLoS Genet. *17*, e1009440. https://doi.org/10.1371/journal.pgen.1009440.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
mQTL associations and annotations	https://doi.org/10.17605/OSF.IO/9R4WF	https://osf.io/9r4wf/
Software and algorithms		
GitHub (code supporting this work)	https://doi.org/10.5281/zenodo.7262352	https://github.com/wilcas/sex_specific_mQTL
LDSC	Bulik-Sulivan et al.,[22,84] Finucane et al.,[23] Gazal, S et al.[85]	https://github.com/bulik/ldsc
GARFIELD	Iotchkova et al.[86]	https://www.ebi.ac.uk/birney-srv/GARFIELD/
CAVIAR	Hormozdiari et al.[87]	http://genetics.cs.ucla.edu/caviar/
Enrichr	Kuleshov et al.[62]	https://maayanlab.cloud/Enrichr/
coloc	Giambartolomei et al.[57]	https://chr1swallace.github.io/coloc/
minfi	Aryee et al.[88]	https://bioconductor.org/packages/release/bioc/html/minfi.html
Other		
Rhode Island Child Health Study (RICHS)	phs001586.v1.p1; GSE75248	https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001586.v1.p1 ; https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE75248
Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD)	phs001717.v1.p1	https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001717.v1.p1

RESOURCE AVAILABILITY

Lead contact

Further information and request for resources should be directed toand will be fulfilled by the lead contact, Jessica K. Dennis (jessica.dennis@bcchr.ca).

Materials availability

The study did not generate new unique reagents.

Data and code availability

- Placental mQTL summary data, along with annotations generated as part of statistical analyses, have been deposited to the Open Science Framework. DOIs are listed in the key resources table. This paper analyzes existing, publicly available data, accession numbers for the datasets are listed in the key resources table.
- All original code has been deposited at Zenodo and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

We obtained placental genotype and DNAm data from two different studies, both of which sampled fetal side placental tissue from pregnancies lasting >=37 weeks. Discovery sample: The National Institute of Child Health and Human Development (NICHD) study[13] included 303 subjects (151 male, 152 female) of diverse reported maternal race/ethnicity (Table S1), the vast majority of which (96%) had no reported pre-eclampsia. Maternal subjects were selected to ensure a low-risk population, which at a minimum excluded subjects with cigarette consumption within six months of delivery, consumption of at least one alcoholic drink per day, chronic hypertension, diabetes, HIV or AIDS, or a history of gestational diabetes in past pregnancies. Additionally, subjects with suspected fetal congenital structural or chromosomal anomalies were excluded, and each pregnancy was screened via ultrasound at the first trimester to ensure accurate dating of gestational age. Birth length was not publicly available for this study, and our discovery and replication samples could therefore not be compared to one another on this measure. Placental samples were taken from a single 0.5 cm x 0.5 cm x 0.5 cm parenchymal biopsy and were frozen within one hour of

delivery. All samples were genotyped using the Illumina HumanOmni2.5 Beadchip, and DNAm was measured using the Illumina HumanMethylation450k array (dbGaP accession phs001717.v1.p1).

Replication sample: The Rhode Island Child Health Study (RICHS)[16,18] included 149 subjects (74 male and 75 female) of primarily non-Hispanic White reported ethnicity. Eligible participants were those with no serious prenatal complications or congenital or chromosomal abnormalities, and large/small for gestational age samples (either > 90% or < 10% on the Fenton growth chart[89]) were matched to average for gestational age samples based on sex, gestational age, and maternal age (Table S1).[81] A total of 12 fetal side placental biopsies were taken from each sample, three from each quadrant of the placenta, and frozen within two hours. Samples were homogenized before measurement to ensure heterogeneous sampling across placental quadrants. Samples were genotyped on the Illumina Expanded Multi-Ethnic Genotyping Array (Mega-EX) (dbGaP accession phs001586.v1.p1). DNAm was measured on the IlluminaHumanMethylation450k array (GEO accession GSE75248). DNAm and SNP genotype data were linked by a common ID file available upon request.

