

## ARTICLE OPEN



# Transcriptome-wide association study for postpartum depression implicates altered B-cell activation and insulin resistance

Jerry Guintivano<sup>1</sup>✉, Karolina A. Aberg<sup>2</sup>, Shaunna L. Clark<sup>3</sup>, David R. Rubinow<sup>1</sup>, Patrick F. Sullivan<sup>1,4,5</sup>, Samantha Meltzer-Brody<sup>1</sup> and Edwin J. C. G. van den Oord<sup>2</sup>

© The Author(s) 2022

Postpartum depression (PPD) affects 1 in 7 women and has negative mental health consequences for both mother and child. However, the precise biological mechanisms behind the disorder are unknown. Therefore, we performed the largest transcriptome-wide association study (TWAS) for PPD (482 cases, 859 controls) to date using RNA-sequencing in whole blood and deconvoluted cell types. No transcriptional changes were observed in whole blood. B-cells showed a majority of transcriptome-wide significant results (891 transcripts representing 789 genes) with pathway analyses implicating altered B-cell activation and insulin resistance. Integration of other data types revealed cell type-specific DNA methylation loci and disease-associated eQTLs (deQTLs), but not hormones/neuropeptides (estradiol, progesterone, oxytocin, BDNF), serve as regulators for part of the transcriptional differences between cases and controls. Further, deQTLs were enriched for several brain region-specific eQTLs, but no overlap with MDD risk loci was observed. Altogether, our results constitute a convergence of evidence for pathways most affected in PPD with data across different biological mechanisms.

*Molecular Psychiatry* (2022) 27:2858–2867; <https://doi.org/10.1038/s41380-022-01525-7>

## INTRODUCTION

Postpartum depression (PPD), a diagnostic subtype of major depressive disorder (MDD) that occurs in the postpartum period, is a common complication of the perinatal period. It affects approximately half a million women annually in the U.S [1–4] and is one of the greatest causes of maternal morbidity and mortality [5, 6]. Additionally, PPD can have long-term adverse consequences on the newborn [7–11]. Despite this impact to public health, there is a lack of studies investigating the biology behind PPD. Although the precise mechanisms are unknown, PPD is a complex disorder that likely involves the culmination of genetic risk factors, response to hormonal fluctuations, and environmental factors.

Pregnancy is characterized by dynamic physiological changes that are expected to return to pre-pregnancy levels during the postpartum period. The stress axis, reproductive system, glucoregulation, and immune activation are a few examples of biological systems that interact and adapt to support the growing fetus. Perturbations in the recovery of these systems after childbirth, in addition to other risk factors, could result in PPD symptoms. Performing transcriptome-wide association studies (TWAS) allows for the interrogation of functional changes associated with case status. Employing TWAS, we can identify the biologically relevant changes in PPD that result from the aforementioned system disruptions, providing crucial insight into specific causes of the disorder.

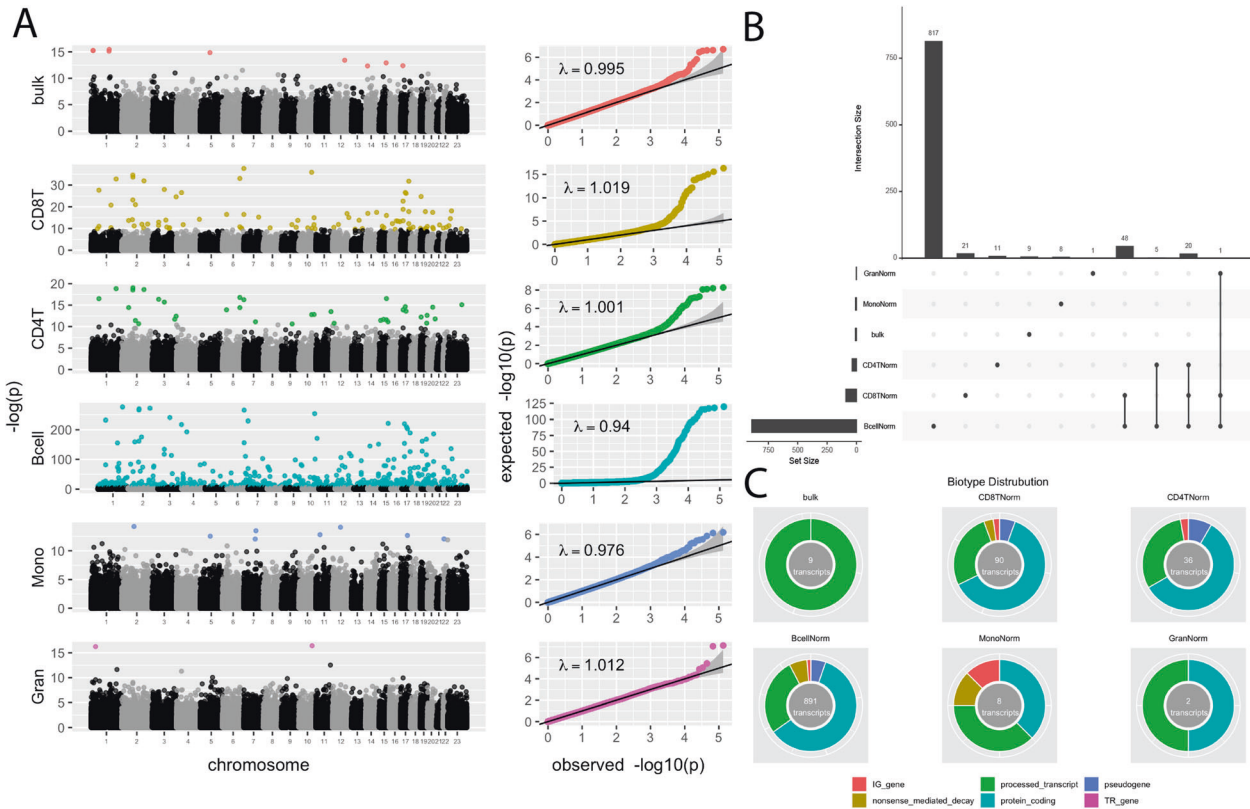
Traditionally, TWAS have been performed on bulk tissues that are composed of multiple diverse cell types. This cellular heterogeneity has a detrimental impact on the ability to detect disease associations [12]. Thus, in bulk tissue, case-control differences will be “diluted” when they affect only one or few cell types, may cancel out if the differences are of opposite signs across cell types, or may be undetectable altogether if the differences involve low abundant cells. Furthermore, identifying the specific cell types from which the association signals originate is key to formulate refined hypotheses of PPD disease pathology, designing proper follow-up experiments, and to develop effective clinical interventions. Efforts to address these issues with cell type-specific effects have been attempted using purified cell populations or single-cell RNA sequencing. However, these methods are labor-intensive and/or cost-prohibitive for most large-scale transcriptomic interrogation. As an alternative, statistical methods have been developed to deconvolute effects of individual cell types using data generated from bulk tissue [12–17].

Transcriptomic information can also be combined with other data types leading to deeper mechanistic insight into the regulation of transcription. For example, single nucleotide polymorphisms (SNPs) have been shown to regulate expression in a cell type-specific manner [18, 19]. In addition, DNA methylation can regulate gene transcription [20, 21]. Identifying the (epi-) genomic regulators of transcriptional differences is a key step for generating

<sup>1</sup>Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. <sup>2</sup>Center for Biomarker Research and Precision Medicine, Virginia Commonwealth University, Richmond, VA, USA. <sup>3</sup>Department of Psychiatry & Behavioral Sciences, Texas A&M University, College Station, TX, USA. <sup>4</sup>Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. <sup>5</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. ✉email: [guinti@email.unc.edu](mailto:guinti@email.unc.edu)

Received: 29 September 2021 Revised: 8 February 2022 Accepted: 9 March 2022

Published online: 1 April 2022



**Fig. 1 PPD TWAS cell type-specific associations.** **A** Manhattan plots and QQ-plots of PPD transcriptome-wide association study (TWAS) results in bulk tissue and individual cell types. Non-significant transcripts are shown in grey/black. Significant transcripts at an FDR < 0.1 are shown in color. **B** Upset plot showing overlap of significant TWAS transcripts across bulk tissue and individual cell types. **C** Biotype distribution of significant TWAS transcripts in bulk tissue and individual cell types.

novel hypotheses about PPD disease etiology. This would allow, for example, designing functional follow-up studies. The identification of (epi-) genomic regulators also has translational value as they are potential targets for correcting aberrant transcription.

In this work, we performed the largest TWAS for PPD, using RNA-sequencing of whole blood, in a cohort of women six-weeks following childbirth. To date, three TWAS studies have been performed with sample sizes ranging from 6 to 15 cases and 10 to 122 controls [22–24]. Further, the analyses presented here are performed on a cell type-specific level. Additionally, SNPs, DNA methylation, and hormone levels were used to identify regulators underlying the observed case-control transcriptional differences. Altogether, this represents the largest and most diverse interrogation into the biology of PPD.

