OBSTETRICS

Discrete placental gene expression signatures accompany diabetic disease classifications during pregnancy

Enrico R. Barrozo, PhD; Diana A. Racusin, MD; Michael D. Jochum, PhD; Brandon T. Garcia, BS; Melissa A. Suter, PhD; Melanie Delbeccaro, MD; Cynthia Shope, MS; Kathleen Antony, MD; Kjersti M. Aagaard, MD, PhD

BACKGROUND: Gestational diabetes mellitus affects up to 10% of pregnancies and is classified into subtypes gestational diabetes subtype A1 (GDMA1) (managed by lifestyle modifications) and gestational diabetes subtype A2 (GDMA2) (requiring medication). However, whether these subtypes are distinct clinical entities or more reflective of an extended spectrum of normal pregnancy endocrine physiology remains unclear.

OBJECTIVE: Integrated bulk RNA-sequencing (RNA-seq), single-cell RNA-sequencing (scRNA-seq), and spatial transcriptomics harbors the potential to reveal disease gene signatures in subsets of cells and tissue microenvironments. We aimed to combine these high-resolution technologies with rigorous classification of diabetes subtypes in pregnancy. We hypothesized that differences between preexisting type 2 and gestational diabetes subtypes would be associated with altered gene expression profiles in specific placental cell populations.

STUDY DESIGN: In a large case-cohort design, we compared validated cases of GDMA1, GDMA2, and type 2 diabetes mellitus (T2DM) to healthy controls by bulk RNA-seq (n=54). Quantitative analyses with reverse transcription and quantitative PCR of presumptive genes of significant interest were undertaken in an independent and nonoverlapping validation cohort of similarly well-characterized cases and controls (n=122). Additional integrated analyses of term placental single-cell, single-nuclei, and spatial transcriptomics data enabled us to determine the cellular subpopulations and niches that aligned with the GDMA1, GDMA2, and T2DM gene expression signatures at higher resolution and with greater confidence.

RESULTS: Dimensional reduction of the bulk RNA-seq data revealed that the most common source of placental gene expression variation was the diabetic disease subtype. Relative to controls, we found 2052 unique and significantly differentially expressed genes ($-2<\text{Log}_2$[fold-change]>2 thresholds; $q<0.05$ Wald Test) among GDMA1 placental specimens, 267 among GDMA2, and 1520 among T2DM. Several candidate marker genes (chorionic somatomammotropin hormone 1 [*CSH1*], period circadian regulator 1 [*PER1*], phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit beta [*PIK3CB*], forkhead box O1 [*FOXO1*], epidermal growth factor receptor [*EGFR*], interleukin 2 receptor subunit beta [*IL2RB*], superoxide dismutase 3 [*SOD3*], dedicator of cytokinesis 5 [*DOCK5*], suppressor of glucose, and autophagy associated 1 [*SOGA1*]) were validated in an independent and nonoverlapping validation cohort ($q<0.05$ Tukey). Functional enrichment revealed the pathways and genes most impacted for each diabetes subtype, and the degree of proximal similarity to other subclassifications. Surprisingly, GDMA1 and T2DM placental signatures were more alike by virtue of increased expression of chromatin remodeling and epigenetic regulation genes, while albumin was the top marker for GDMA2 with increased expression of placental genes in the wound healing pathway. Assessment of these gene signatures in single-cell, single-nuclei, and spatial transcriptomics data revealed high specificity and variability by placental cell and microarchitecture types. For example, at the cellular and spatial (eg, microarchitectural) levels, distinguishing features were observed in extravillous trophoblasts (GDMA1) and macrophages (GDMA2). Lastly, we utilized these data to train and evaluate 4 machine learning models to estimate our confidence in predicting the control or diabetes status of placental transcriptome specimens with no available clinical metadata.

CONCLUSION: Consistent with the distinct association of perinatal outcome risk, placentae from GDMA1, GDMA2, and T2DM-affected pregnancies harbor unique gene signatures that can be further distinguished by altered placental cellular subtypes and microarchitectural niches.

Key words: diabetes, GDM, GDMA, RNA-seq, scRNA-seq, T2DM, transcriptomics, Visium

Introduction

Up to 10% of pregnant persons in the United States are diagnosed with gestational diabetes mellitus (GDM) annually, with a global prevalence approximating 17%.[1,2] Pragmatic distinctions between GDM requiring medications (gestational diabetes subtype A2 [GDMA2]), GDM managed by lifestyle and not medication (gestational diabetes subtype A1 [GDMA1]), and preexisting type 2 diabetes mellitus (T2DM) can be challenging.[3–5] Regardless, duration and severity of hyperglycemia in pregnancy drive short- and long-term perinatal, maternal, and offspring outcomes.[6,7]

Historically, GDM was described as any manifestation of glucose intolerance discovered during pregnancy.[5] However, this definition is imprecise because it fails to recognize well-known differences in underlying disease pathophysiology. Outside of pregnancy, beyond clinical parameters such as age of onset and severity of hyperglycemia, development of clinical laboratory testing reflective of disease pathophysiology has enabled timely and accurate differential diagnosis of type 1 (immune-mediated T1DM, driven by autoimmune destruction of the pancreatic islet cells as detected by diagnostic anti-islet autoantibodies to insulin, glutamic acid

Cite this article as: Barrozoa ER, Racusina DA, Jochum MD, et al. Discrete placental gene expression signatures accompany diabetic disease classifications during pregnancy. Am J Obstet Gynecol 2024;XXX:XX–XX.

0002-9378

https://doi.org/10.1016/j.ajog.2024.05.014

Click Supplemental Materials and Video under article title in Contents at ajog.org

AJOG at a Glance

Why was this study conducted?

The underlying placental gene signatures associated with preexisting type 2 diabetes mellitus (T2DM) and gestational diabetes subtype A1 and gestational diabetes subtype A2 (GDMA1 and GDMA2) remain poorly understood.

Key findings

Discrete placental gene expression profiles distinguishing GDMA1, GDMA2, T2DM, and healthy controls were identified and validated in complementary cohorts. Interestingly, gene signature profiles from GDMA1 and T2DM were more alike than GDMA1 and GDMA2, or GDMA2 and T2DM.

What does this add to what is known?

These findings suggest that placental gene expression profiles cannot be solely attributed to diabetic medication usage, but are more related to the diabetes subtype. While results from term placentae are inherently not diagnostic, these findings may inspire future research to identify sensitive and specific biomarkers that could guide earlier initiation and optimization of therapy.

decarboxylase, tyrosine phosphatase-like protein tyrosine phosphatase-like protein islet cell antigen 512, and zinc transporter 8) and type 2 (T2DM, peripheral insulin resistance and relative impairment of insulin secretion) diabetes.[5] In contrast, while GDM is also a disorder of insulin resistance, one key driving force behind the development of GDM is thought to arise from an inability to fully adapt or accommodate physiological insulin resistance, primarily driven by human placental lactogen (hPL).[3,4,6]

Distinguishing between T2DM and GDM poses a significant challenge, as both conditions are characterized by peripheral insulin resistance and a relative deficiency in insulin secretion, and the challenge can become further compounded by unequal access to healthcare and racial health disparities. Amongst patients with access to pre- and early-pregnancy care, diabetes established before pregnancy or identified in the first trimester or early midgestation by either elevated fasting glucose or increased hemoglobin A1C (HbA1C) is considered preexisting diabetes.[5] Conversely, GDM is diagnosed between weeks 24 and 28 of pregnancy by oral glucose tolerance test (GTT) and is considered a complication of later pregnancy.[5] GDM can also be subcategorized into either diet-controlled (GDMA1) or medication-controlled (GDMA2). Currently, the most accurate way to distinguish T2DM from GDM in affected gravidae requires postpartum GTT. While advantageous to subsequent maternal or future pregnancy care, postpartum testing is of no value to the index-affected pregnancy, is reliant upon access to care outside of pregnancy, and is disparate amongst populations with the highest prevalence of diabetic risk factors.[8–10]

The placenta is presumed to modulate the physiological development of insulin resistance during pregnancy via the secretion of steroid hormones, and hPL in particular.[3,4,6] Thus, several studies have examined the changes to placental gene expression through transcriptomic and proteomic analysis with noted heterogeneity of findings.[11–25] We hypothesized that since gestational diabetes is a distinct clinical entity, we should be able to detect unique molecular signature differences in placental gene expression between subjects with GDMA1, GDMA2, preexisting diabetes, and nondiabetic controls. To this end, we applied bulk RNA-sequencing (RNA-seq) and robust computational analysis to a discovery case-control cohort with subsequent validation in a separate and nonoverlapping validation cohort using reverse transcription and quantitative polymerase chain reaction (RT-qPCR) to identify differential gene expression specific for GDMA1, GDMA2, and T2DM. We then integrated single-cell RNA-sequencing (scRNA-seq) and spatial transcriptomics to assess these gene expression signatures in the context of placental cell subtypes and their microenvironments, which compartmentalize diverse functions at the maternal-fetal interface. We aimed to combine these high-resolution technologies with a rigorous classification of gene expression alterations associated with diabetes subtypes in pregnancy, including T2DM and nondiabetic controls.