METHOD DETAILS

Genotype data processing and quality control

Genotyping data from NICHD and RICHS were processed separately, and were subject to the same processing and quality control (QC) protocol, which was based on the RICOPILI pipeline.[90] Using plink 1.9,[91] we removed SNPs that were strand ambiguous, had a call rate < 0.05, or MAF < 0.01. We removed individuals with a mismatch between recorded and genotyped sex, SNP missingness > 0.02, and with excess heterozygosity (F_{het} > 0.2). SNPs on the X chromosome were kept, using the default 0/2 encoding in plink for SNPs on the X chromosome in males.

We used the GRAFpop algorithm[30] on autosomal SNPs to assign individuals to one of 7 ancestry groups: EUR (European), AFR (African), AFR_AM (African American), LAT_AM_1 (Latin American 1), LAT_AM_2 (Latin American 2), EAS (East Asian), PAC (Asian Pacific Islander), and SAS (South Asian), which we used to help us decide the number of ancestry PCs to include in downstream analyses. Genetic ancestry assignments agreed well with maternal self-reported race/ethnicity (Table S2), and samples were filtered to ensure identity by descent did not exceed an $|F_{ST}|$ > 0.2 between samples within each ancestry group.

To define ancestry PCs, we first pruned SNPs based on LD in a pairwise manner using plink (running –indep-pairwise 200 100 0.2 twice), and then removed SNPs with an MAF < 0.05, and SNPs in highly recombinant regions (the HLA region, chr6:28,477,797-33,448,354 in GRCh37, and regions of long range LD).[92] Ten ancestry PCs were computed from this set of SNPs using FastPCA.[93] From these 10, we ultimately selected the first five to include in all downstream analyses on the basis of their ability to separate distinct populations of subjects (defined by self-reported race or GRAF-pop estimated ancestry) in each study (Figures S4A–S4D). Individuals of the same GRAF-pop estimated ancestry clustered together along the axes of PC1, PC2, PC3, and PC4. No further clustering by ancestry was observed along the axes of PC5 and above, suggesting that the first five PCs would sufficiently account for population stratification in downstream analyses.

Before imputing SNP genotypes, we removed variants that were associated with genotype batch (Bonferroni corrected p < 0.05) and variants with a Hardy-Weinberg equilibrium p < 1e-10. Variants were then aligned and imputed to the cosmopolitan 1000 genomes phase 3 reference panel[29] using the Michigan imputation server.[94] Variants were filtered to have an imputation quality R2 > 0.3 and an ancestry-specific MAF within 0.1 of their corresponding 1000 genomes phase 3 group. Last, any remaining imputed SNPs associated with genotype batch were removed. The NICHD study had 275 samples with 14,110,641 SNPs remaining for analysis, whereas RICHS had 136 samples with 6,139,984 SNPs remaining for analysis. This between-study difference in the total number of imputed SNPs is likely due to differences in sample size, array, and ancestral diversity.[95] However, stratification in these groups was not apparent in genotyping PCs greater than five (Figures S4B and S4D), and we adjust for these five PCs in each mQTL set.

DNA methylation array processing and quality control

For NICHD and RICHS samples, we used publicly available DNAm data measured by the Illumina HumanMethylation450k array. The data available from NICHD were already partially processed, having already undergone both background correction using internal control probes, and quantile normalization in Genome Studio (v2011.1) software available from Illumina.[13] As such, internal control probes were unavailable in NICHD for use in functional normalization, which can outperform other normalization approaches.[96] Thus, to ensure comparability across samples, we repeated background correction in both NICHD and RICHS using using the single sample normal-exponential via the out-of-band (noob) method, which does not require internal control probes (*preprocessNoob, minfi*[88,97,98]), and re-computed quantile normalization in each sample after applying the following quality control measures as implemented in *minfi*:[88]