**RESULTS**

We recruited a case-control cohort of 1551 women (579 PPD cases, 972 controls) at six weeks postpartum with PPD case-control status established using clinical interviews. Participants were racially (66.5% Black, 32.9% White, and 0.6% Asian) and ethnically (15.9% Hispanic) diverse (Table S1) [25]. We generated whole blood-derived transcription data using RNA-sequencing (RNA-seq), which resulted in 134,302 known transcripts from 51,079 genes (88.2% of all Ensembl annotated genes) [26]. In addition, to study regulators of PPD expression differences we generated array-based DNA methylation data (Illumina 450k), assaying 373,635 CpGs, and genome-wide SNP data which, with imputation, assayed ~12.5 million variants. Cell type proportions were estimated from DNA methylation data using standard methods [14, 27]. The average estimated cell type proportions were 9.7% CD8<sup>+</sup> T-cells, 16.3% CD4<sup>+</sup> T-cells, 5.6% B-cells, 4.7%

monocytes, 59.4% granulocytes, and 3.0% natural killer cells. The validity of the estimated proportions was established through high correlations with complete blood counts available for a subset of participants (Supplemental Results) [14, 27]. We observe differences in cell proportions between cases and controls for CD4<sup>+</sup> T-cells ( $\beta = -0.06, p = 0.03$ ) and granulocytes ( $\beta = 0.06, p = 0.02$ ), but not for the other cell types (Table S1). All downstream analyses control for cell type proportions.

**TWAS for PPD identifies cell type-specific dysregulation**

In addition to TWAS of whole blood, we performed cell-type-specific TWAS using a statistical deconvolution approach [12, 16, 18, 28, 29]. Figure 1A shows Manhattan plots, QQ-plots and lambdas for the TWAS results for whole blood and each cell type separately. Across all analyses, we correct for multiple testing using false discovery rate (FDR), with a q-value < 0.1 threshold indicating significance. Decreasing the q-value to 0.05 would, for example, result in only a modest reduction in false positives but decreases the probability of finding true effects exponentially as shown previously [30]. The TWAS in whole blood and in the different cell types each revealed transcripts that were significantly associated with PPD (see Tables S2–S7 for all transcripts with p-value < 0.05). The QQ-plots and lambdas from the TWAS results (Fig. 1A) in combination with TWAS of permuted case-control status for each analysis (Fig. S1), which yielded average lambdas of approximately one, indicated no evidence of test-statistic inflation.

*Whole blood.* TWAS in whole blood revealed nine transcripts from nine genes that were significantly associated with PPD (Table S2). The most significant transcript was for the small nuclear RNA, *RNVU1-9* ( $p$ -value =  $1.92 \times 10^{-7}$ ). Using the Ensembl

biotype annotations [26], which provide an indication of transcript function, we see that all nine transcripts are processed transcripts (Fig. 1C).

**Granulocytes.** TWAS in granulocytes had two transcripts from two genes significantly associated with PPD case status: a protein-coding transcript for *SH3PXD2A* ( $p$ -value =  $8.81 \times 10^{-8}$ ) and a processed transcript for *CSF3R* ( $p$ -value =  $7.45 \times 10^{-8}$ ) (Fig. 1C; Table S3).

**Monocytes.** Monocyte TWAS resulted in eight transcripts from eight genes that reached transcriptome-wide significance (Table S4). Transcript biotypes are varied with three protein-coding, three processed transcripts, one nonsense-mediated decay, and one immunoglobulin (IG) gene (Fig. 1C). The top hit ( $p$ -value =  $6.60 \times 10^{-7}$ ) is a protein-coding transcript for the gene *VWA3B*.

**CD8<sup>+</sup> T-cells.** There were 90 transcripts in 87 genes that reached transcriptome-wide significance in the TWAS of CD8<sup>+</sup> T-cells (Table S5). These transcripts are composed of 56 protein-coding, 24 processed transcripts, five pseudogenes, three nonsense-mediated decay, and two IG genes (Fig. 1C). In this cell population, a protein-coding transcript for the gene, *HEATR2*, was the most significantly associated transcript ( $p$ -value =  $6.46 \times 10^{-12}$ ). We found that the significant protein-coding transcripts were enriched for 18 pathways, which group into three clusters related to protein secretion, signal transduction, and neuronal processes (Table S8).

**CD4<sup>+</sup> T-cells.** In CD4<sup>+</sup> T-cells, 36 transcripts from 35 genes passed transcriptome-wide significance (Table S6). These significant transcripts included 21 protein-coding, 11 processed transcripts, three pseudogenes, and one IG gene (Fig. 1C). *IGKV1D-13*, was the top transcript ( $p$ -value =  $2.94 \times 10^{-4}$ ). Top protein-coding transcripts from the CD4<sup>+</sup> T-cell TWAS were enriched for 19 pathways, which segregate into three clusters related to neuronal development, signal transduction, and cell receptor signaling (Table S9).

**B-cells.** The TWAS in B-cells yielded 891 transcripts representing 789 genes that were transcriptome-wide significant (Table S7). The altered transcripts are comprised of 534 protein coding, 242 processed transcripts, 56 nonsense-mediated decay, 46 pseudogenes, 12 IG genes, and one T-cell receptor (TR) gene (Fig. 1C). The most significant ( $p$ -value =  $1.76 \times 10^{-120}$ ) is the only protein-coding transcript for the gene *FMOD*. Pathway analyses of the protein-coding findings show enrichment for 98 pathways. Pathway clustering results in ten clusters including those related to B-cell activation, apoptotic pathways, cellular starvation, cellular metabolism, nucleic acid metabolism, neuron cell morphology, organ development, glucose metabolism, and cell adhesion and cytoskeleton organization (Table S10).

Overall, individual cell types have unique profiles of transcripts that are differentially expressed between cases and controls (Fig. 1B). However, an overlap of 20 significant transcripts is seen among CD4<sup>+</sup> T-cells, CD8<sup>+</sup> T-cells and B-cells, which all are lymphocytes. These overlapping differently expressed transcripts could reflect common functions associated to PPD that are shared by the different lymphocytes cell types. Among the overlapping

transcripts are three pathway clusters that involve nervous system development, regulation of signal transduction, and regulation of macromolecule metabolism (Table S11).

### Top Loci in Whole Blood Overlap Findings From Previous Transcriptome Studies of PPD

Three other transcriptome studies of PPD have been performed [22–24]. These previous studies were performed using whole blood and are independent of the current study, which allowed us to test whether our top bulk results overlap the most significant genes in the previous studies.

As shown in Table 1, we found the top 5% of our whole blood results shared genes with the significant genes reported by Landsman et al. Further, for the remaining two studies (Pan et al. and Mehta et al.) we observed overlap between the top 5% of our whole blood results and the top 5% of the reported genes. We used sign tests to compare the overall patterns of results between our results and the previous PPD TWAS. Under the null, the expectation is that 50% of the signs of the overlapping genes will be the same between two independent sets of results. The significance of the observed proportion was evaluated using the binomial distribution. From this, we find that our results were concordant with results by Mehta et al. ( $p < 0.001$ ), which may be due to the small sample sizes of the other studies.

### SNPs and DNA methylation, but not hormones, may regulate PPD-associated transcripts

Whole blood and cell type-specific TWAS identified multiple transcripts that differed between cases and controls. We performed regulation analyses to study whether these differences were regulated by pregnancy-related hormones (estradiol, E2; progesterone, P4; oxytocin; BDNF), DNA methylation, or SNPs measured in the same samples. These analyses followed the model depicted in Fig. S4, which assumes the tested marker regulates transcription (a mediator), which in turn alters PPD risk. As there may be additional mechanisms through which the marker may affect PPD, the model also allows for a direct effect of the marker on PPD. The null hypothesis states that the marker does not regulate differentially expressed transcripts ( $H_0: a \times b = 0$ ). SNPs and CpGs were tested as putative regulators if they had a nominal association with case status ( $p < 0.05$ ) and were within a 10 kb window (cis-acting elements) of the genes tested. Hormones were tested as putative regulators if they had a nominal association with case status ( $p < 0.05$ ).

The postpartum period is characterized by a large fluctuation in hormones, making dysregulation as a result of hormone changes an attractive mechanism. We observed significant association between PPD case status and oxytocin levels ( $p$ -value =  $2.38 \times 10^{-4}$ ), but not for any other measured hormones (Table S1). However, when we examine oxytocin as a potential regulator of transcriptional differences, we did not observe any significant effects (Table S12). This is not to say hormones are not contributing to case-control differences, but we do not see supporting evidence that hormones regulate the specific transcriptional differences observed in this study.