Materials and methods

Participants and clinical disease classifications

Protocols utilizing human tissue were approved by the Baylor College of Medicine Institutional Review Board (H-28623). An extensive medical record review screened participants for inclusion and exclusion criteria, and informed consent was obtained before deidentification and utilization in the current study. Inclusion and exclusion criteria were set a priori as follows: For participants to be eligible, they must have been pregnant, at least 18 years of age, and able to provide informed consent to provide a placental sample at delivery. Clinical metadata were collected from the electronic medical record and curated in a secured database by trained research personnel are reported in Table and Supplemental Table, A. Selected cases were routinely audited and adjudicated by board-certified or board-eligible maternal-fetal medicine specialists (K.M.A., D.R., and K.A.) to ensure diagnostic accuracy. Diagnosis of gestational diabetes was based on uniform established institutional criteria using the Carpenter-Coustan GTT, with a screening glucose challenge test value of greater than or equal to 140 mg/dL defining a positive screen. Pregestational diabetes was classified by 1 or more of the following: established diagnosis of diabetes before pregnancy, positive GTT, or elevated HbA1C early in pregnancy (HbA1C>6.4). Hypertensive disorders of pregnancy, including both gestational hypertension and preeclampsia, were defined using the American College

TABLE
Study population characteristics (n=176 total)

Variable	Classification group				Adjusted P values		
	Control: 11 RNA-seq & 29 RT-qPCR, n=40 total	GDMA1: 5 RNA-seq & 25 RT-qPCR, n=30 total	GDMA2: 9 RNA-seq & 35 RT-qPCR, n=44 total	T2DM: 5 RNA-seq & 33 RT-qPCR, n=38 total	Control vs GDMA1	Control vs GDMA2	Control vs T2DM
Maternal age, mean years (SEM)	28.5 (1.0)	32.4 (1.0)	31.6 (0.5)	31.8 (1.0)	.279	.391	.410
Gravidity, mean (SEM)	3.6 (0.3)	3.0 (0.3)	3.3 (0.2)	3.3 (0.3)	>.99	>.99	>.99
Parity, mean (SEM)	2.3 (0.3)	1.8 (0.3)	2.1 (0.2)	2.1 (0.2)	.267	.206	1.000
Gestational age, mean weeks (SEM)	39.7 (0.2)	39.1 (0.2)	38.5 (0.2)	38.2 (0.2)	.319	**.0012**	**<.0001**
Birthweight, mean grams (SEM)	3504 (80.1)	3219 (84.1)	3465 (86.9)	3499 (140)	>.99	>.99	.862
Hypertensive disorders of pregnancy, n= positive/ total (%)	6/40 (15)	7/30 (23.3)	6/44 (13.6)	21/38 (55.2)	>.99	>.99	**.0002**
Body mass index at delivery, mean (SEM)	32.3 (1.6)	29.1 (1.0)	35.3 (1.7)	34.8 (1.3)	.269	.126	.099
Race (%)					>.99	>.99	>.99
White	37 (92.5)	26 (86.7)	42 (95.4)	37 (97.4)			
Black	2 (5)	0 (0)	1 (2.3)	1 (2.6)			
Asian	1 (2.5)	1 (3.3)	0 (0)	0 (0)			
Other	0 (0)	3 (10)	1 (2.3)	0 (0)			
Ethnicity (%)					.631	>.99	>.99
Hispanic/LatinX	38 (95)	30 (100)	43 (97.7)	37 (97.4)			
Non-Hispanic/LatinX	2 (5)	0 (0)	1 (3.3)	1 (2.6)			
Fetal sex (%)					>.99	>.99	>.99
Female	14 (35)	11 (37.7)	21 (47.7)	21 (55.2)			
Male	26 (65)	19 (63.3)	20 (45.5)	17 (44.8)			
Not reported	0 (0)	0 (0)	3 (6.8)	0 (0)			
Diabetes medication (%)					>.99	>.99	>.99
None	40 (100)	30 (100)	0 (0)	0 (0)			
Glyburide	0 (0)	0 (0)	39 (88.6)	9 (22.7)			
Insulin	0 (0)	0 (0)	5 (11.4)	29 (76.3)			
Prenatal 3 h GTT					>.99	>.99	>.99
Pass	40 (100)	0 (0)	10 (22.7)	0 (0)			
Fail	0 (0)	30 (100)	34 (77.3)	0 (0)			
NA	0 (0)	0 (0)	0 (0)	38 (100)			
Postpartum 2 h GTT					>.99	>.99	>.99
Pass	0 (0)	30 (100)	34 (77.3)	0 (0)			
Fail	0 (0)	0 (0)	0 (0)	0 (0)			
NA	40 (100)	0 (0)	10 (22.7)	38 (100)			

Statistically significant comparisons (adjusted P value<.05) in bold as determined by Kruskal-Wallis (numerical) or Friedman (categorical) test with Dunn's multiple comparisons. There were 24 deidentified subject's specimens with no available clinical metadata, and thus allowed us to categorize them as unknowns in the RNA-seq analysis. In Figures 5 and 6 there were n=2 subjects scRNA-seq controls, n=1 GDMA1, and n=1 GDMA2 subjects from Yang et al (2021) analyzed independently.

GDMA1, gestational diabetes subtype A1; *GDMA2*, gestational diabetes subtype A2; *GTT*, glucose tolerance test; *NA*, not applicable, data not determined or reported; *RNA-seq*, RNA-sequencing; *RT-qPCR*, reverse transcription and quantitative polymerase chain reaction; *scRNA-seq*, single-cell RNA-sequencing; *SEM*, standard error of the mean; *T2DM*, type 2 diabetes mellitus.

of Obstetricians and Gynecologists' classification.[26]

Sample collection and processing

All samples were collected by personnel trained in perinatal and placental pathology under strict uniform protocol. Briefly, following standard obstetrical practice, the placenta was delivered and immediately passed to trained personnel in a sterile clean container. In a separate room, 2 samples were collected from midway between the cord insertion and placental margin by incision through the fetal surface into the parenchyma but not to the maternal surface. All samples were collected within 1 hour of delivery under clean and sterile conditions as detailed above, placed on dry ice in sterile closed vials, transported to the laboratory, and stored at −80°C until messenger RNA (mRNA) extraction.

Bulk RNA-seq

Ribonucleic acid (RNA) was extracted from placental tissue as previously described.[27] The Machery Nagel Nucleospin II kit was used to extract RNA from each sample, and samples were stored at −80°C. Each sample was analyzed for quality control on the Agilent Bioanalyzer, with a minimum RNA Integrity Number of 4.0 accepted for transcriptomic analysis. For each library, mRNA was purified from 10 μg of total RNA using the DynaBeads mRNA Purification Kit (Invitrogen) and fragmented using the RNA Fragmentation Reagents (Ambion). Double-stranded Complementary deoxyribonucleic acid (cDNA) was synthesized from fragmented mRNA using the Superscript Double-Stranded cDNA Synthesis Kit (Invitrogen) and Random Hexamer primers (50 ng/μL, Invitrogen). cDNA libraries were prepared according to the manufacturer's protocol (TruSeq RNA Library Prep Kit v2) and sequenced on an Illumina Hiseq 2000.

Bulk RNA-seq analysis

Reads were preprocessed and quality-filtered using FastqQC (v0.11.9), and barcodes and adapters were trimmed using Trimmomatic (v0.33).[28] Reads were aligned to the human transcriptome (GRCh38.p13) using STAR (v2.7.8).[29] PCR duplicate reads were removed using Picard (v2.24.0).[30] Unique reads were counted using HTseq (v0.11.1).[31] Counts were used for differential expression analysis in R (v4.0.2) using DEseq2 (v3.12)[32] relative to the healthy control placenta samples. Gene Ontology (GO) biological processes functional enrichment analysis utilized clusterProfiler (v4)[33] to identify significantly enriched terms.

Reverse transcription and quantitative PCR

Total RNA was isolated from placental tissue (n=122), purified using the Nucleospin II kit, and total RNA was reverse transcribed in a final volume of 20 μL using the High Capacity cDNA Archive Kit with random primers (Applied Biosystems) as described by the manufacturer. Commercially available primer and probe sets (Applied Biosystems) were used for real-time PCR using 2 μL of cDNA samples (50 ng), 5 μL of TaqMan PCR Universal Master Mix (Applied Biosystems), 2 μM of each primer and 1 μM of gene-specific TaqMan probe in a total volume of 5 μL. The ABI PRISM 7900HT Sequence Detection System (Applied Biosystems) was used to detect the amplification level. Fold-change for each group was calculated using the threshold value method[34] as previously described[35] utilizing glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) as the endogenous control. The following genes were analyzed for RT-qPCR validation using off-the-shelf primers: *PER1*, *SOGA1*, *IL2RB*, *SOD3*, *DOCK5*, *CSH1*, *EGFR*, *PIK3CB*, and *FOXO1*.

Single-cell and spatial transcriptomics

The methods of the term placenta single-cell and spatial transcriptomics atlas have been described previously.[36] For this study, we reexamined the GDMA1, GDMA2, and T2DM gene signatures identified in the bulk RNA-seq data in the atlas regions where the placenta was sampled (decidua, basal plate, chorioamniotic membrane, chorionic villi, or uniformly from the parenchyma), the atlas cell type or spatial niches, and leveraged the atlas control vs disease classifications, which included GDM single-cell cases from Yang et al (2021).[11]

Machine learning random forest algorithms

Gene counts, and metadata were imported into R (v4.3.1) and underwent preprocessing via i) center and scaling of the gene counts, ii) filtering genes with zero or near zero variance, and iii) filtering genes with >90% correlation. After preprocessing, datasets were copied into individual data frames for each group (GDMA1, GDMA2, T2DM, control, and unknown), and annotated as a target column for a 2-component classification model where the target for each group was converted into logical class form. Each data frame was then sliced into 80% training and 20% testing weighted by the target classification outcome. A 5-fold 3 repeat cross-validation random forest classification model was then trained for each dataset. Models for each group were evaluated via model metrics i) accuracy, ii) kappa scores, iii) out-of-the-box (OOB) error rate estimation, and iv) postresample metrics derived from prediction on the testing dataset. Following the training of individual random forest classification models for each target class, Unknown samples underwent prediction with each model, resulting in a data frame with predicted classification and probability scores.