First the pattern of methylation across all probes was checked to confirm that it followed a hemi-methylated pattern characteristic of the placenta.[48] Next, we removed individuals with sex mismatches, i.e., individuals with sex chromosome DNAm values that did not cluster with samples in their reported sex (two samples total, neither of which were flagged as contaminated during genotyping QC). We then removed probes with a detection p > 0.01, and probes that either failed or were missing in > 20% of samples. Next, we removed probes containing a SNP (MAF > 0.01 across all 1000 Genomes subjects) in their single base extension site (annotated by Illumina and identified by others),[99] which affects hybridization of these sequences to the array. We then removed Y chromosome probes, and non-specific cross-hybridizing probes.[100] Lastly, we ensured that the data were normalized across remaining samples, using *preprocessQuantile* from the R package *minfi,* which normalizes values across samples for each probe while accounting for biases inherent to 450K array data.[88,101] After implementing these QC steps, the NICHD and RICHS studies had 447,232 and 446,976 probes remaining for analysis.

Criteria for GWAS summary statistics

We define childhood onset traits and conditions as those which have a median age of onset < 25,[54] which is roughly considered to be the end of brain maturation and adolescence.[55] Summary statistics were formatted for all LD score analyses using the munge_sumstats.py script made available by the LDSC developers. This includes the following quality control measures: only biallelic SNPs are kept, strand ambiguous SNPs are excluded, duplicate SNPs are excluded, SNPs have an imputation INFO score > 0.9, a MAF > 0.01, $0 < p \leq 1$, SNPs with a number of samples < 90th quantile divided by 1.5 are excluded, and the median *Z* score of all SNPs are within 0.1 of its expected null value (0 for signed statistics, 1 for an odds ratio). This largely matches what has been done in previous publications dealing with sex-stratified summary statistics, with the exception of differences MAF threshold used across different studies.[70,76]

Genome-wide summary statistics for all traits were downloaded in April 2021 and are publicly available (Table S5). Sex-stratified GWAS from the Psychiatric Genomics Consortium (PGC) were obtained from Martin et al., 2021.[76] Summary statistics for preeclampsia were requested separately via an InterPreGen access request. In S-LDSC analyses, we imposed a filter of a GWAS h^2_{SNP} z-score[22,23] > 1.644854, corresponding to a h^2_{SNP} estimate that was significant at a nominal $p < 0.05$.

QUANTIFICATION AND STATISTICAL ANALYSIS

Mapping ancestry-stratified mQTL in NICHD

In order to investigate the degree to which mQTL computed in NICHD, which is an ancestrally diverse cohort (Tables S1 and S2) were reflective of a given ancestry, we elected to compute mQTL in the three GRAF-pop ancestry groups with the largest number of samples : AFR_AM (N=82), EUR (N=72), and LAT_AM (N=76). We combined samples from the LAT_AM_1 (N=30) and LAT_AM_2 (N=46) groups into the LAT_AM group to reach a similar sample size to EUR and AFR_AM samples. We used predicted ancestries here as only reported maternal ethnicity was present, and discrepancies between predicted ancestry and self-reported maternal ethnicity could be due to paternal genetic background. As in our main mQTL analyses, we used linear regression, as implemented in matrixEQTL,[102] for associating imputed genotype dosages to normalized DNAm beta values.[101] In line with our main QTL analysis, each SNP within 75 kb upstream or downstream of each CpG site was regressed onto DNAm, accounting for gestational age, sex, methylation array batch, 5 genetic ancestry PCs, and 9 DNAm PCs. We chose to include genetic ancestry PCs in this analysis to account for any remaining relatedness between samples.[103]

Mapping cross-sex, sex-dependent, male-stratified, and female-stratified mQTL

We conducted four mQTL analyses to detect four different cis-mQTL effects: a cross-sex analysis, an interaction analysis (in which the effect of SNP on DNAm differed by sex, as captured by a genotype by sex interaction term), an analysis of males only, and an analysis of females only (Figure 1). Each analysis was conducted separately in NICHD and RICHS.