Among the disease-associated DNA methylation sites (10 CpGs in bulk, 158 CpGs in CD8<sup>+</sup> T-cells, 67 CpGs in CD4<sup>+</sup> T-cells, 1,520 CpGs in B-cells, 13 CpGs in monocytes, 5 CpGs in granulocytes), we identified five CpGs that were significantly associated with four

**Table 1.** Overlap of top TWAS findings versus existing transcriptome studies of PPD.

Authors	Year	PMID	Cases	Controls	Tissue	Overlapping genes	<i>n</i> Shared	direction	sign test <i>p</i> -value
Landsman et al.	2017	27816578	6	10	PBMC	10	7		0.17
Pan et al.	2018	29973662	8	10	PBMC	19	7		0.92
Mehta et al.	2021	33664235	15	122	Whole Blood	2894	1793		<0.001

**Table 2.** Overlap of PPD deQTLs in bulk brain tissue or neurons from various brain regions.

	Enrichment Target	Enrichment OR	p-value
Bulk eQTLs	Amygdala	2.31	0.06
	Anterior Cingulate Cortex (BA24)	2.32	0.04
	Caudate Basal Ganglia	2.77	0.01
	Cerebellar Hemisphere	2.59	3.02E-03
	Cerebellum	2.58	2.19E-03
	Cortex	2.95	1.05E-03
	Frontal Cortex (BA9)	3.27	9.40E-04
	Hippocampus	2.54	0.03
	Hypothalamus	3.90	5.90E-04
	Nucleus Accumbens (Basal Ganglia)	2.88	2.43E-03
	Putamen (Basal Ganglia)	2.33	0.02
	Substantia Nigra	2.19	0.07
	Neuronal eQTLs	Amygdala	1.47
Anterior Cingulate Cortex (BA24)		0.73	0.74
Caudate Basal Ganglia		0.00	1.00
Cerebellar Hemisphere		0.76	0.72
Cerebellum		0.00	1.00
Cortex		1.55	0.37
Frontal Cortex (BA9)		1.53	0.37
Hippocampus		2.21	0.16
Hypothalamus		1.52	0.38
Nucleus Accumbens (Basal Ganglia)		1.53	0.37
Putamen (Basal Ganglia)		0.00	1.00

differentially expressed transcripts in B-cells ( $q$ -value  $< 0.1$ ; Table S13). Thus, these methylation marks may serve as potential cis-acting regulators for their corresponding protein-coding transcripts for *CD22*, *CXXC5*, *MYO1D*, and *KCNQ1* [31]. It should be noted that methylation was assayed with a commonly used array-based platform resulting in 373,635 high-quality methylation markers, which corresponds to  $\sim 1.3\%$  out of the 28.3 million CpGs in the human genome [32]. Thus, it is possible, and likely, that many additional regulatory methylation marks are present that were not assayed in our dataset.

Using our SNP data, we examined whether disease-associated eQTLs (deQTL) exist for any of the differentially expressed transcripts. We tested 36 SNPs in whole blood, 1696 SNPs in  $CD8^+$  T-cells, 322 SNPs in  $CD4^+$  T-cells, 15,154 SNPs in B-cells, 149 SNPs in monocytes, and 58 SNPs in granulocytes as putative regulators. No deQTLs were identified for whole blood and granulocytes. However, we detected 17 deQTLs for seven transcripts in  $CD8^+$  T-cells, seven deQTLs for three transcripts in  $CD4^+$  T-cells, 523 deQTLs for 124 transcripts in B-cells, and four deQTLs for two transcripts in monocytes ( $q$ -value  $< 0.1$ ; Table S14). The majority of the transcripts with deQTLs were protein-coding (78%). These deQTLs are genomic regulators of transcriptional differences associated with PPD and have translational value as potential targets for correcting aberrant transcription.

#### PPD deQTLs are enriched for multiple brain eQTLs but not for MDD GWAS loci

We tested whether the PPD deQTLs we identified were enriched for significant MDD GWAS loci [33, 34]. However, we did not observe any overlap between PPD deQTLs and significant MDD GWAS loci. The lack of overlap may be due to differences in ancestry between studies. Our PPD cohort is comprised mainly of Black and Hispanic women, whereas the MDD GWAS was limited to individuals of European ancestry. Alternatively, our results could

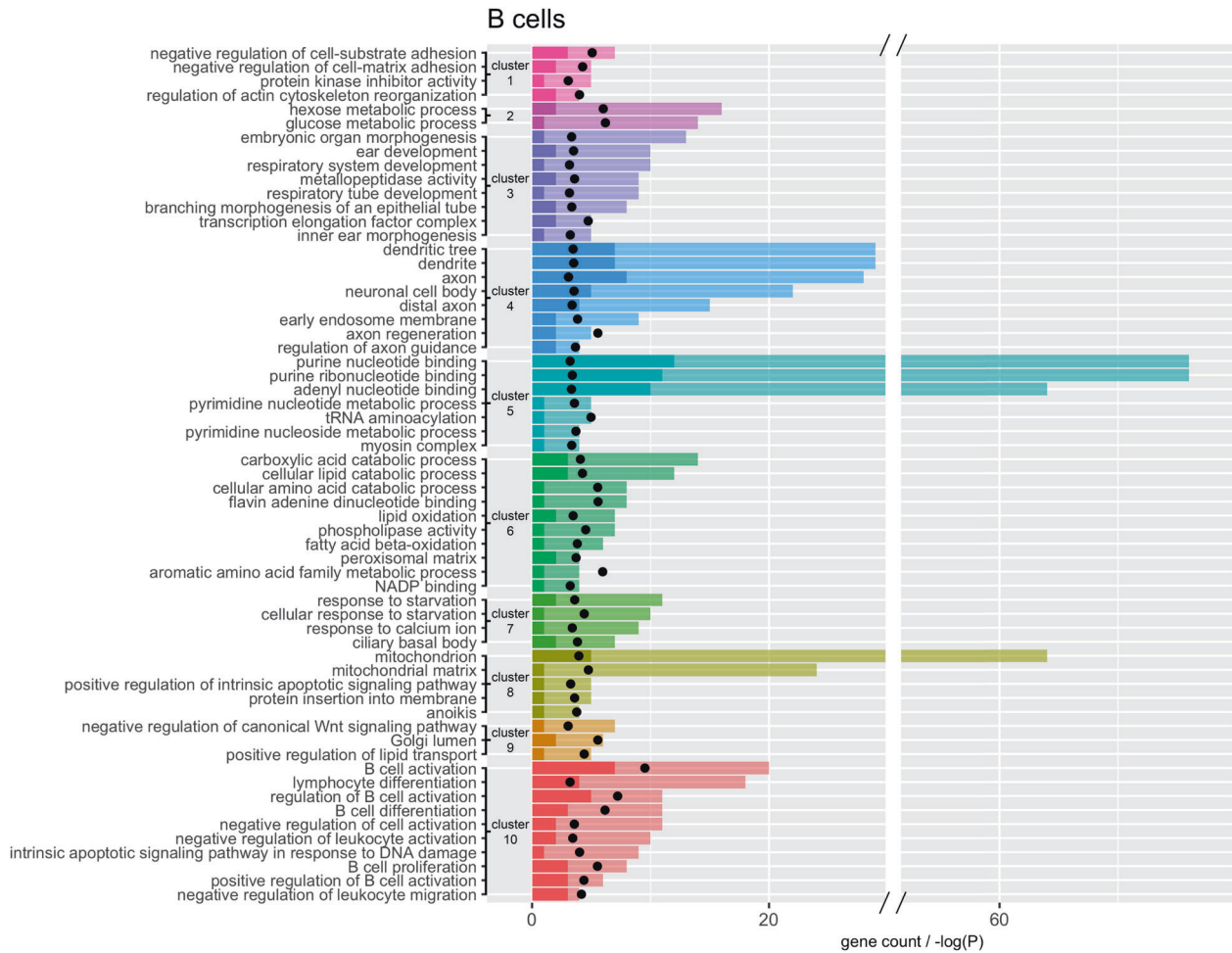
support literature that PPD is a distinct disorder with a different underlying etiology [35–38].

Further, to examine the potential impact of PPD deQTLs across tissues, we tested whether they overlapped eQTLs in bulk brain tissue or neurons from various brain regions as reported by the most recent GTEx analyses [18, 19]. We found significant enrichment with our deQTLs for eQTLs in nearly all brain regions (Table 2; 10 out of 12, 83.3%), with the exception of amygdala and substantia nigra. However, we did not identify any enrichment for eQTLs in neurons. Although additional investigations would be required, this could suggest that the PPD deQTLs detected in this study do not exert their effects on neurons, but rather non-neuronal cell types, such as glia. In total, there were 45 deQTL containing genes overlapping with eQTLs in various brain tissues (Table S14).