Data and code availability

The bulk RNA-seq data have been deposited to the Gene Expression Omnibus (GEO) (accession GSE249311). The custom scripts used for bulk and single-cell transcriptomics analyses are available at https://github.com/Aagaardlab/GDMA-Subtype-Transcriptomics and https://github.com/Aagaardlab/placenta-spatial-transcriptomics, respectively. Published term human placenta scRNA-seq, snRNA-seq, and spatial transcriptomics datasets were downloaded and analyzed independently including GEO (accession GSE222987),[36] GEO accession (GSE173193),[11] European Genome-Phenome Archive accession

(EGAD00001003705),[37] and database of genotypes and phenotypes accessions (phs001886.v1.p1, phs001886.v2.p1, and phs001886.v3).[38–40]

Statistical analyses

Significance (adjusted P value<.05) of clinical metadata was determined by Kruskal-Wallis (numerical) or Friedman (categorical) test with Dunn's multiple comparisons with Prism (v10.0.3). In differential expression analyses, significance was determined with Wald tests with False Discovery Rate multiple test correction where significance was defined as a $q<0.05$ in R with DEseq2. Differentially expressed genes (DEGs) (significance thresholds defined as $q<0.05$ and $-2<\text{Log}_2[\text{fold-change}]>2$) were subject to GO biological processes functional enrichment analysis utilizing clusterProfiler (v4) to identify significantly enriched terms (adjusted P value Bonferroni correction, $q<0.05$).

Results

Gene expression analysis of placentae from participants with preexisting or clinically characterized gestational diabetes

This case-cohort study was designed to rigorously examine placental gene expression differences between gestational and pregestational diabetics. The only significant differences among the cases relative to controls included (1) gestational age at delivery: controls, group mean of 39.7 weeks (±0.2), GDMA2 38.5 weeks (±0.2; adjusted P value=.0012; Kruskal-Wallis), and T2DM 38.2 weeks (±0.2 adjusted P value<.0001) and (2) hypertensive disorders of pregnancy where 55% prevalence among gravidae with T2DM group was significantly greater (non-T2DM range 13%–23%, adjusted P value=.0002) (Table). Additional clinical metadata are available in Supplemental Table, A. Consistent with our patient population and those at highest risk for GDM and T2DM, the majority of our participants identified as Hispanic and were multiparous. Consistent with participants clinical documentation demonstrating adequate glycemic control during pregnancy (eg, fasting glucose values <95 mg/dL; 2 hours postprandial <120 mg/dL), there was no statistical difference in birth weight in our participant cohort.

To determine the GDMA1, GDMA2, and T2DM gene expression changes at the placental maternal-fetal interface, we used a 2-phase case-cohort study design in which an initial discovery cohort (phase 1) of $n=30$ of 54 participants samples were comprised of well-characterized specimens ($n=11$ controls, 5 GDMA1, 9 GDMA2, 5 T2DM) which were compared to $n=24$ placentae from an uncharacterized population-based controls ("unknown") for analysis by bulk RNA-seq (Figure 1, A). With dimensional reduction, we found that our participants' placental gene expression signatures independently parsed by diabetes classification (Figure 1, A and B). Moreover, differential expression analysis comparing each diabetes classification group to the controls (eg, control vs GDMA1, control vs GDMA2, control vs T2DM) allowed for the discovery of 8749 unique significantly DEGs for each diabetes subtype relative to the controls (significance thresholds defined as $-2<\text{Log}_2[\text{fold-change}]>2$; adjusted P value (q)<.05, Wald test; unfiltered differential expression result tables are available in Supplemental Table, B).

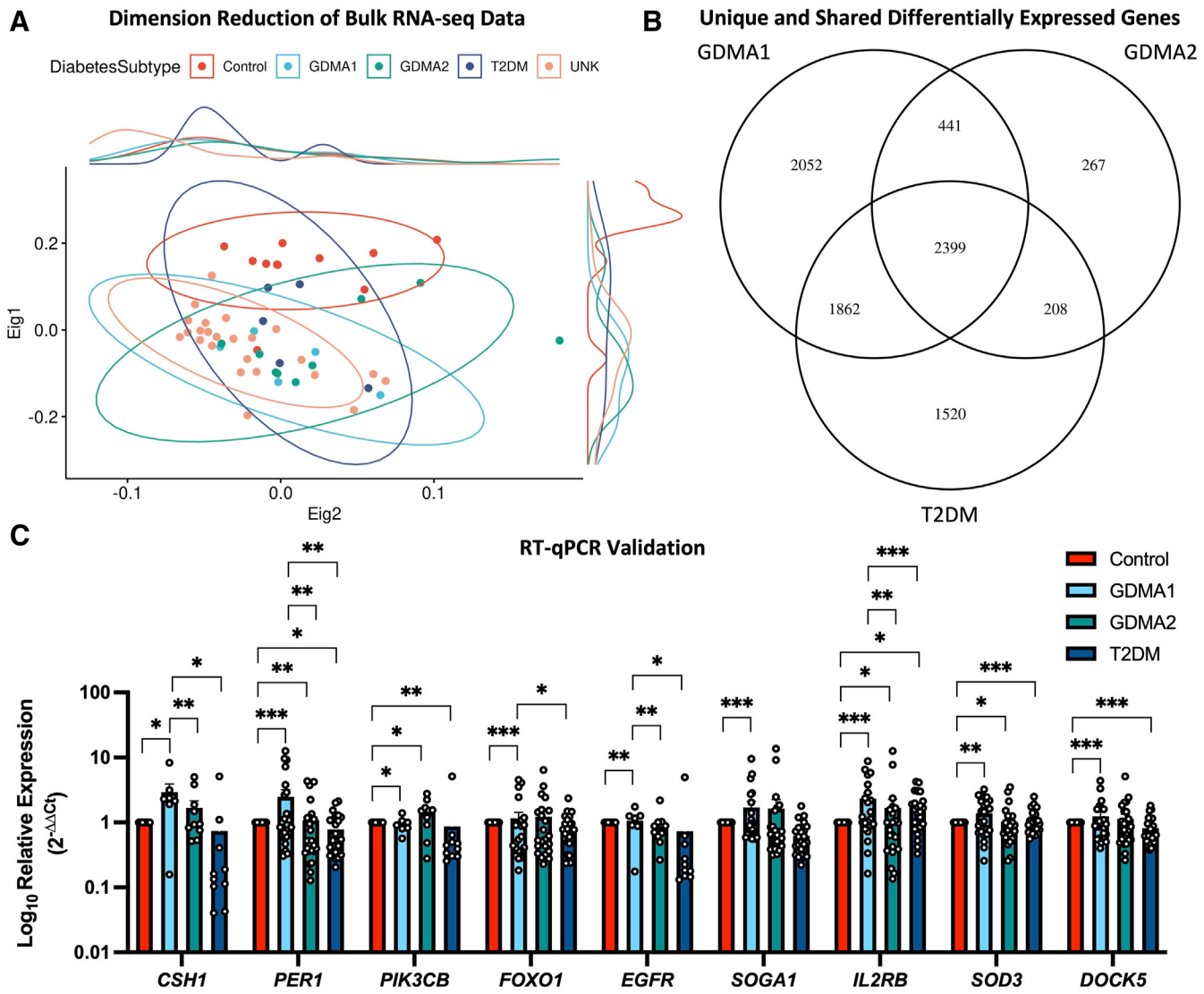
qRT-PCR measures of biologically significant genes of interest in a validation cohort of nonoverlapping subjects

In our a priori designed second phase, a subset of marker genes were selected for validation by virtue of (i) significant differential transcript expression in the RNA-seq analysis (eg, unbiased), or (ii) presumptive or putative biological relevance (eg, curated).[29–32] We defined 'presumptive or putative biological relevance' of curated select genes as having had at least one other study reported in the literature identifying the gene to be potentially important regulators of glucose modulation, obesity or insulin resistance, and/or glucose metabolism. With RT-qPCR, we validated 9 of the top markers from the discovery phase in the second phase validation cohort of $n=122$ placental specimens from nonoverlapping participants which had postpartum GTTs to reliably classify their diabetic disease status or control. Specifically, our phase 2 validation cohort was comprised of $n=29$ nondiabetic controls, $n=25$ GDMA1, $n=35$ GDMA2, and $n=33$ T2DM (Figure 1C). We confirmed that *CSH1* was differentially expressed between T2DM subjects and controls ($P=.02$), GDMA2 ($P=.008$), and GDMA1 ($P=.002$). Significant differences in expression of *EGFR* were demonstrated between T2DM subjects and controls ($P=.02$). Expression of *PER1* differed significantly between controls vs GDMA1 ($P<.001$), GDMA2 ($P=.02$), and T2DM subjects ($P=.02$). Additionally, differential gene expression was noted between GDMA1 when compared to T2DM subjects ($P<.001$) as well as GDMA2 ($P<.001$). *SOGA1* expression differed between controls and GDMA1 ($P=.003$) and GDMA2 ($P=.01$). Expression of *SOD3* differed in GDMA1 as compared to GDMA2 ($P<.001$) and T2DM subjects ($P=.003$). There were also significant expression differences noted between controls and GDMA1 ($P<.001$). Expression of *DOCK5* also differed significantly in controls as compared to GDMA1 ($P=.02$) and GDMA2 ($P=.003$). Expression of *IL2RB* differed between controls and GDMA1 ($P=.008$). Significant differences in expression of *FOXO1* were also noted between controls and GDMA1 ($P=.002$). Expression of *PIK3CB* differed significantly between GDMA2 and T2DM subjects ($P=.04$). Next, we aimed to determine the defining features and biological pathways most perturbed in the placentae of the GDMA1, GDMA2, and T2DM subjects.