We used linear regression, as implemented in matrixEQTL,[102] for associating imputed genotype dosages to quantile normalized DNAm beta values.[101] Each SNP within 75 kb upstream or downstream of each CpG site was regressed onto DNAm, accounting for gestational age, sex, methylation array batch, 5 genetic ancestry PCs, and 9 DNAm PCs. As interaction effects (i.e., sex-dependent effects) are harder to detect, and require a mQTL effect, we elected to exclude trans-mQTL, or SNPs > 75kb from an associated CpG site, from this analysis.[75,104,105] We chose a 75 kb window as a compromise between the observation that as much as 47% of cis-mQTL can be detected within 2 kb of a CpG site, and 1 Mb, which is the upper range of what is considered to be "cis" in QTL analysis.[11,13,32,106–109] The number of DNAm PCs to include was determined based on the number of mQTL declared on chromosome 21 at a Bonferroni corrected $p < 0.05$, varying the number of PCs included while keeping other covariates fixed and selecting the number of PCs for which the number of mQTL declared did not improve (Figures S4E and S4F).[108,110,111]

Since sex-dependent molQTL detection can be biased by sex differences in cell type proportions,[19] we ensured that the top 10 DNAm PCs were correlated with placental cell type proportions (Figures S3A and S3B). Placental cell type proportions were estimated using the data available on cell-type specific methylation in placenta in the R package *PlaNET*,[112] and the Houseman algorithm implemented in *minfi*.[88] Ultimately, we elected to use DNAm PCs in place of estimated cell type proportion as covariates in mQTL mapping, as the PCs captured variation attributable to both known (i.e., cell type proportions) and unknown sources of confounding. We also checked if there was a difference in estimated cell-type proportion between males and females in each study (Figures S4C and S4D). We observed a sex difference in average estimated stromal cell proportion in the NICHD study, which is accounted for by DNAm PCs.

Additionally, we assessed whether DNAm correlated with any phenotypic or maternal characteristics recorded for NICHD (e.g., gestational age, fetal sex, birth weight, mode of delivery, mode of labor, and maternal ethnicity; Figure S3A) and RICHS (e.g., gestational age, fetal sex, birth weight, fenton growth curve group, maternal BMI, and maternal ethnicity; Figure S3B). DNAm PC1 in NICHD and DNAm PC2 in RICHS showed high correlation with fetal sex, which is unsurprising given our inclusion of the X chromosome. Additionally, DNAm PCs beyond PC5 showed correlation with self-reported maternal ethnicity of Black and White, and comparatively lower correlation with remaining variables. We chose not to account for the effect of these covariates when computing PCs since accounting for additional covariates when computing hidden covariates, such as DNAm PCs, does not consistently perform better than accounting for both PCs and fixed covariates separately.[111]

Before meta-analyzing results from the NICHD and RICHS samples, we tested cross-sex and sex-dependent mQTL identified in NICHD for replication in RICHS using the π_1 statistic,[27] which is a better measure of comparability than the raw proportion of overlapping sites because it accounts for between-study differences in the number of tested SNP-CpG pairs, which can artificially inflate or deflate estimates of overlap.

For the π_1 analysis, we relied on cross-sex and sex-dependent mQTL called in each study at an FDR < 0.05 because calling mQTL at a Bonferroni corrected p < 0.05 resulted in unstable estimates of π_1. Within matrixEQTL,[102] we identified cross-sex mQTL with linear models testing for association between genotype and DNAm. Sex-dependent mQTL were those with a statistically significant genotype by sex interaction term. In the NICHD study, we found 275,806 CpGs with at least one mQTL and 60,537 CpGs with at least one mQTL that interacted with sex. In RICHS, we found 129,361 CpGs with at least one mQTL and 19,530 CpGs with at least one mQTL that differed by sex. The π_1 estimate for the cross-sex mQTL was 0.74, which is the proportion of mQTL in the NICHD study that have similar effects in the RICHS study (computed over 2,691,024 SNP-CpG pairs available in both studies). Sex-dependent mQTL showed modest replication (π_1 = 0.28, computed over 80,363 SNP-CpG pairs available in both studies).