#### Convergent evidence implicates pathways most affected in PPD

Pathway analyses of TWAS results identified pathways that are altered in PPD cases compared with controls. However, RNA transcription is potentially regulated by many biological processes. Our analyses suggested that SNPs (deQTLs), and to a lesser extent DNA methylation, may be involved in the regulation of PPD-related transcriptional differences. By examining the overlap of these pathways with deQTLs we can identify which pathways are initially disrupted by genetic loci in cases compared to the effect of other factors. We identified 138 transcripts regulated by deQTLs in whole blood and across the five cell types examined. More specifically, 124 of these transcripts were found in B-cells and overlap pathways in every pathway cluster identified. Figure 2 illustrates the B-cell pathways that contain at least one deQTL containing gene organized by cluster. Additional results for other cell types can be found in Supplemental Tables. Overall, our results suggest the pathway





**Fig. 2 Clusters of significant TWAS pathways in B-cells that contain at least one deQTL gene.** Each cluster is represented by a different color. Opaque bars are the total number of genes overlapping the pathway. Solid bars are the number of genes with a deQTL. Black points are  $-\log(p\text{-value})$  for the pathway enrichment.

clusters shown in Fig. 2 may be potentially dysregulated due to genetic factors in women with PPD.

**DISCUSSION**

To generate novel hypotheses for PPD disease pathology, we studied the biological underpinnings of PPD in a large cohort of women six weeks after childbirth. Results showed cell type-specific transcriptional differences associated with PPD, with a majority of the changes seen in B-cells. Furthermore, these associations were significantly overrepresented in multiple sets of pathways. These pathways reflected the significant effects of SNPs regulating the PPD-associated transcriptional changes. This constitutes a convergence of evidence with data from two different biological mechanisms.

Pregnancy is characterized by substantial changes in multiple physiological systems. Failure to return to pre-pregnancy levels during the postpartum period may contribute to PPD symptoms, making these systems candidate mechanisms for PPD. Our association and pathway results implicate two, potentially co-occurring, such mechanisms: B-cell activation and insulin resistance (IR). In addition to showing pregnancy-related changes, both mechanisms have previously been linked to depression [39–44].

Our results specifically implicate B-cell activation (Fig. 2, cluster 10), which plays a critical role in the immune system. B-cells become activated when their receptor recognizes an antigen and binds to it. Activated B-cells then produce antibodies, along with

secreting pro- and anti-inflammatory factors. During pregnancy, B-cells undergo dynamic changes as the maternal immune system has to balance tolerance of the foreign-growing fetus with maintaining vigilance against pathogens [45, 46]. Thus, B-cell concentrations are significantly lower during the third trimester and immediately following delivery compared to non-pregnant women, but levels typically return to those seen in non-pregnant women by six-weeks postpartum [47].

A growing body of evidence suggests that inflammatory processes may play a significant role in PPD [24, 48–50]. However, the specific role of B-cells has yet to be elucidated. Recent work has shown an increase in B-cell densities in the brains of those with mood disorders compared to controls [39]. Furthermore, in whole blood, altered B-cell homeostasis was observed in those with MDD compared to controls [40, 41]. A possible mechanism contributing to increased B-cell activation could be related to autoimmunity [51]. Depression is often co-morbid with autoimmune disease; risk of depression is 1.25–3.56 times higher in people with autoimmune disease than without [52–54]. Additionally, a feature of many autoimmune disorders is a loss of B-cell tolerance coinciding with the inappropriate production of autoantibodies [51, 55]. Thus, an aberrant autoimmune response could potentially contribute to PPD.

Further, we did not observe significant differences in B-cell proportions between cases and controls ( $p\text{-value} = 0.78$ ). As multiple subtypes of B-cells exist, it may be that we did not observe

differences in overall B-cell proportions but there may be differences in more specific B-cell subsets.

Not only do we observe a pathway cluster, composed of 12 pathways, directly related to B-cell activation (cluster 10), we see multiple pathway clusters associated with cellular metabolism, which supports our hypothesis of B-cell activation. Activation initializes cellular reprogramming of quiescent naïve B-cells to drive re-entry into the cell cycle [56]. This rapid expansion requires the production of biomolecules (lipids, proteins, nucleotides in clusters 2, 5, 6, and 9) at an increased rate. Additionally, work in mice has shown that B-cell activation results in increased glucose uptake (cluster 2) and mitochondrial remodeling (cluster 8) [56]. Upon B-cell activation, not only is there a slew of metabolic changes, but there are changes to the cellular structure (cluster 1). Antigen binding triggers substantial remodeling of the cell cytoskeleton, which induces cell spreading, the formation of the immune synapse, and the gathering of antigen for endocytosis [57]. Additionally, apoptosis is a carefully regulated process through the lifecycle of B-cells. Disruptions to apoptotic pathways (cluster 8) affect multiple processes including homeostasis, quality control of the antibody response, and tolerance [58].

The second implicated mechanism, insulin resistance (IR), is supported by several factors. Insulin promotes the absorption of excess blood glucose into other tissues for energy storage. IR occurs when cells become insensitive to the effects of insulin leading to a buildup of blood glucose and insulin. Starting in the second trimester of pregnancy, insulin sensitivity is progressively reduced as much as 60–80% [59]. This coincides with steady increases in insulin [60]. These changes serve as a physiological adaptation of the mother to ensure adequate carbohydrate supply for the rapidly growing fetus [61]. After delivery, insulin returns to pre-pregnancy levels [62–64].

IR is a risk factor for depression. Rodent studies have shown that brain IR alters dopamine turnover and induces anxiety and depressive-like behaviors in mice [65]. In humans, greater glycemic variability has been associated with negative moods [66]. IR typically predates the development of diabetes. A meta-analysis of 27 studies demonstrated that depression is significantly associated with hyperglycemia for both type 1 and type 2 diabetes [67]. Studies further suggest that insulin-sensitizing agents could play a significant role in the treatment of major depression, particularly in patients with documented IR [68, 69]. Pregnancy is known to increase the risk of developing Type 2 diabetes after giving birth [70]. Furthermore, pre-pregnancy or gestational diabetes was independently associated with perinatal depression, including new onset of PPD [71–73].

We tested whether genes implicated by our top results were significantly overrepresented for genes related to A1C [74] and IR [75] in whole blood. Hemoglobin A1C levels are measure of a person's blood sugar levels over the past three months and are highly correlated with measures of IR [76]. We found the top 5% of our whole blood findings for PPD were enriched for the top 5% of associations with A1C ( $p$ -value =  $4.79 \times 10^{-7}$ ) and IR ( $p$ -values = 0.04). Databases can not directly implicate IR as such a pathway does not exist. IR is a disorder characterized by disruptions of multiple biological functions. However, IR can be implicated by nearly all clusters in our pathway analyses (Fig. 2). With the evidence linking IR and B-cells [77, 78], it is reasonable to observe a signature of IR in B-cells. For example, B-cells contribute to the development of IR (cluster 10). These cells can promote IR through T-cell modulation and production of pathogenic antibodies [77, 78]. Insulin signaling regulates diverse cellular functions including metabolic pathways, apoptosis, mitogenesis, and membrane trafficking through protein kinases (cluster 1) [79, 80]. Insulin directly affects glucose metabolic processes (cluster 2). Circulating levels of purines (cluster 5) [81, 82], amino acids, and fatty acids (cluster 6) [83, 84] are also associated with IR. The administration of carboxylic acids (cluster 6) improved

glycemic control, potentially by reducing IR [85]. IR may lead to inadequate intracellular glucose potentially leading to apoptosis and intracellular starvation (cluster 7 and 8) [86]. Wnt signaling (cluster 9) is involved in the regulation of glucose homeostasis in multiple organs, particularly in insulin-responsive tissues [87].

A number of limitations of the present study should be mentioned. While we studied blood, the pathogenic processes for PPD most likely manifest in the brain. It is likely that in studying blood, other possible PPD-related mechanisms might have been missed. However, there is cross-talk between the two tissues across the blood-brain barrier [88]. This would allow altered B-cell activation and IR to affect the brain and be observed in our study. Furthermore, we observed deQTLs that affect genes in both blood and brain, specifically in brain regions implicated in mood disorders (e.g., hippocampus, cingulate cortex, frontal cortex). These deQTLs can be studied in model systems for functional follow-up to evaluate causality and their downstream biological effects [89]. Additionally, the B-cell activation and IR hypotheses for PPD requires further validation through direct measurements in PPD cases versus controls.

In conclusion, we have performed the largest and most comprehensive biological interrogation of PPD, to date. Our results suggest that PPD is associated with an increase in B-cell activation, a finding previously unreported in the literature. While we do not know the precise mechanisms behind this increase in B-cell activation, we hypothesize it could be due to co-occurring dysregulation in IR. Additionally, we identified genetic variants, deQTLs, that regulate, in part, the transcriptional differences between cases and controls. Our findings require further validation and follow-up studies. However, these novel hypotheses for PPD provide promising avenues for future research.