GDMA1 subtype gene signature was associated with a differential regulation of pathways mapping to functional pathways including methylation and chromatin remodeling

Compared to controls, we found 6754 significantly DEGs (Figure 2, A; $q<0.05$), *RMRP* stood out as the most significantly upregulated gene, and the GDMA1 gene signature was

FIGURE 1
Naïve transcriptomic analysis of participants with clinically well-characterized preexisting or gestational diabetes and a comparative cohort of deidentified and unclassified participant specimens



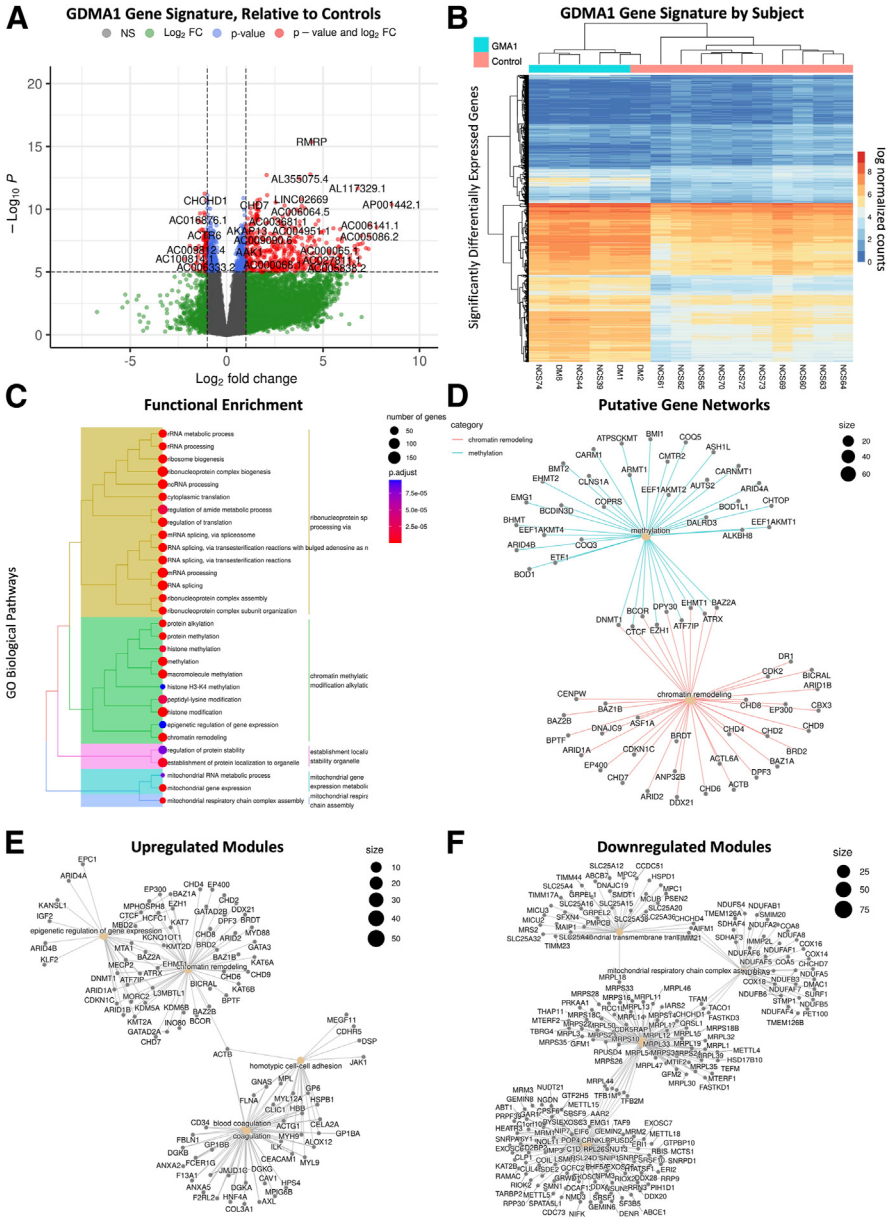
A, Clustering of the bulk RNA-seq data from the n=30 participants in the discovery cohort, and the n=24 placenta bulk RNA-seq participant samples with unknown diabetic status. **B,** Comparison of the 8749 significantly differentially expressed genes (DEGs) for each diabetes subtype relative to controls (adjusted P value (q)<.05, Wald test). **C,** In the nonoverlapping validation cohort of n=145 subjects with clinically well-characterized diabetes (vs control) status, 9 putative disease classifying DEGs identified in the initial bulk RNA-seq experiments were quantitated using reverse transcription and qPCR. Significance (*q<0.05, **q<0.01, ***q<0.001) determined by ordinary 2-way analysis of variance, with Tukey's multiple corrections test and error bars represent the standard error of the mean.

GDMA1, gestational diabetes subtype A1; *GDMA2*, gestational diabetes subtype A2; *RNA-seq*, ribonucleic acid-sequencing; *RT-qPCR*, reverse transcription and quantitative polymerase chain reaction; *T2DM*, type 2 diabetes mellitus.

reproducible across multiple subjects (Figure 2, B). We then performed functional enrichment on these significantly differently expressed genes (Figure 2, C–F) utilizing the GO Biological Pathways database, which identified 263 significantly enriched terms (adjusted P value Bonferroni correction, q<0.05; full GO Biological Pathway results available in Supplemental Table, C). To determine if there was overlap in the genes comprising these significantly enriched terms, we selected 2 closely related categories of chromatin remodeling and methylation (Figure 2, D) and found overlap in 9 genes, including DNA methyltransferase 1 (*DNMT1*), enhancer of zeste homolog 1 (*EZH1*), and ATRX chromatin remodeler (*ATRX*) which are key epigenetic regulators. Next, we assessed the directionality of the gene expression changes in these pathways. Out of the 4287 significantly upregulated genes, the top 3000 were reanalyzed by

FIGURE 2

GDMA1 diabetes subtype gene signature is associated with a differential regulation of transcription pathways mapping to methylation and chromatin remodeling



A, Volcano plot of significant differentially expressed genes (DEGs) in GDMA1 placentae relative to controls. **B,** Heatmap with hierarchical clustering of samples by GDMA1 gene signature (6754 DEGs). **C**—**F,** GO biological processes functional enrichment analysis utilizing clusterProfiler (v4) identified 263 significantly enriched terms (adjusted *P* value Bonferroni correction, $q<0.05$). **C,** Treeplot with hierarchical clustering of biological pathways significantly perturbed. **D,** Selection of categories from (**C**) and visualization of the significant genes and their overlap in these pathways as a cnetplot. **E** and **F,** Visualization of the top 3000 significantly upregulated (**E**—out of 4287 significantly upregulated genes) and downregulated (**F**—out of 2467) pathways and genes visualized as a cnetplot revealing the directionality of gene expression in each pathway and the specific genes perturbed (as determined by $-2<\text{Log}_2[\text{fold-change}]>2$; $q<0.05$).

GDMA1, gestational diabetes subtype A1.

pathway analysis and the top 5 nodes were visualized, revealing top modules in the same pathways of chromatin remodeling, epigenetic regulation of gene expression, coagulation, and homotypic cell-cell adhesion (Figure 2, E). Conversely, the 2467 significantly downregulated genes revealed modules centering around mitochondrial functions including transmembrane transport, respiratory chain complex assembly, translation, and ribonucleoprotein complex biogenesis (Figure 2, F).

GDMA2 subtype gene signature was associated with differential regulation of pathways related to cellular metabolism, cell-substrate adhesion, and wound healing

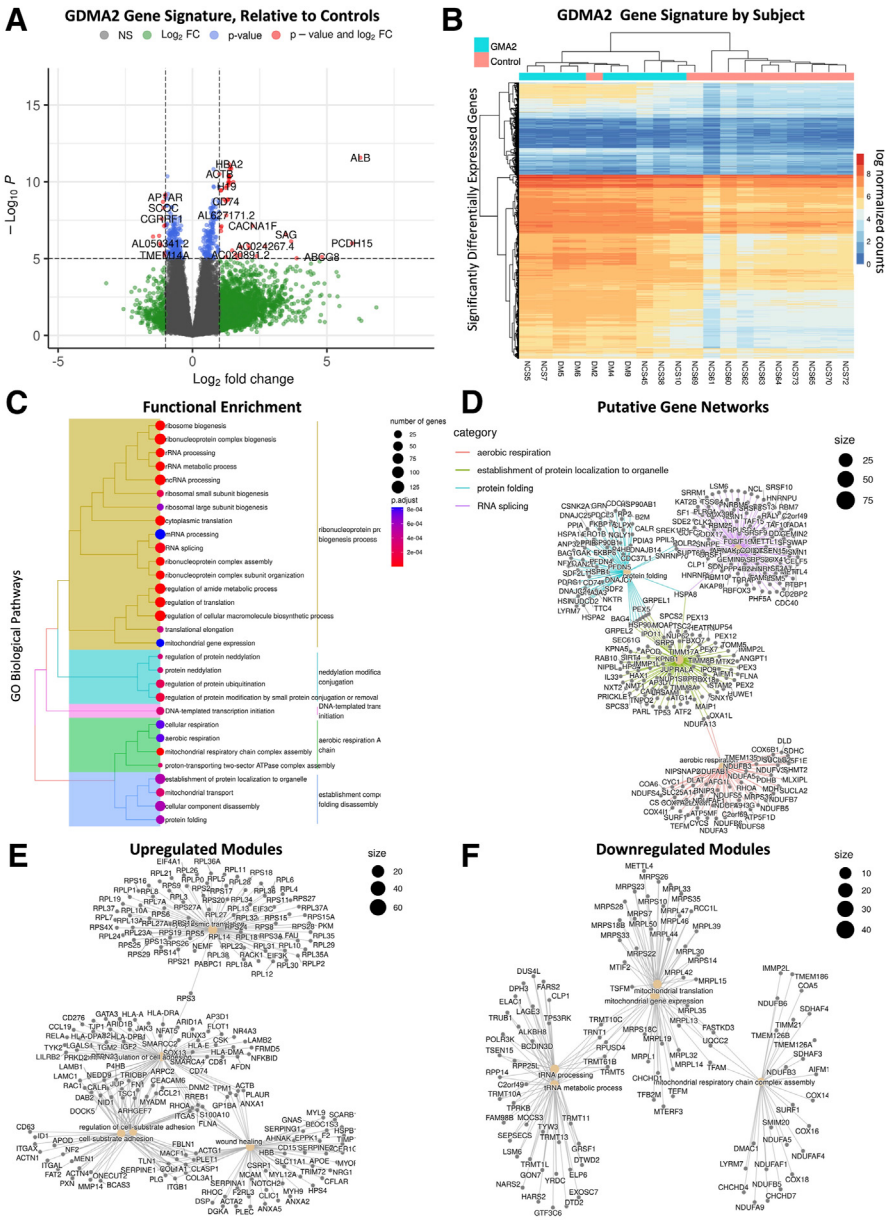
Compared to controls, we found 3315 significantly DEGs (Figure 3, A; $q<0.05$), with *ALB* as the most upregulated and significantly DEG amongst 139 significantly enriched terms ($q<0.05$) from placentae of GDMA2 affected participants (Figure 3, B). We found that heat shock protein family A member 8 (*HSPA8*) was a part of 3 pathways and identified 11 other significant DEGs represented in multiple pathways of biological significance. Next, we assessed the directionality of the gene expression changes in these pathways and sorted the 1678 upregulated and 1637 downregulated genes into distinct GDMA2 gene signature modules based on molecular functions (Figure 3, E and F). Cell-substrate adhesion, cytoplasmic translation, and wound healing were upregulated while mitochondrial gene expression, mitochondrial translation, respiratory chain complex assembly, and transfer ribonucleic acid metabolic processing were downregulated.