We then meta-analyzed across the NICHD and RICHS samples for each of the four analyses using MeCS software,[31] which was designed to account for the highly correlated nature of molecular cis-QTL effects between independent studies. Prior to calling significance and defining mQTL sets, we ensured that the distribution of uniform (expected) -$\log_{10}$ p-values vs. computed (observed) -$\log_{10}$ p-values for each of the mQTL analysis showed no systematic inflation at low p-values (Figures S1J–S1M). As a point of reference, deviation from a uniform distribution is typically observed at -$\log_{10}$ p-values of 2 in cis-mQTL analysis in blood, and in past mQTL analysis in placenta.[13,36]

Within each analysis, we retained significant mQTL, or mQTL with a significant genotype by sex interaction effect in the interaction analysis, where significance was declared at a Bonferroni corrected p < 0.05. From these four analyses, we defined six sets of mQTL: (i) cross-sex; (ii) sex-dependent; (iii) male-stratified; (iv) female-stratified; (v) male-specific; and (vi) female-specific. The (v) male- and (vi) female-specific mQTL were those with significant stratified and interaction effects (i.e., the intersection of sex-dependent and male- or female-stratified mQTL). The reported effect size for each male- and female–specific mQTL was taken from male- and female-stratified mQTL. In counting mQTL, associated CpG sites, and the genes to which CpG sites were annotated, SNPs with a MAF < 0.05 were excluded in order to ensure mQTL were less likely to be a result of outliers in DNAm.

Assessing overlap of placental mQTL with mQTL in other prenatal tissues

As a main focus of this study is outlining the developmental origins of complex traits, we compared placental mQTL to mQTL from other tissues collected before and after birth using π_1 statistics. This assesses the proportion of placental mQTL effects that are similar to mQTL computed in other tissues.[27,28] Here we detail how mQTL were computed in these other tissues in comparison to our own mQTL protocol.

We accessed summary statistics for umbilical cord blood mQTL from the accessible resource for integrative epigenomic studies (ARIES; N=711; available via http://www.mqtldb.org/), originally from the Avon longitudinal study of parents and children,[32] and fetal brain mQTL summary statistics from the human developmental biology resource[113] (HDBR; N=173; available via https://epigenetics.essex.ac.uk/mQTL/). The ARIES study consisted of 51% male samples, and in mQTL mapping authors accounted for sex, 10 ancestry PCs, batch, and cell-type proportion estimates for white blood cell counts. mQTL in the ARIES study were called at a $p < 1 * 10^{-14}$ for all pairwise associations between 8,074,398 imputed SNPs at an MAF > 0.05 and 395,625 CpG sites passing their QC (i.e., cis- and trans-mQTL effects were identified). This threshold is more stringent than our Bonferroni corrected alpha < 0.05 computed for associations in cis, which corresponds to a threshold of $p < 3.77 * 10^{-10}$. Meanwhile, the HDBR study in homogenized fetal brain tissue consisted of 54% male samples ranging from 8-24.1 weeks post conception. When mapping mQTL, they accounted for sex, age, and 2 genotyping PCs. Additionally, they had a more stringent methylation QC and included only directly genotyped SNPs at MAF > 0.05, resulting in 430,304 SNPs and 314,554 CpG sites tested for all pairwise associations (i.e., cis- and trans-mQTL effects). mQTL were called at a Bonferroni-corrected threshold of $p < 3.69 * 10^{-13}$. Notably, the HDBR study excluded the sex chromosomes from their mQTL analysis.