## METHODS

### Study population

Detailed information about the study can be found elsewhere [90]. Briefly, we followed the 2010 US Census terminology for describing the self-reported “race” and “ethnicity” (Hispanic or Non-Hispanic) of subjects. We refer to the participants as Asian, Latina (“of Latino, Hispanic, or Spanish origin”), Black (or African-American), and White (i.e., European ancestry, non-Hispanic).

Recruitment of postpartum women aged 17–45 years occurred from 9/2012 to 6/2016 in four outpatient obstetrical clinics (University of North Carolina Women's Hospital, Wake County Health Department, Alamance County Health Department, East Carolina University School of Medicine) during routine six-week postpartum visits ( $\pm 1$ –2 weeks). Detailed recruitment procedures can be found in Supplemental Methods.

Case-control status was determined using clinical interview. All women attending these clinics were first screened for study inclusion using the Edinburgh Postnatal Depression Scale (EPDS). The 10-item EPDS is a commonly used PPD screening instrument. High EPDS scores are consistent with a PPD diagnosis by structured clinical interview [91]. Women with high EPDS scores ( $\geq 11$ ) or low EPDS scores ( $\leq 7$ ) were invited to participate. All women then had PPD case status determined using the MINI diagnostic interview. For a full list of inclusion/exclusion criteria, see Supplemental Methods. Briefly, all participants included no indication of MDD during the first or second trimesters of pregnancy, singleton pregnancy, and live term birth ( $\geq 34$  weeks gestation). This study was approved by the University of North Carolina Institutional Review Board Committee for the Protection of Human Subjects. All subjects provided written informed consent and signed the Health Insurance Portability and Accountability Act release.

### Subject Assessments

All participants were administered the MINI International Neuropsychiatric Interview (MINI-Plus, version 6.0), a structured clinical interview for the assessment of psychiatric disorders [92, 93]. Experienced and certified ( $\kappa > 0.8$  versus criterion ratings) psychiatric research nurses working in each clinic setting administered the MINI-Plus. Cases for this study were defined by having current MDD as assessed by the MINI-Plus. Controls did not have

current MDD using the MINI-Plus. All study procedures could be performed in Spanish with a native speaker.

### Biological Sampling

Peripheral blood was sampled and immediately processed on-site at the time of subject assessment. For plasma separation, blood aliquots were centrifuged at 3300 rpm for 10 min at room temperature immediately after sampling. For serum separation, blood aliquots were centrifuged at 3300 rpm for 10 min at 2–8 °C, 60 min after blood draw. All plasma and serum samples were then snap-frozen and kept at –80 °C until analysis. Aliquots were drawn into PAXgene RNA tubes and stored at –80 °C until RNA extraction. Genomic DNA was extracted from aliquots of whole blood using Qiagen Autopure LS, which utilized Qiagen Puregene chemistry. Samples that were missing, had insufficient sample, or that did not meet minimum detection thresholds were excluded from analyses.

### RNA Sequencing

For RNA extraction, samples are pulled from –80 °C freezers, allowed to thaw at +4 °C overnight, and extracted using the QIA Symphony platform. A detailed RNA extraction protocol can be found in Supplemental Methods. Fresh-frozen total RNA was prepared for sequencing following the Nugen Ovation Human Blood RNA-seq library prep kit according to the manufacturer's instructions. RNA libraries were sequenced as 2 × 50 bp paired-end reads with 24 samples per lane on an Illumina HiSeq 4000 sequencer. Each sample was run on two different lanes at two different times. Samples were balanced by case status, age, race, ethnicity, and recruitment site across sequencing pools to reduce technical biases. Preliminary sample and read quality control (QC) was performed using FastQC using default settings. Briefly, raw sequence reads are read in and reports are generated on read quality and composition. No samples were dropped or required resequencing. Reads were aligned with HISAT2 (v2.1.0) and transcriptomes were reconstructed using StringTie (v.1.3.3), both within the macocktail pipeline [94]. Reads from runs 1 and 2 for every sample were merged prior to quantification. Reference transcriptome was downloaded from ENSEMBL (GRCh37, release 92; [http://ftp.ensembl.org/pub/grch37/release-92/gtf/homo\\_sapiens/Homo\\_sapiens.GRCh37.87.gtf.gz](http://ftp.ensembl.org/pub/grch37/release-92/gtf/homo_sapiens/Homo_sapiens.GRCh37.87.gtf.gz)) [26] and used for alignment, transcriptome reconstruction, and quantification steps. This reference includes all available biotypes including protein coding genes, as well as pseudogenes, lncRNA, and ncRNA. Following transcriptome assembly, the StringTie merge option was used to combine all assembled transcriptomes across all samples and then re-quantified against the merged transcriptome (stringtie -eB) so expression measures are consistent across all samples. Transcripts were excluded if they were depletion targets for library prep (ENSEMBL gene\_biotype "rRNA", *HBA1*, *HBA2*, *HBB*, *HBD*), unannotated (not associated with an ENSEMBL ID), present in < 1% of samples, had an average TPM < 1 (low expression outlier) or > 20,000 (high expression outlier). Following this quality control, data for 108,474 transcripts remained for association testing.

For association testing, technical variables were measured for each sample including: i) the total number of reads, the number of uniquely aligned reads, and the proportion of reads aligned using StringTie, ii) sequencing pool, and iii) calculation of the first ten principal components across all transcript counts for depletion targets for library prep (ENSEMBL gene\_biotype "rRNA", *HBA1*, *HBA2*, *HBB*, *HBD*). Final association models included maternal age, race/ethnicity, estimated cell proportions, proportion of reads aligned, number of uniquely aligned reads, and sequencing pool. Additionally, principal components of TPM values were used to capture any remaining unmeasured source of variation. One principal component (PC1) was included based on the scree test. Lastly, multidimensional outliers were excluded, resulting in data for 482 cases and 859 controls. Quantile-quantile plots for each cell type-specific TWAS (Fig. 1A), along with TWAS of permuted case-control status for each analysis yielded average lambdas of approximately 1 (Figure S1), indicated no evidence of test-statistic inflation under the empirical null.

### DNA Methylation Assessment

A detailed DNA methylation pipeline can be found in Supplemental Methods. DNA sample bisulfite conversion and microarray hybridization were through the Illumina Fast Track Genotyping service. DNA methylation was assessed using the Infinium Human 450k array. Quality control steps are described elsewhere [95]. Briefly, we employed a stringent quality control pipeline comprised of the following steps: i) removal of samples

with > 1% of probes with detection P-value > 0.001, ii) removal of probes with > 1% of samples with detection P-value > 0.001, iii) removal of cross-hybridizing probes, iv) and probes containing a SNP with minor allele frequency > 1% within 10 bp of the single-base extension position [96]. Normalization of the DNA methylation data was performed using the BMIQ function [97] within the minfi package. Following quality control of probes, data for 373,635 CpGs remained for regulation analyses.

Residuals were used for regulation (deQTL) analyses. Covariates were selected using multiple regression analyses in RaMWAS [98] from a pool of multiple types of variables (see Supplemental Methods for description of variables tested). Final association models included maternal age, race/ethnicity, estimated cell proportions, slide and array (batch), and median methylated and unmethylated signal intensities, and three PCs from raw control probes (PCs 2, 8, 10). Additionally, PCs of beta values were used to capture any remaining unmeasured source of variation. PCs 1–5 were included based on the scree test. As a final step, multidimensional outliers across PC1–15 were identified using the mvoutliers R package and excluded, resulting in data for 503 cases and 897 controls.

### SNP Genotyping

Genotypes were assessed using the Illumina Multi-Ethnic Genome Arrays (MEGA; Illumina, San Diego, CA, USA) through the Illumina Fast Track Genotyping service. GenomeStudio software version 2.0 (Illumina, San Diego, CA, USA) was used to call genotypes from raw Illumina data. We have described our quality control procedures for SNPs elsewhere [99, 100]. Briefly, SNPs are removed for bad genome mapping of array probe, missingness (>0.01), and low MAF (<0.01). Any individual with high missingness was excluded (>0.01). Genotypes were imputed against the Haplotype Reference Consortium [101] using the University of Michigan Imputation Server [102]. Following imputation, genotypes underwent another round of quality control. Genotypes were excluded for having low quality scores ( $r^2 < 0.8$ ), missingness (>0.01), and low MAF (<0.01). Again, any individual with high missingness was excluded (>0.01).

For deQTL analyses, principal components of SNPs were used to capture any unmeasured source of variation in our genetic data (e.g population stratification). Two principal components (PCs 1 and 2) were included based on the scree test, resulting in data for 487 cases and 864 controls.

### Cell type-specific analyses

Whole blood contains a mixture of different cell types. Cell type-specific TWAS are critical as expression changes may remain undetectable in whole blood as changes may cancel each other out when they have opposites affects across cell types or involve effects of low abundant cells that are diluted by effects of common cell types [12, 18]. Cell type-specific TWAS also improve the biological interpretation of findings as knowing the cell type in which the change occurred may provide further clues about the underlying biological mechanisms. Therefore, we performed cell type-specific deconvolution analyses, which were first introduced about 20 years ago [103]. Most of the initial deconvolution papers include sections showing the validity of the approach. For example, the 2010 paper by Shen-Orr [29] experimentally validated the method through tests with pre-designed mixtures.