T2DM diabetes subtype gene signature associated with upregulation in chromatin remodeling

We found 5989 significantly DEGs (Figure 4, A; $q<0.05$). Functional enrichment identified 256 significantly enriched terms ($q<0.05$), including chromatin remodeling, methylation,

FIGURE 3

A GDMA2 diabetes subtype gene signature is associated with upregulation of pathways involved in wound healing



A, Volcano plot showing significant differentially expressed genes (DEGs) in GDMA2 placentae relative to controls. **B,** Heatmap with hierarchical clustering of samples by the 3315 DEGs unique to the GDMA2 gene signature. **C**–**F,** GO biological processes functional enrichment analysis utilizing clusterProfiler (v4) identified 139 significantly enriched terms (adjusted *P* value Bonferroni correction, $q<0.05$). **C,** Treeplot with hierarchical clustering of biological pathways significantly perturbed. **D,** Selection of 4 categories from (**C**) and the associated genes significantly perturbed visualized as a cnetplot. **E** and **F,** Visualization of the top 3000 significantly upregulated (**E**—out of 1678 upregulated genes) and downregulated (**F**—out of 1637) pathways.

GDMA2, gestational diabetes subtype A2.

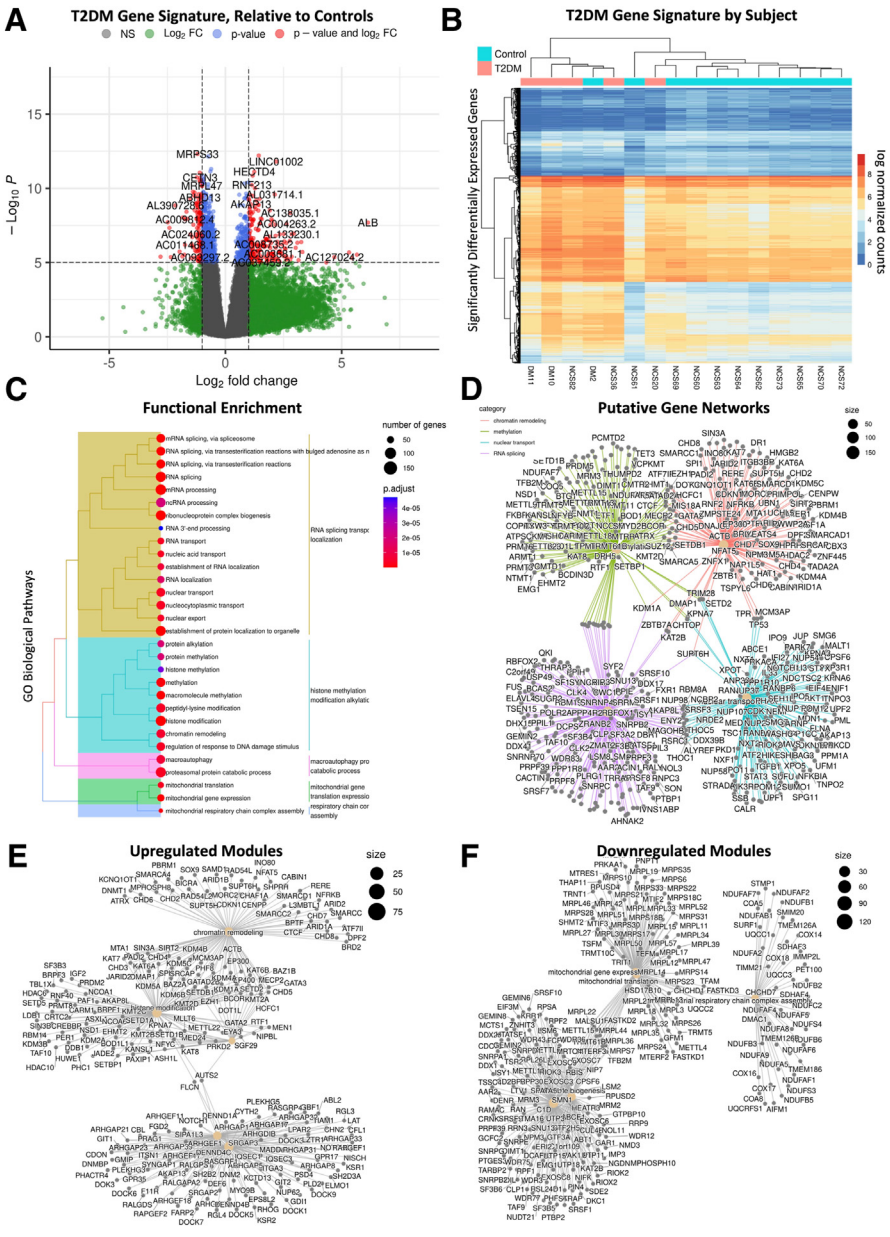
nuclear transport, and RNA splicing pathways, which were analyzed for putative gene networks. We found SET domain containing 2 (*SETD2*), tripartite motif containing 28 (*TRIM28*), SPT6 homolog, histone chaperone (*SUPT6H*) were involved in 3 of those 4 pathways and identified over 50 additional significant DEGs that are involved in multiple pathways. The 1678 upregulated and 1637 downregulated genes revealed distinct T2DM gene signature modules (Figure 3, E and F). Surprisingly and similar to GDMA1 participants placental transcription profiling, that of T2DM participants also showed upregulation of chromatin remodeling, histone modification, histone lysine methylation, and small guanosine triphosphatase mediated signaling were upregulated, while mitochondrial gene expression, translation, respiratory chain complex assembly, translation, ribonucleoprotein complex assembly, and ribosome biogenesis were downregulated.

Spatial and single-cell resolution of placental DEGs parsed by diabetes disease classification in pregnancy

From each of our diabetic and control participants, we found 28 of the marker genes in the GDMA1 bulk RNA-seq analysis aligned with genes that were significantly associated with cell types or spatial niches ($-2<\log_2[\text{fold-change}]>2$ thresholds; $q<0.05$) from the term placenta transcriptomics atlas.[36] The same analysis identified 24 genes from the GDMA2 signature and 33 from the T2DM signature. Therefore, we integrated these analyses to ask (1) within the microarchitecture of the placenta, what part and (2) which cell types harbored the most DEGs when comparing diabetic disease classifications to controls. We performed dimensional reduction of the term placenta atlas on all 291,871 transcriptomes and included an independent analysis of a subset of 19,324 transcriptomes from Yang et al (2021) derived from controls ($n=2$), GDMA1 ($n=1$), or GDMA2 ($n=1$) placental specimens (Figure 5, A). We then performed differential expression analysis on the unique diabetes gene signatures from the bulk RNA-seq analysis (Figures 1–4) projected onto placental regions sampled in the term placental atlas (Figure 5, B; differential expression results available in Supplemental Table, D). The average expression of each

FIGURE 4

A T2DM diabetes subtype gene signature is associated with upregulation of pathways involved in chromatin remodeling



A, Volcano plot showing significant differentially expressed genes (DEGs) in T2DM placentae relative to controls. **B,** Heatmap with hierarchical clustering of samples by the 5989 DEGs in the T2DM gene signature. **C**—**F,** GO biological processes functional enrichment analysis utilizing clusterProfiler (v4) identified 256 significantly enriched terms (adjusted *P* value Bonferroni correction, $q<0.05$). **C,** Treeplot with hierarchical clustering of biological pathways significantly perturbed. **D,** Selection of 4 categories from (**C**) and the associated genes significantly perturbed visualized as a cnetplot. **E** and **F,** Visualization of the top 3000 significantly upregulated (**E**—out of 1678 upregulated genes) and downregulated (**F**—out of 1637).

T2DM, type 2 diabetes mellitus.