Assessing enrichment of placental mQTL in regulatory regions across tissues using GARFIELD

We used GARFIELD[38] to quantify the enrichment of mQTL in different genomic regions across tissues, accounting for linkage disequilibrium (LD) and redundant annotations (i.e., annotations with similar enrichment to one another across mQTL). We used the genomic regions from the ENCODE and NIH Roadmap Epigenomics projects, which were provided as defaults by the software developers, and which excluded the HLA region.[39–42] As our mQTL are derived from subjects of diverse ancestry, we computed LD tag SNPs using all 1000 Genomes subjects, which include representatives of all ancestries predicted for both NICHD and RICHS subjects.[29,38,86] GARFIELD was run for cross-sex, sex-dependent, male-stratified, and female-stratified mQTL sets by mapping each mQTL to its minimum p value regardless of the probe with which it was associated. As male- and female-specific mQTL were defined on two sets of p values, we elected to exclude them for this analysis. We then ran GARFIELD considering only mQTL that were significant at a threshold of $1 * 10^{-9}$ (accounting for the roughly 100 million SNP-CpG pairs tested in cross-sex, sex-dependent, male-stratified, or female-stratified mQTL analysis). To simplify our results, rather than reporting the enrichment result from each regulatory mark measured for each individual cell line, we chose to report the mean odds ratio of a regulatory mark across all cell lines taken from a single tissue.

Linkage disequilibrium score regression analyses

We estimated GWAS trait h^2_{SNP} using linkage disequilibrium score regression (LDSC) software[22,23,84,85] (https://github.com/bulik/ldsc). We then partitioned h^2_{SNP} by placental mQTL sets using stratified LDSC (S-LDSC). S-LDSC is most powerful when the annotation (in this case, mQTL sets) cover at least 1% of the genome. Calling mQTL at Bonferroni-corrected p < 0.05 resulted in too few mQTL to meet this criterion, so we expanded our mQTL set using a previously developed method.[24] Briefly, for each CpG site with at least one mQTL at an FDR < 0.05, we fine-mapped meta-analyzed mQTL associations with a nominal p < 0.05 using CAVIAR.[87] The output from CAVIAR is a set of SNPs with a 95%

likelihood of containing the SNP causal to changes in DNAm (i.e., 95% credible set). As fine-mapping for interaction effects is difficult to interpret, we did not consider sex-dependent mQTL as its own category for this analysis. Instead, for CpG sites with at least one male-stratified or male-specific mQTL, as well as female-stratified or female-specific mQTL, we ran fine-mapping with male- and female-stratified mQTL p values respectively. We chose not to use the maximum causal posterior probability (max CPP) measure for each SNP as the weighted sum of our male- and female-specific annotations was small, less than 50,000 compared to the over 1 million per annotation that we achieved when using the 95% credible set (i.e., assigning 95% credible set SNPs a weight of 1, Figures S3E and S3F). As the mQTL used in this annotation are ultimately called at a less stringent criteria (FDR < 0.05, followed by fine-mapping) than the significant mQTL called at a Bonferroni-corrected p < 0.05, we elected not to analyze these mQTL beyond their use in S-LDSC. Nonetheless, all annotations are publicly available for download (key resources table).

Once we had defined the mQTL sets for use in S-LDSC, the stratified LD scores for our annotations (i.e., mQTL sets) were built using all subjects available in the 1000 Genomes Phase 3 EUR population, the publicly available baseline v2.2 LD scores (consisting of 97 annotations) provided by the software developers, and excluding SNPs within the MHC region of the human genome.[22,23] The EUR population was chosen as opposed to a transancestral population as all GWAS available for our traits of interest were conducted in European populations (see Availability and criteria for GWAS summary statistics, below), and a core assumption of LDSC is that LD patterns are similar across the GWAS population and population in which scores are computed.[22–24,85]