For TWAS, cell type-specific analyses were conducted using a deconvolution approach described and validated previously [16, 28, 29]. Briefly, cell proportions are estimated from bulk (cellularly heterogenous) DNA methylation data using available reference panels [27]. These predicted cell type proportions are then used to test case-control differences on a cell type-specific level using all study samples with available bulk transcription data. The statistical model we use is:

$$Y^{bulk} = \sum_{c=1}^{n_c} m_c P_c + \sum_{c=1}^{n_c} m_c^{PPD} (PPD \times P_c) + E$$

Thus, measurements from bulk tissue  $Y^{bulk}$  are regressed on  $c = 1$  to  $n_c$  cell type proportions  $P_c$ , and the product of disease status for PPD coded as 0 or 1 by cell type proportions ( $PPD \times P_c$ ). The model allows for covariates (not shown) and residual effects  $E$ . Coefficient  $m_c$  is the effect of cell type  $c$ . The case-control difference  $m_c^{PPD}$  for cell type  $c$  is used to test the null hypothesis that cell type means are equal for cases and controls. Note that the model has no constant due to  $\sum_{c=1}^{n_c} P_c \cong 1$ . Alternatively, the model is sometimes written with a constant whereby one of the cell type proportions is omitted [17] but this produces identical results [16, 28, 29].



## Pathway analyses

Pathway analyses were performed with Reactome [104] and GO [105, 106] databases. These analyses also used circular permutations (For a detailed description see Circular Permutations section of Supplemental Methods) that properly control the Type I error in the presence of correlated sites. Furthermore, as the permutations are performed on a marker level they also properly account for gene size, as larger genes with more markers are more likely to be among the top results in the permutations. Specifically, we first mapped each marker to genes (Ensembl gene annotations GRCh37, release 92: <ftp://ftp.ensembl.org/pub/grch37/release-92/>) [26] using the Bioconductor GRanges package. Gene boundaries were extended to include a 10 kb upstream flank (i.e., promoters). Markers were allowed to map to multiple genes if their genomic position overlapped multiple unique gene annotations. After mapping, we performed 10,000 circular permutations at the marker level. For each permutation, a two-by-two table was created by cross classifying whether or not the genes were among the top findings for each analysis (TWAS) versus whether or not the gene was in the tested pathway. Each gene was counted only once when creating this table (thus, if there were three markers in the gene, this was counted as one and not as three). Cramér's  $V$  (sometimes referred to as Cramér's  $\phi$ ) was used as the test statistic to measure whether genes from the pathway were overrepresented among the top analysis results.  $P$  values were calculated as the proportion of permutations that yielded a value equal to or greater than that of Cramér's  $V$  observed in the empirical data. A minimum of three input genes were required to be present in each queried pathway and were considered significant after controlling the family-wide error rate at  $\alpha = 0.05$ . Pathway tests were run on all markers with a  $q$ -value  $< 0.1$ . As many pathways share a large number of common gene members, we used the Louvain method [107] in the igraph R package to cluster enriched pathways by similarity.

## Regulation Analyses

To study regulatory effects of SNPs, we perform mediation analyses using the model in Fig. S4 (covariates not shown) assuming that the SNP regulates (path  $a$ ) transcription (a mediator), which in turn alters (path  $b$ ) PPD risk. As there may be additional mechanisms through which the marker may affect PPD, the model also allows for a direct effect of the SNP (path  $c$ ) on PPD. The null hypothesis states that the marker does not regulate differentially expressed transcripts ( $H_0: a \times b = 0$ ).

The model in Fig. S4 is the standard approach underlying Mendelian Randomization [108] that aims to improve causal inferences in correlational studies. In this model, the SNP is the instrumental variable where the direction of effect has to be from SNP to transcript abundance (path  $a$ ) and SNP to PPD status (path  $c$ ). Thus, the causal direction cannot be reversed as neither gene expression nor PPD status can change the SNP. The direction of effect from transcript abundance to PPD status (path  $b$ ) can in principle be reversed. However, if we reverse the direction of effect there can no longer be an indirect effect of the SNP on PPD so that in these instances the null hypothesis,  $H_0: a \times b = 0$ , is true. Therefore, rejecting the null hypothesis in the SNP regulatory analyses essentially provides evidence for the causal direction of effects displayed in Fig. S4 where transcriptional changes alter PPD risk and not the other way around.

We also used the model depicted in Fig. S4 to study possible regulatory effects of hormones and DNA methylation. However, as both the direction of path  $a$  and path  $c$  can be reversed it no longer provides evidence for the causal direction of effects so strictly speaking significant findings cannot be interpreted as meaning that hormones and DNA methylation are regulators.

For each gene with a differentially expressed transcript, all annotated transcripts with expression data were tested. SNPs and CpGs were tested as putative regulators if they had a nominal association with case status ( $p < 0.05$ ) in a 10 kb window of the genes tested. Hormones were tested as putative regulators if they had a nominal association with case status ( $p < 0.05$ ). This pre-selection avoids running regulatory analyses with a large number of markers (SNPs, CpGs, hormones), the majority of which cannot be regulators because they do not affect transcript abundance. All raw data had their respective covariates regressed out prior to mediation testing. Causal mediation analyses were conducted with the mediate package (v4.5.0) in R. Specifically, we used the mediate function which implements a quasi-Bayesian approach with 1000 to 1,000,000 Monte Carlo draws for the approximation of the  $p$ -values for the mediation effect [109]. All analyses begin with 1000 simulations. If a  $p$ -value cannot be approximated ( $p$ -value = 0), another round is performed using 10,000 simulations. This process continues with a 10-fold increment in simulations

until a  $p$ -value can be approximated, or the number of simulations reaches 1,000,000. A  $q$ -value  $< 0.1$  was used to declare significance [30].

## DATA AVAILABILITY

Full transcriptomic, DNA methylation, and SNP data used to support the findings of this study have, or will be, deposited in dbGaP (accession: phs002103.v1.p1). [[https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study\\_id=phs002103.v1.p1](https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs002103.v1.p1)]. but will be embargoed until results from full datasets are in press.

## REFERENCES

- Wisner KL, Moses-Kolko EL, Sit DK. Postpartum depression: A disorder in search of a definition. *Arch Women's Ment Health*. 2010;13:37–40.
- Marmorstein NR, Malone SM, Iacono WG. Psychiatric disorders among offspring of depressed mothers: associations with paternal psychopathology. *Am J Psychiatry*. 2004;161:1588–94.
- Flynn HA, Davis M, Marcus SM, Cunningham R, Blow FC. Rates of maternal depression in pediatric emergency department and relationship to child service utilization. *Gen Hosp Psychiatry*. 2004;26:316–22.
- Hamilton BE, Martin JA, Osterman MJ, Curtin SC, Matthews TJ. Births: Final data for 2014. *Natl Vital-Stat Rep*. 2015;64:1–64.
- Gavin NI, Gaynes BN, Lohr KN, Meltzer-Brody S, Gartlehner G, Swinson T. Perinatal depression: A systematic review of prevalence and incidence. *Obstet Gynecol*. 2005;106:1071–83.
- Gaynes BN, Gavin N, Meltzer-Brody S, Lohr KN, Swinson T, Gartlehner G, et al. Perinatal depression: Prevalence, screening accuracy, and screening outcomes. *Evid Rep/Technol Assess*. 2005;119:1–8.
- Sloman J, Slomov G, Emonts P, Reginster JY, Bruyere O. Consequences of maternal postpartum depression: A systematic review of maternal and infant outcomes. *Women's Health (Lond)*. 2019;15:1745506519844044.
- Gelaye B, Rondon MB, Araya R, Williams MA. Epidemiology of maternal depression, risk factors, and child outcomes in low-income and middle-income countries. *Lancet Psychiatry*. 2016;3:973–82.
- Farias-Antunez S, Xavier MO, Santos IS. Effect of maternal postpartum depression on offspring's growth. *J Affect Disord*. 2018;228:143–52.
- Netsi E, Pearson RM, Murray L, Cooper P, Craske MG, Stein A. Association of persistent and severe postnatal depression with child outcomes. *JAMA Psychiatry*. 2018;75:247–53.
- Oh Y, Joung YS, Baek JH, Yoo N. Maternal depression trajectories and child executive function over 9 years. *J Affect Disord*. 2020;276:646–52.
- Shen-Orr SS, Gaujoux R. Computational deconvolution: Extracting cell type-specific information from heterogeneous samples. *Curr Opin Immunol*. 2013;25:571–8.
- Aran D, Hu Z, Butte AJ. xCell: Digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol*. 2017;18:220.
- Koestler DC, Christensen B, Karagas MR, Marsit CJ, Langevin SM, Kelsey KT, et al. Blood-based profiles of DNA methylation predict the underlying distribution of cell types: a validation analysis. *Epigenetics*. 2013;8:816–26.
- Guintivano J, Aryee MJ, Kaminsky ZA. A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. *Epigenetics*. 2013;8:290–302.
- Zheng SC, Breeze CE, Beck S, Teschendorff AE. Identification of differentially methylated cell types in epigenome-wide association studies. *Nat Methods*. 2018;15:1059–66.
- Montano CM, Irizarry RA, Kaufmann WE, Talbot K, Gur RE, Feinberg AP, et al. Measuring cell-type specific differential methylation in human brain tissue. *Genome Biol*. 2013;14:R94.
- Kim-Hellmuth S, Aguet F, Oliva M, Munoz-Aguirre M, Kasela S, Wucher V, et al. Cell type-specific genetic regulation of gene expression across human tissues. *Science*. 2020;369:6509.
- Consortium GT. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*. 2020;369:1318–30.
- Harris CJ, Scheibe M, Wongpalee SP, Liu W, Cornett EM, Vaughan RM, et al. A DNA methylation reader complex that enhances gene transcription. *Science*. 2018;362:1182–6.
- Neri F, Rapelli S, Krepelova A, Incarnato D, Parlato C, Basile G, et al. Intragenic DNA methylation prevents spurious transcription initiation. *Nature*. 2017;543:72–7.
- Pan D, Xu Y, Zhang L, Su Q, Chen M, Li B, et al. Gene expression profile in peripheral blood mononuclear cells of postpartum depression patients. *Sci Rep*. 2018;8:10139.
- Landsman A, Aidelman R, Smith Y, Boyko M, Greenberger C. Distinctive gene expression profile in women with history of postpartum depression. *Genomics*. 2017;109:1–8.