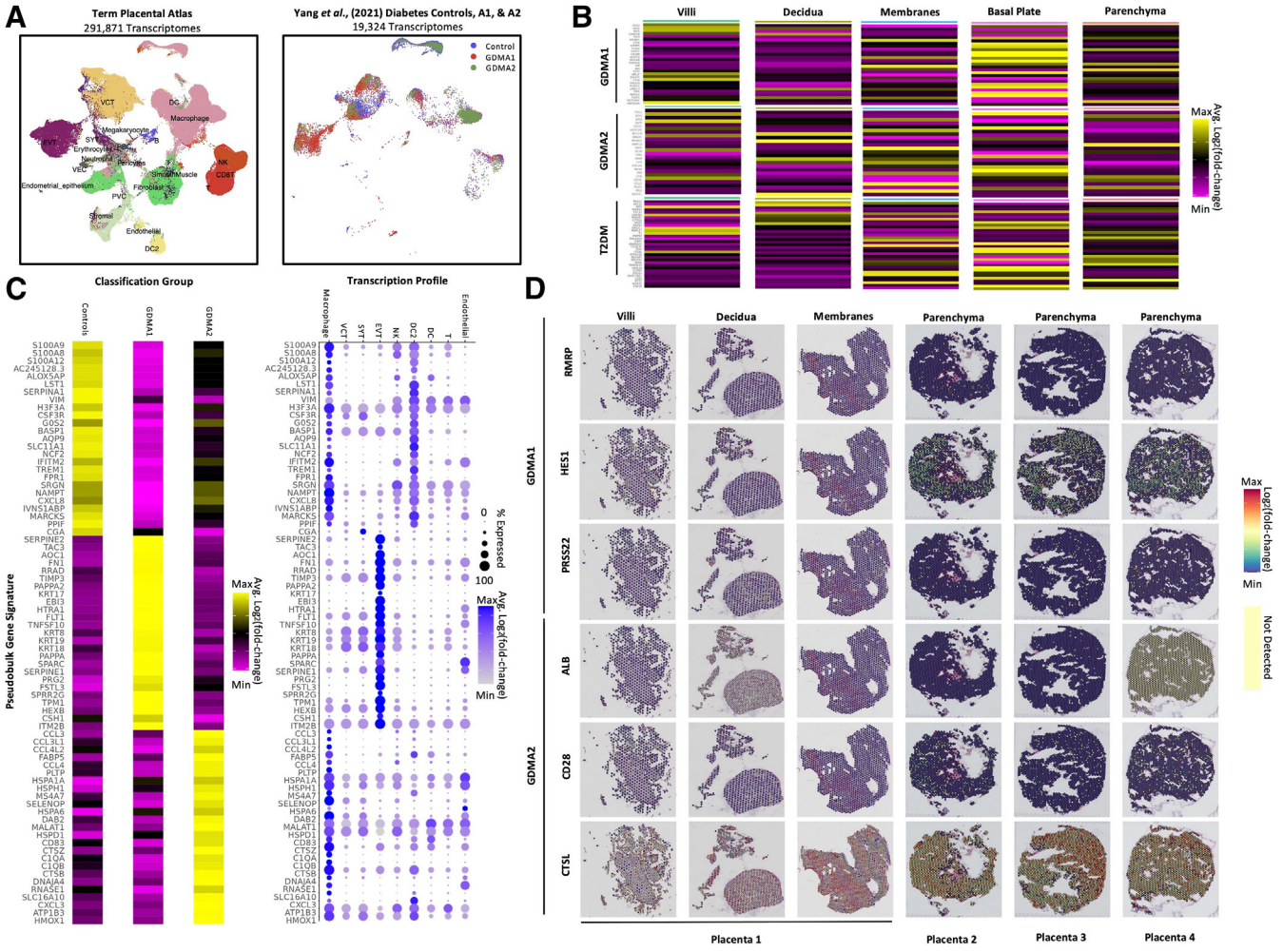
marker was highly variable in each region of the placenta, warranting higher resolution via subset analyses of the GDMA1 and GDMA2 scRNA-seq placental specimen data from Yang et al (2021) projected as classification group or single-cell profile (Figure 5, C; full differential expression results available in Supplemental Table, E). Consistent with our bulk RNA-seq analysis, we observed unique distinctions of diabetic disease classifications were faithfully recapitulated in the placental transcriptomics analysis at a single-cell level. Specifically, we found that GDMA1 single-cell gene signature aligned with extravillous trophoblasts, while the GDMA2 signature aligned with macrophages. To visualize the spatial gene expression and identify the transcriptional microarchitecture of top mapping gene markers in the placenta (Figure 5, D), we collected placentae from n=4 well-characterized healthy and nondiabetic controls for spatial transcriptomic analyses and visualized key genes of interest in the chorionic villi, decidua, chorioamniotic membranes, and parenchyma (Figure 5, D). Based on our independent analysis of the dataset of Yang et al (2021), we visualized 2 genes known to be independent of diabetic status (Hes family BHLH transcription factor 1 [*HES1*], protease serine 22 [*PRSS22*]) as positive controls for spatial discrimination in the placenta. Faithfully recapitulating our bulk RNA-seq data, we found that the spatial expression of GDMA1 and GDMA2 top mapping genes (RNA component of mitochondrial RNA processing endoribonuclease [*RMRP*], albumin [*ALB*] CD28 molecule [*CD28*], cathepsin L [*CTSL*]) were highly dependent on the microenvironment niches, suggesting that placental cell type and microarchitectural transcript profiles varied by diabetes subtype. With this additional layer of cellular and spatial complexity, we hypothesized that established computational methodologies for artificial intelligence (eg, machine learning prediction models) may provide naïve and unbiased aid in identifying predictive features of GDMA1 and GDMA2 placental gene expression.

Machine learning prediction models from placental transcriptional data yield accurate diabetes disease classification

The ground truths and workflow for our machine learning approach are outlined

FIGURE 5

Spatial and single-cell resolution of placental differentially expressed genes parsed by diabetes disease classification in pregnancy



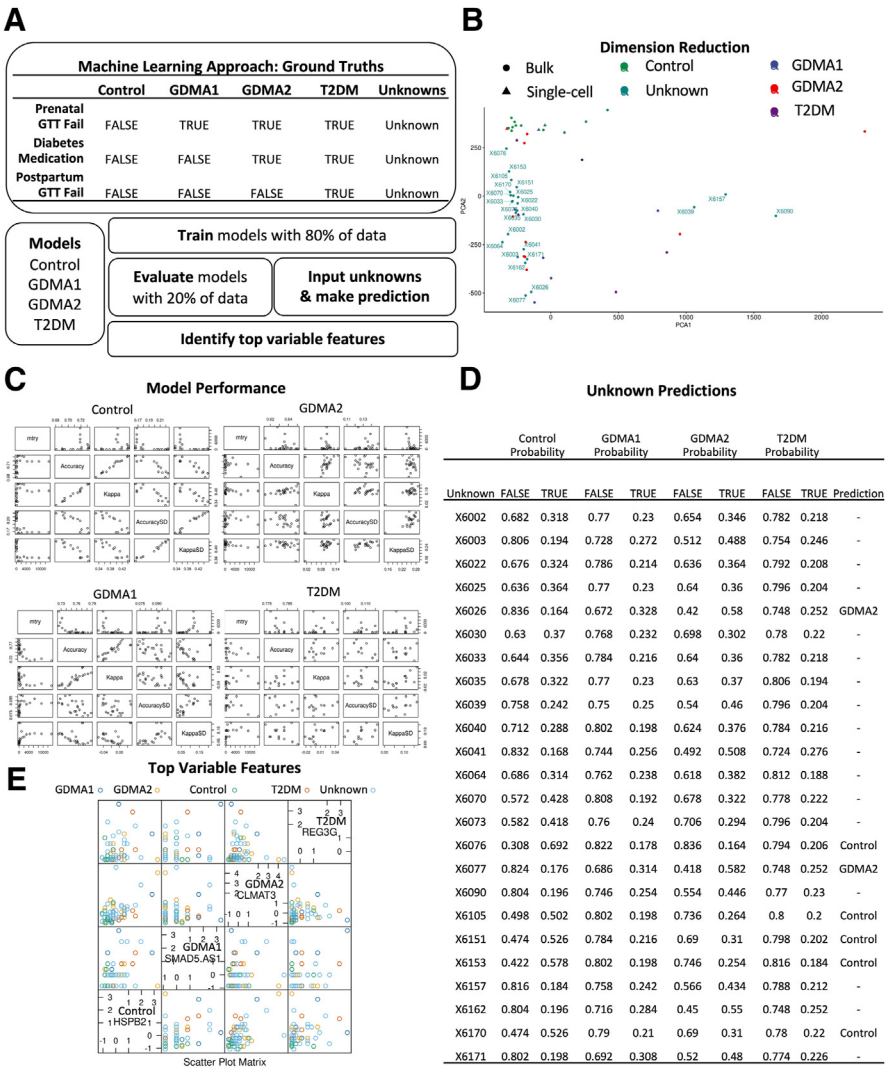
A, Dimensional reduction of the term placenta atlas (291,871 transcriptomes) or independent analysis of a subset of 19,324 transcriptomes from Yang et al (2021) derived from controls (n=2), GDMA1 (n=1), or GDMA2 (n=1) specimens. **B,** Differential expression of the unique diabetes gene signatures from the bulk RNA-seq analysis arising from the term placental atlas. **C,** Subset analysis of data from Yang et al (2021) projected as classification group (left) or single-cell profile (right). **D,** Spatial gene expression of the top DEGs from the bulk and scRNA-seq analysis in the villi, decidua, or membranes of 1 placenta, or the parenchyma from 3 separate placentae.

DEGs, differentially expressed genes; *GDMA1*, gestational diabetes subtype A1; *GDMA2*, gestational diabetes subtype A2; *RNA-seq*, ribonucleic acid-sequencing; *scRNA-seq*, single-cell RNA-sequencing.

in Figure 6, A, and benefited from our incorporation of not only placentae from known control cases and diabetic participants, but also bulk RNA-seq placenta data with unknown clinical metadata or GDM classification (ie, termed "unknowns"). We generated 4 random forest models and evaluated their performance (Figure 6, C). Amongst the 80% of data in the training models, the OOB error rate for each model was 29.63% for control, 18.52% for GDMA1, 37.04% for GDMA2, and 18.52% for T2DM. We then tested the accuracy of each model with the remaining 20% of data, which were untouched but had known and well-characterized diabetes classifications and all 4 models demonstrated 100% sensitivity and specificity (kappa=1). Given the high kappa of our 20% validation dataset relative to the modest performance of our training models, we inputted our "unknowns" into each of the models and reported the probabilities in Figure 6, D. The models could predict the phenotypes of 7 specimens with confidence and aligned with the dimensional reduction (Figure 1, A), as the "unknowns" called controls were most proximal in their distance metric from known control participants placental gene expression profiles.