We report enrichment as the proportion of h^2_{SNP} explained by the SNPs in each mQTL set, divided by the proportion of SNPs included in the mQTL set. We also quantify enrichment using τ^* (Table S5), which is defined previously as:[85]

$$\tau^* = \frac{\tau s.d.(c)}{h^2_{SNP}/M}$$

Where h^2_{SNP} is the SNP heritability of the trait, s.d.(c) is the standard deviation of the annotation, τ is the coefficient of that annotation, and M is the number of variants used to compute h^2_{SNP}. Thus, tau can be interpreted as the standardized average contribution of SNPs in an annotation to the total h^2_{SNP} of a trait. An advantage of τ^* over enrichment is that it quantifies effects that are specific to that annotation, i.e., after taking into account the overlap between the annotation of interest, and other annotations in the model. An elevated tau value suggests that enrichment is not explained by overlap with other annotations in the model. Negative τ^* values indicate that the annotation on its own reduces h^2_{SNP}, on average.

S-LDSC was run separately for cross-sex, male- and female-specific, and male- and female-stratified mQTL, accounting for the 97 baseline annotations. Following S-LDSC analysis in single traits, we meta-analyzed enrichment and τ^* estimates across trait groups using the *rmeta* package.[114] Estimates from trait categories with sex-stratified GWAS summary statistics available were analyzed with mQTL annotations matching the sex of their component GWAS (GWAS labeled _female or _male in Table S4). τ^* and its standard error from meta-analysis were used to generate a normally distributed *Z* score for computing meta-analyzed p values. No such procedure is currently defined for computing significance from meta-analyzed S-LDSC enrichment values, so significance is not computed.

Colocalization of placental mQTL and GWAS loci

We performed a colocalization analysis to characterize the extent to which the proportion of h^2_{SNP} attributed to placental mQTL corresponded to shared genetic variants between mQTL and GWAS loci. We focused on all sets of SNPs within 75kb of CpG sites with a mQTL associated at a p < 5e-8, that were also associated with a GWAS trait at p < 5e-8, and we computed colocalization using the *coloc* R package.[57] Notably, unlike our protocol in S-LDSC, we included the HLA region in colocalization analysis given that this analysis is not dependent on genome-wide patterns of LD, and the HLA region is of particular interest to the immune-related traits analyzed. All genes considered to be within the HLA region were selected from chr6:28,477,797-33,448,354 in genome build hg19, and were gathered from their corresponding table on the UCSC genome browser at http://genome.ucsc.edu.[115] Colocalization was computed with cross-sex GWAS summary statistics. We defined colocalization having an overall $P(H_4) > 0.9$ per CpG site, which corresponds to the likelihood of both the level of DNAm and the trait sharing a single causal genetic variant in that region. Coloc assumes a single causal variant per colocalized site. A recent modification to coloc, called Sum of Single Effects (SuSiE), allows for multiple causal variants.[116] Since we were primarily interested in whether there was *any* match in GWAS-associated SNPs and mQTL across many candidate loci, we opted to only use coloc. While this approach may have missed some colocalized SNPs, it does not result in any additional false positives.[116] However, if a particular variant in *cis* to a given CpG site is of interest, we suggest applying SuSiE to our set of mQTL in order to capture all possible colocalized SNPs.

In plotting these results (Figure 6A), we wanted to highlight the colocalizations that were found only through sex-stratified analyses. Therefore, all colocalizations identified in the cross-sex analysis were labeled as such, even if they were also identified in the male- and female-stratified analyses. A small number of colocalizations (N=30 CpG sites) were identified in both male- and female-stratified analyses (annotated to N=4 genes unique to this set), but not in cross-sex analyses. We label these "Male- and Female-specific" colocalizations. Finally, we plot colocalizations that were found in the male- and female-stratified analyses only, which we refer to as male-specific and female-specific colocalizations respectively.