24. Mehta D, Grewen K, Pearson B, Wani S, Wallace L, Henders AK, et al. Genome-wide gene expression changes in postpartum depression point towards an altered immune landscape. *Transl Psychiatry*. 2021;11:155.
25. Guintivano J, Sullivan PF, Stuebe AM, Penders T, Thorp J, Rubinow DR, et al. Adverse life events, psychiatric history, and biological predictors of postpartum depression in an ethnically diverse sample of postpartum women. *Psychol Med*. 2018;48:1190–200.
26. Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, et al. Ensemble 2016. *Nucleic Acids Res*. 2016;44:D710–6.
27. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinforma*. 2012;13:86.
28. Chan RF, Turecki G, Shabalín AA, Guintivano J, Zhao M, Xie LY, et al. Cell type-specific methylome-wide association studies implicate neurotrophin and innate immune signaling in major depressive disorder. *Biol Psychiatry*. 2020;87:431–42.
29. Shen-Orr SS, Tibshirani R, Khatri P, Bodian DL, Staedtler F, Perry NM, et al. Cell type-specific gene expression differences in complex tissues. *Nat Methods*. 2010;7:287–9.
30. van den Oord EJ, Sullivan PF. False discoveries and models for gene discovery. *Trends Genet*. 2003;19:537–42.
31. Luo C, Hajkova P, Ecker JR. Dynamic DNA methylation: In the right place at the right time. *Science*. 2018;361:1336–40.
32. Babenko VN, Chadaeva IV, Orlov YL. Genomic landscape of CpG rich elements in human. *BMC Evol Biol*. 2017;17:19.
33. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet*. 2018;50:668–81.
34. Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat Neurosci*. 2019;22:343–52.
35. Di Florio A, Meltzer-Brody S. Is postpartum depression a distinct disorder? *Curr Psychiatry Rep*. 2015;17:76.
36. Batt MM, Duffy KA, Novick AM, Metcalf CA, Epperson CN. Is postpartum depression different from depression occurring outside of the perinatal period? A review of the evidence. *Focus (Am Psychiatr Publ)*. 2020;18:106–19.
37. Bernstein IH, Rush AJ, Yonkers K, Carmody TJ, Woo A, McConnell K, et al. Symptom features of postpartum depression: are they distinct? *Depress Anxiety*. 2008;25:20–6.
38. Pawluski JL, Lonstein JS, Fleming AS. The neurobiology of postpartum anxiety and depression. *Trends Neurosci*. 2017;40:106–20.
39. Schlaaff K, Dobrowolny H, Frodl T, Mawrin C, Gos T, Steiner J, et al. Increased densities of T and B lymphocytes indicate neuroinflammation in subgroups of schizophrenia and mood disorder patients. *Brain Behav Immun*. 2020;88:497–506.
40. Syed SA, Beurel E, Loewenstein DA, Lowell JA, Craighead WE, Dunlop BW, et al. Defective inflammatory pathways in never-treated depressed patients are associated with poor treatment response. *Neuron*. 2018;99:914–24 e913.
41. Ahmetshpach D, Schwarte K, Ambree O, Burger C, Falcone V, Seiler K, et al. Altered B cell homeostasis in patients with major depressive disorder and normalization of CD5 surface expression on regulatory B cells in treatment responders. *J Neuroimmune Pharm*. 2018;13:90–9.
42. Hamer JA, Testani D, Mansur RB, Lee Y, Subramaniapillai M, McIntyre RS. Brain insulin resistance: A treatment target for cognitive impairment and anhedonia in depression. *Exp Neurol*. 2019;315:1–8.
43. Kan C, Silva N, Golden SH, Rajala U, Timonen M, Stahl D, et al. A systematic review and meta-analysis of the association between depression and insulin resistance. *Diabetes Care*. 2013;36:480–9.
44. Frangou S, Shirali M, Adams MJ, Howard DM, Gibson J, Hall LS, et al. Insulin resistance: Genetic associations with depression and cognition in population based cohorts. *Exp Neurol*. 2019;316:20–6.
45. Muzzio D, Zenclussen AC, Jensen F. The role of B cells in pregnancy: the good and the bad. *Am J Reprod Immunol*. 2013;69:408–12.
46. Dutta S, Sengupta P, Haque N. Reproductive immunomodulatory functions of B cells in pregnancy. *Int Rev Immunol*. 2020;39:53–66.
47. Lima J, Martins C, Leandro MJ, Nunes G, Sousa MJ, Branco JC, et al. Characterization of B cells in healthy pregnant women from late pregnancy to post-partum: A prospective observational study. *BMC Pregnancy Childbirth*. 2016;16:139.
48. Brann E, Fransson E, White RA, Papadopoulos FC, Edvinsson A, Kamali-Moghaddam M, et al. Inflammatory markers in women with postpartum depressive symptoms. *J Neurosci Res*. 2020;98:1309–21.
49. Brann E, Papadopoulos F, Fransson E, White R, Edvinsson A, Hellgren C, et al. Inflammatory markers in late pregnancy in association with postpartum depression—A nested case-control study. *Psychoneuroendocrinology*. 2017;79:146–59.
50. Leff-Gelman P, Mancilla-Herrera I, Flores-Ramos M, Cruz-Fuentes C, Reyes-Grajeda JP, Garcia-Cuetara Mdel P, et al. The immune system and the role of inflammation in perinatal depression. *Neurosci Bull*. 2016;32:398–420.
51. Rawlings DJ, Metzler G, Wray-Dutra M, Jackson SW. Altered B cell signalling in autoimmunity. *Nat Rev Immunol*. 2017;17:421–36.
52. Andersson NW, Gustafsson LN, Okkels N, Taha F, Cole SW, Munk-Jorgensen P, et al. Depression and the risk of autoimmune disease: A nationally representative, prospective longitudinal study. *Psychol Med*. 2015;45:3559–69.
53. Siegmann EM, Muller HHO, Luecke C, Philipsen A, Kornhuber J, Groher TW. Association of depression and anxiety disorders with autoimmune thyroiditis: A systematic review and meta-analysis. *JAMA Psychiatry*. 2018;75:577–84.
54. Euesden J, Danese A, Lewis CM, Maughan B. A bidirectional relationship between depression and the autoimmune disorders - New perspectives from the National Child Development Study. *PLoS One*. 2017;12:e0173015.
55. Carter RH. B cells in health and disease. *Mayo Clin Proc*. 2006;81:377–84.
56. Waters LR, Ahsan FM, Wolf DM, Shirihai O, Teitell MA. Initial B cell activation induces metabolic reprogramming and mitochondrial remodeling. *Science*. 2018;5:99–109.
57. Tolar P. Cytoskeletal control of B cell responses to antigens. *Nat Rev Immunol*. 2017;17:621–34.
58. Defrance T, Casamayor-Palleja M, Krammer PH. The life and death of a B cell. *Adv Cancer Res*. 2002;86:195–225.
59. Buchanan TA, Metzger BE, Freinkel N, Bergman RN. Insulin sensitivity and B-cell responsiveness to glucose during late pregnancy in lean and moderately obese women with normal glucose tolerance or mild gestational diabetes. *Am J Obstet Gynecol*. 1990;162:1008–14.
60. Butte NF. Carbohydrate and lipid metabolism in pregnancy: Normal compared with gestational diabetes mellitus. *Am J Clin Nutr*. 2000;71:1256S–61S.
61. Sivan E, Homko CJ, Chen X, Reece EA, Boden G. Effect of insulin on fat metabolism during and after normal pregnancy. *Diabetes*. 1999;48:834–8.
62. Flores DL, Hendrick VC. Etiology and treatment of postpartum depression. *Curr Psychiatry Rep*. 2002;4:461–6.
63. Di Cianni G, Miccoli R, Volpe L, Lencioni C, Del Prato S. Intermediate metabolism in normal pregnancy and in gestational diabetes. *Diabetes Metab Res Rev*. 2003;19:259–70.
64. Chen TH, Lan TH, Yang CY, Juang KD. Postpartum mood disorders may be related to a decreased insulin level after delivery. *Med Hypotheses*. 2006;66:820–3.
65. Kleinridders A, Cai W, Cappellucci L, Ghazarian A, Collins WR, Vienberg SG, et al. Insulin resistance in brain alters dopamine turnover and causes behavioral disorders. *Proc Natl Acad Sci USA*. 2015;112:3463–8.
66. Penckofer S, Quinn L, Byrn M, Ferrans C, Miller M, Strange P. Does glycemic variability impact mood and quality of life? *Diabetes Technol Ther*. 2012;14:303–10.
67. Lustman PJ, Anderson RJ, Freedland KE, de Groot M, Carney RM, Clouse RE. Depression and poor glycemic control: a meta-analytic review of the literature. *Diabetes Care*. 2000;23:934–42.
68. Lin KW, Wroolie TE, Robakis T, Rasgon NL. Adjuvant pioglitazone for unremitted depression: Clinical correlates of treatment response. *Psychiatry Res*. 2015;230:846–52.
69. Colle R, de Laminat D, Rotenberg S, Hozer F, Hardy P, Verstuylt C, et al. PPAR-gamma agonists for the treatment of major depression: A review. *Pharmacopsychiatry*. 2017;50:49–55.
70. Kim C. Maternal outcomes and follow-up after gestational diabetes mellitus. *Diabet Med*. 2014;31:292–301.
71. Kozhimannil KB, Pereira MA, Harlow BL. Association between diabetes and perinatal depression among low-income mothers. *JAMA*. 2009;301:842–7.
72. Miller ES, Peri MR, Gossett DR. The association between diabetes and postpartum depression. *Arch Women's Ment Health*. 2016;19:183–6.
73. Ruohomaki A, Toffol E, Upadhyaya S, Keski-Nisula L, Pekkanen J, Lampi J, et al. The association between gestational diabetes mellitus and postpartum depressive symptomatology: A prospective cohort study. *J Affect Disord*. 2018;241:263–8.
74. Sliker RC, van der Heijden A, van Leeuwen N, Mei H, Nijpels G, Beulens JWJ, et al. HbA1c is associated with altered expression in blood of cell cycle- and immune response-related genes. *Diabetologia*. 2018;61:138–46.
75. Razny U, Polus A, Goralska J, Zdzienicka A, Gruca A, Kapusta M, et al. Effect of insulin resistance on whole blood mRNA and microRNA expression affecting bone turnover. *Eur J Endocrinol*. 2019;181:525–37.
76. Umeno A, Yoshida Y. Utility of hemoglobin A1c in detecting risk of type 2 diabetes: comparison of hemoglobin A1c with other biomarkers. *J Clin Biochem Nutr*. 2019;65:59–64.
77. Winer DA, Winer S, Shen L, Chng MH, Engleman EG. B lymphocytes as emerging mediators of insulin resistance. *Int J Obes Suppl*. 2012;2:54–57.
78. Winer DA, Winer S, Shen L, Wadia PP, Yantha J, Paltser G, et al. B cells promote insulin resistance through modulation of T cells and production of pathogenic IgG antibodies. *Nat Med*. 2011;17:610–7.
79. Virkamaki A, Ueki K, Kahn CR. Protein-protein interaction in insulin signaling and the molecular mechanisms of insulin resistance. *J Clin Invest*. 1999;103:931–43.