In addition, we examined the components of each machine learning model to determine the DEGs that contributed

FIGURE 6
Machine learning prediction models from placental transcriptional data yielding diabetes classification



A, Ground truths and workflow of the machine learning process. **B,** Principal component analysis dimension reduction of bulk RNA-seq and single-cell RNA-sequencing samples with unknowns labeled by sample identification. **C,** Each model was trained and evaluated with a randomized 80/20 scheme. **D,** Following evaluations with each model, we predicted 4 of the unknowns were controls, 0 were GDMA1, 2 were GDMA2, and 0 were T2DM based on their placental gene signatures. **E,** Counts for the transcript that accounted for the highest degree of variation in each model (see Supplemental Table, F for all genes).

GDMA1, gestational diabetes subtype A1; *GDMA2*, gestational diabetes subtype A2; *RNA-seq*, ribonucleic acid-sequencing; *T2DM*, type 2 diabetes mellitus.

the most variation. The contribution values of each transcript to the model are available in Supplemental Table, F and the top transcript for each model is plotted with the counts for each sample in Figure 6, E. Interestingly, the top features identified by our 4 machine learning models were not significantly differentially expressed in our bulk RNA-seq analysis; however, there were significantly DEGs found within the top 8 in the control model (*CHA2*), 7 for GDMA1 (*CTS4*), 6 for GDMA2 (*HIST1D*), and the 3 with T2DM (thymocyte selection associated [*THEMIS*]). Lastly, several long noncoding RNAs (lncRNAs) with no known relevance to GDM or pregnancy were found to have high-ranking contributions to these models. Notably, lncRNA AC110597-1 ranked second in the control model with a contribution score of 64.1. AC005034-6 was the fourth highest contributor to the GDMA1 model with a score of 91.5. In the GDMA2 model, UBE2Q1-AS1 had the second highest contribution at 80.9, and in the T2DM model, AL645465-1 ranked second with a high contribution score of 98.1. These findings suggest a promising potential of artificial intelligence and machine learning as tools for unbiased hypothesis generation, serving as valuable complements to the rigorous orthogonal methods essential for hypothesis testing.

Comment

Principal findings

From GDM and T2DM cases and nondiabetic controls comprising our discovery cohort, rigorously defined by postpartum GTT, we demonstrate unique molecular signatures in placental gene expression by virtue of diabetes subtype. In a nonoverlapping validation cohort, we validate placental expression of several key genes (*CSH1*, *PER1*, *PIK3CB*, *EGFR*, *SOD3*, *DOCK5*). We then utilize a term placenta transcriptomics atlas to assess the GDMA1, GDMA2, and T2DM gene signatures by placenta region, cell type, and niche microenvironment. Lastly, we integrate these bulk and high-resolution datasets to create machine learning models to predict the diabetes subtype based on placenta gene expression profiles. Collectively, these data demonstrate that molecular profiles significantly differ among diabetic classifications in pregnancy and those with pregestational diabetes, relative to nondiabetic controls.

This study provides molecular evidence demonstrating that the placental transcriptome is distinctly different by diabetic disease classification. Since the placenta is established within days of implantation, these findings raise the possibility that earlier diagnostic measures may be developed, and could

potentially lead to improved maternal and fetal outcomes. It stands to reason that variation in clinical diabetic disease states (and their associated manifestations) should have distinct molecular signatures.[27,41–44] The placenta is an active metabolic organ crucial to glucose transport and homeostasis,[45] and we have shown here that the placenta maintains a footprint of the DEGs which are associated with diabetic disease classifications. These unique transcriptomic signatures of each of the distinctive diabetic disease states beg the question: are the changes in gene expression pathways in the placenta a cause or consequence of maternal diabetes? The current study is cross-sectional and observational, and thus cannot provide causal links. However, by utilizing prenatal and postpartum verified GDM subtype classifications, we were able to demonstrate significant differences in placental transcriptomes and functional profiles detected by unbiased methods and validated with RT-qPCR. Further, we extend these GDMA1, GDMA2, and T2DM placental gene signatures to specific placental regions, cell types, and spatial niches. Lastly, we integrate bulk and high-resolution data to generate 4 machine learning models which utilized the placental gene expression profiles to make predictions of participants diabetes classification.

Although the OOB error performance of our machine learning models were relatively modest, with increasing sample sizes and additional improvements in resolution, machine learning models will likely improve, and our data reported herein provides a rationale for longitudinal studies that will track the correlations of top biomarkers with noninvasive maternal diagnostics. Compared to other models,[46] this study demonstrated significant results utilizing placental transcriptomic variation across different diabetic disease states to predict diabetes classification even with limited sample sizes, suggesting that even small, well-curated 'omics datasets can be informative with the appropriate AI modeling and applications. An additional strength of the machine learning approach included focusing on specific sets of molecular markers and less reliance on extensive clinical metadata, mitigating privacy concerns associated with alternative machine learning models that rely on substantial amounts of protected data to be trained. Moreover, compared with other models that can be difficult to interpret, this study's clinical applicability and interpretability offer tangible, understandable insights for clinicians and provide clinically relevant findings. Lastly, in contrast to other models, these analyses harness AI as a complement to rigorous transcriptomics analyses and existing diabetes diagnosis ground truths (Figure 6, A), avoiding the potential pitfall of rewriting diagnosis rules in a black box, and maintaining alignment with the ethical use of AI as a support tool in clinical research.

Results in the context of what is known

We based our selection of genes for further validation with RT-qPCR on their capacity to drive principal components clustering along both PC1 and PC2 axes (Figure 1). Notably, not only did these candidates drive significant differential expression per our RNA-seq analysis, but they are summarily described in the scientific literature as important modulators in metabolism, glucose utilization, and obesity. *PER1*, like other circadian clock proteins, is instrumental in encoding the circadian rhythms of metabolism and has been implicated in gestational diabetes.[47] *IL2RB*, plays a key role in T cell mediated immune response and has been associated with the development of type I diabetes.[48] Others have shown a relationship between *SOD3* levels, fasting plasma glucose, body mass index, and insulin resistance.[49] El-Sayed Moustafa et al,[50] describe *DOCK5* as a susceptibility gene for severe obesity. The cluster of *CSH* genes, also known as placental lactogen, is thought to play a role in maternal adaptation to pregnancy, preeclampsia, intrauterine growth, and gestational diabetes.[51] Specifically, placental expression of *CSH1* was shown to be significantly higher in large for gestational age offspring compared to average and small for gestational age counterparts.[51] Here, we found *CSH1* was significantly higher between controls and GDMA1 (2.91-fold, $P=0.002$) and GDMA2 (1.67-fold, $P=0.008$) subjects, but found no significant differences in birthweight between groups (Table). *EGFR* signaling has been described as a link to the growth of macrosomic fetuses in mothers with gestational diabetes.[52] *PIK3CB* influences glucose metabolism,[53] and others have associated overexpression of *PIK3CB* with development of gestational diabetes.[54] *FOXO1* is an important factor in the regulation of insulin and glucose metabolism.[55]

Research implications

Albumin was the most commonly upregulated gene for both GDMA2 and T2DM relative to controls (Figures 3, A and 4, A), and the diagnostic value of glycated albumin as a biomarker for diabetes has been previously studied inside and outside the context of pregnancy.[56–59] *RMRP*, the most upregulated gene in GDMA1 placentae, is a lncRNA that was found to act as an antagomir for endogenous miR-1a-3p, leading to *JunD* expression and diabetic nephropathy in a mouse model.[60] CD28 gene expression was a consistent marker for GDMA2 in the bulk and scRNA-seq datasets (Figure 5, C and D). In a case-control flow cytometry study of maternal peripheral blood, GDM subjects had prolonged CD28 expression with high frequencies of T-cell activation.[61] In a recent study, placental gene expression changes in GDM subjects relative to controls also observed alterations in epigenetic regulation and profiled differential methylation profiles by bisulfite sequencing in paired neonatal umbilical cord blood samples.[62] They found several genes that had altered methylation and gene expression patterns associated with GDM including suppressor of cytokine signaling 3 (*SOCS3*), which, according to the term placenta transcriptomics atlas map to the chorioamniotic membranes, endometrial and dendritic cells, low-density lipoprotein receptor (*LDLR*), which map to the basal plate and endothelial cells,

and growth arrest and DNA damage inducible alpha (*GADD45A*), which map to syncytiotrophoblasts and megakaryocytes. In an effort to correlate adverse pregnancy outcomes with indicator biomarkers, first- and second-trimester cell-free DNA methylation and cell-free transcriptomes were analyzed,[63] 2 of their top markers S100A8 and *MS4A3* aligned with macrophages and natural killer cells according to the placenta transcriptomics atlas.

Clinical implications

Comparison of genes differentially expressed between diabetic subjects and controls revealed that a subset of the genes is shared between diabetic subtypes, while retaining genes unique to gestational diabetes and preexisting type 2 DM (Figure 1, B). Each cohort's unique gene expression pattern, alongside cellular and spatial specificity, are highly consistent at a molecular level, and suggest that individualized placental signatures faithfully recapitulate the clinical distinct entities of GDMA1, GDMA2, and T2DM, and cannot be attributed to medication alone. Transcriptomic pathway analysis revealed significant enrichment in metabolic pathways in all classes of diabetics as compared to controls, highlighting the clinical relevance of the DEGs. These findings are further corroborated by the recent examination of genome-wide association studies which found several loci uniquely associated with GDM, expanding upon earlier findings of shared genes that predisposed individuals toward GDM and T2DM.[64,65] While diabetes subtypes may have some phenotypic similarities, clinically relevant genotypic and transcriptional differences setting each subtype apart will become paramount to correlative longitudinal noninvasive biomarker discovery studies.