80. Manning BD, Toker A. AKT/PKB signaling: Navigating the network. *Cell*. 2017;169:381–405.
81. Romeo GR, Jain M. Purine metabolite signatures and type 2 Diabetes: Innocent bystanders or actionable items? *Curr Diab Rep*. 2020;20:30.
82. Weber G, Lui MS, Jayaram HN, Pillwein K, Natsumeda Y, Faderan MA, et al. Regulation of purine and pyrimidine metabolism by insulin and by resistance to tiazofurin. *Adv Enzym Regul*. 1985;23:81–99.
83. Yoon MS. The emerging role of branched-chain amino acids in insulin resistance and metabolism. *Nutrients* 2016;8:405.
84. Sears B, Perry M. The role of fatty acids in insulin resistance. *Lipids Health Dis*. 2015;14:121.
85. Mingrone G, Castagneto-Gissey L, Mace K. Use of dicarboxylic acids in type 2 diabetes. *Br J Clin Pharm*. 2013;75:671–6.
86. Fournier AM. Intracellular starvation in the insulin resistance syndrome and type II diabetes mellitus. *Med Hypotheses*. 1998;51:95–9.
87. Abiola M, Favier M, Christodoulou-Vafeiadou E, Pichard AL, Martelly I, Guillet-Deniau I. Activation of Wnt/beta-catenin signaling increases insulin sensitivity through a reciprocal regulation of Wnt10b and SREBP-1c in skeletal muscle cells. *PLoS One*. 2009;4:e8509.
88. Daneman R, Prat A. The blood-brain barrier. *Cold Spring Harb Perspect Biol*. 2015;7:a020412.
89. Anzalone AV, Koblan LW, Liu DR. Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat Biotechnol*. 2020;38:824–44.
90. Guintivano J, Sullivan PF, Stuebe AM, Penders T, Thorp J, Rubinow DR, et al. Adverse life events, psychiatric history, and biological predictors of postpartum depression in an ethnically diverse sample of postpartum women. *Psychol Med*. 2018;48:1190–1200.
91. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry*. 1987;150:782–6.
92. Otsubo T, Tanaka K, Koda R, Shinoda J, Sano N, Tanaka S, et al. Reliability and validity of Japanese version of the Mini-International Neuropsychiatric Interview. *Psychiatry Clin Neurosci*. 2005;59:517–26.
93. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 1998;59:22–33.
94. Sahraeian SME, Mohiyuddin M, Sebra R, Tilgner H, Afshar PT, Au KF, et al. Gaining comprehensive biological insight into the transcriptome by performing a broad-spectrum RNA-seq analysis. *Nat Commun*. 2017;8:59.
95. Guintivano J, Shabalín AA, Chan RF, Rubinow DR, Sullivan PF, Meltzer-Brody S, et al. Test-statistic inflation in methylome-wide association studies. *Epigenetics*. 2020;15:1163–6.
96. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*. 2013;8:203–9.
97. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics*. 2013;29:189–96.
98. Shabalín AA, Hattab MW, Clark SL, Chan RF, Kumar G, Aberg KA, et al. RaMWAS: Fast methylome-wide association study pipeline for enrichment platforms. *Bioinformatics*. 2018;34:2283–5.
99. International Schizophrenia C, Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. 2009;460:748–52.
100. Wang KS, Liu XF, Aragam N. A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. *Schizophr Res*. 2010;124:192–9.
101. Loh PR, Danecek P, Palamara PF, Fuchsberger C, AR Y, KF H, et al. Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet*. 2016;48:1443–8.
102. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016;48:1284–7.
103. Venet D, Pecasse F, Maenhaut C, Bersini H. Separation of samples into their constituents using gene expression data. *Bioinformatics*. 2001;17:5279–87.
104. Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R, et al. The Reactome pathway Knowledgebase. *Nucleic Acids Res*. 2016;44:D481–7.
105. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25:25–9.
106. The Gene Ontology C. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res*. 2019;47:D330–8.
107. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J Stat Mech-Theory Exp*. 2008;12:P10008.
108. Davey Smith G, Ebrahim S. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? *BMJ*. 2005;330:1076–9.
109. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychological Methods*. 2010;15:309–34.

## AUTHOR CONTRIBUTIONS

EVDO, JG, and KA conceived the idea, designed and supervised the implementation of the study, interpreted the results, revised the paper. JG conducted the analyses, interpreted the results, drafted, and revised the paper. SC contributed to the analyses. PFS, DR, SMB contributed to the study design and interpretation of the results. All authors discussed and commented on the paper.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41380-022-01525-7>.

**Correspondence** and requests for materials should be addressed to Jerry Guintivano.

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

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022