Strengths and limitations

There are both strengths and limitations to our study. First, in translational and reference resource studies such as ours, efforts aimed at rigorous clinical classification are important. Other investigators have previously argued that there is potentially significant overlap and risk of misclassification between T2DM and GDMA2, especially in populations who have limited access to care.[66] Despite the challenges for postpartum assessment described in the literature,[67] our population demonstrated strong adherence to care with a 71.3% clinical testing follow-up rate. Second, we set out to enroll a largely Hispanic population. Compared to non-Hispanic whites, Hispanic patients suffer significantly higher rates of diabetes, with 50% of those with gestational diabetes developing T2DM within 5 years.[68] Despite this high rate of disease progression, Hispanics have the lowest rates of postpartum diabetes follow-up, with as few as 1 in 5 getting screened for T2DM at their postpartum visit, potentially increasing the number who enter a subsequent pregnancy with undiagnosed T2DM.[67] Additional strengths of our study include our computational methodology with extensive quality control, with purposeful utilization of an unsupervised learning approach to minimize bias in our results. The choice of genes for RT-qPCR validation was supported not only by our RNA-seq data but also by biological plausibility and corroboration in the literature. We compared the overlap of the differential expression results from the bulk RNA-seq datasets, validated 9 genes in an independent nonoverlapping RT-qPCR cohort, and assessed the gene signatures in a term transcriptomics atlas. We found the cell-type niches for these genes of interest to align with the heterogeneous cell types and complex microarchitecture of the placenta. Together, these results suggest the cell types within the placenta responsible for GDM subtypes are potentially rare subpopulations, which gene expression changes may become masked with bulk analysis.

We acknowledge that there are weaknesses of this study, which limit the applicability of these findings to its generalizability. The cohorts analyzed were not powered to stratify nor parse comparisons by fetal sex, ethnicity, or race. Additional studies assessing potential differences associated with preterm birth, infection status, mode of delivery, placental pathology, sexual dimorphisms, and racial/ethnic health disparities are warranted. Since our high-risk cohort reflects a predominantly Hispanic/Latino population with a well-described risk for both pregestational and gestational diabetes, the alterations in placental gene signature may be specific to the particular molecular mechanisms driving disease in this population. However, this may arguably be a strength of this study given the attributable population risk as previously detailed. Additionally, there are certainly other factors involved in placental response to altered glucose homeostasis not tested in the current study, including additional maternal influences such as hormonal signaling or maternal-fetal genetic interactions. However, our unbiased approach in large discovery and nonoverlapping validation cohorts (n=176 subjects in total, inclusive of both our discovery and validation cohorts) should minimize such misclassification errors. Moreover, employing targeted confirmation of findings in our expanded and nonoverlapping validation cohort (n=122 subjects) enabled a broad examination of the placental reaction to disruptions in maternal metabolism by examining the downstream integration of these signals on the critical mediator of fetal nutrient supply.

Conclusions

In summary, this work moves us closer to appreciating, and eventually understanding, the difference between the spectrum of normal placental physiology and the effects of diabetes in pregnancy at a mechanistic level. Our robust placental gene expression analysis, both with bulk RNA-seq and RT-qPCR validation, demonstrates that despite expected phenotypic overlap, placental gene expression differs among gestational diabetics, those with pregestational diabetes, and nondiabetic controls. Our use of a naïve analytical pipeline, blind to disease classification, was able to group specimens by similar gene expression profiles, and functionally and faithfully recapitulate clinically meaningful distinctions. Furthermore, functional pathway analyses confirmed

the clinical relevance of the DEGs. This lends credence to the current classification of pregestational and gestational diabetes and potentially lays the groundwork for the future development of distinct clinical algorithms aimed at earlier and more accurate screening for and treatment of underlying diabetic pathology. Further studies are needed to achieve the long-term goal of identifying novel biomarkers to improve the diagnosis and management of diabetes in pregnancy. ■

Acknowledgments

We would like to acknowledge the participants who willingly donated their placental tissue, J. Chen, and the PeriBank team of coordinators for their collection of placental specimens and abstraction and entry of data. The authors wish to acknowledge members of the Aagaard Lab for the critical review of the manuscript.

References

1. Wang H, Li N, Chivese T, et al. IDF diabetes atlas: estimation of global and regional gestational diabetes mellitus prevalence for 2021 by International Association of Diabetes in Pregnancy Study Group's criteria. Diabetes Res Clin Pract 2022;183:109050.

2. Gregory EC, Ely DM. Trends and characteristics in gestational diabetes: United States, 2016-2020. Natl Vital Stat Rep 2022;71:1–15.

3. Desoye G, Hauguel-de Mouzon S. The human placenta in gestational diabetes mellitus. The insulin and cytokine network. Diabetes Care 2007;30 Suppl 2:S120–6.

4. Robitaille J, Grant AM. The genetics of gestational diabetes mellitus: evidence for relationship with type 2 diabetes mellitus. Genet Med 2008;10:240–50.

5. ElSayed NA, Aleppo G, Aroda VR, et al. 2. Classification and diagnosis of diabetes: standards of care in diabetes-2023. Diabetes Care 2023;46:S19–40.

6. Radaelli T, Varastehpour A, Catalano P, Hauguel-de Mouzon S. Gestational diabetes induces placental genes for chronic stress and inflammatory pathways. Diabetes 2003;52:2951–8.

7. Gestational diabetes mellitus. Washington, DC: American College of Obstetricians and Gynecologists; 2018.

8. Janevic T, McCarthy K, Liu SH, et al. Racial and ethnic inequities in development of type 2 diabetes after gestational diabetes mellitus. Obstet Gynecol 2023;142:901–10.

9. Herrick CJ, Keller MR, Trolard AM, Cooper BP, Olsen MA, Colditz GA. Factors associated with postpartum diabetes screening in women with gestational diabetes and medicaid during pregnancy. Am J Prev Med 2021;60:222–31.

10. Sinha DD, Williams RC, Hollar LN, et al. Barriers and facilitators to diabetes screening and prevention after a pregnancy complicated by gestational diabetes. PLoS One 2022;17:e0277330.

11. Yang Y, Guo F, Peng Y, et al. Transcriptomic profiling of human placenta in gestational diabetes mellitus at the single-cell level. Front Endocrinol (Lausanne) 2021;12:679582.

12. Gauster M, Desoye G, Totsch M, Hiden U. The placenta and gestational diabetes mellitus. Curr Diab Rep 2012;12:16–23.

13. Armistead B, Johnson E, VanderKamp R, et al. Placental regulation of energy homeostasis during human pregnancy. Endocrinology 2020;161:bqaa076.

14. Jayabalan N, Nair S, Nuzhat Z, et al. Cross talk between adipose tissue and placenta in obese and gestational diabetes mellitus pregnancies via exosomes. Front Endocrinol (Lausanne) 2017;8:239.

15. Carrasco-Wong I, Moller A, Giachini FR, et al. Placental structure in gestational diabetes mellitus. Biochim Biophys Acta Mol Basis Dis 2020;1866:165535.

16. Kedziora SM, Obermayer B, Sugulle M, et al. Placental transcriptome profiling in subtypes of diabetic pregnancies is strongly confounded by fetal sex. Int J Mol Sci 2022;23:15388.

17. Meyrueix LP, Gharaibeh R, Xue J, et al. Gestational diabetes mellitus placentas exhibit epimutations at placental development genes. Epigenetics 2022;17:2157–77.

18. Zhao YH, Wang DP, Zhang LL, Zhang F, Wang DM, Zhang WY. Genomic expression profiles of blood and placenta reveal significant immune-related pathways and categories in Chinese women with gestational diabetes mellitus. Diabet Med 2011;28:237–46.

19. Napso T, Yong HEJ, Lopez-Tello J, Sferruzzi-Perri AN. The role of placental hormones in mediating maternal adaptations to support pregnancy and lactation. Front Physiol 2018;9:1091.

20. Lapolla A, Traldi P. Proteomic approaches in the study of placenta of pregnancy complicated by gestational diabetes mellitus. Biomedicines 2022;10:2272.

21. Huynh J, Dawson D, Roberts D, Bentley-Lewis R. A systematic review of placental pathology in maternal diabetes mellitus. Placenta 2015;36:101–14.

22. Moyce BL, Dolinsky VW. Maternal beta-cell adaptations in pregnancy and placental signalling: implications for gestational diabetes. Int J Mol Sci 2018;19:3467.

23. Lekva T, Lyle R, Roland MC, et al. Gene expression in term placentas is regulated more by spinal or epidural anesthesia than by late-onset preeclampsia or gestational diabetes mellitus. Sci Rep 2016;6:29715.

24. Magee TR, Ross MG, Wedekind L, Desai M, Kjos S, Belkacemi L. Gestational diabetes mellitus alters apoptotic and inflammatory gene expression of trophobasts from human term placenta. J Diabetes Complications 2014;28:448–59.

25. Shah NS, Wang MC, Freaney PM, et al. Trends in gestational diabetes at first live birth by race and ethnicity in the US, 2011-2019. JAMA 2021;326:660–9.

26. Gestational hypertension and preeclampsia: ACOG Practice Bulletin summary, number 222. Obstet Gynecol 2020;135:1492–5.

27. Suter M, Ma J, Harris A, et al. Maternal tobacco use modestly alters correlated epigenome-wide placental DNA methylation and gene expression. Epigenetics 2011;6:1284–94.

28. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30:2114–20.

29. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.

30. Broad Institute. Picard Toolkit. GitHub Repository: Broad Institute. 2019. Available at: https://broadinstitute.github.io/picard/. Accessed December 28, 2022.

31. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 2015;31:166–9.

32. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:1–21.

33. Wu T, Hu E, Xu S, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. Innovation 2021;2:100141.

34. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 2001;25:402–8.

35. Suter M, Abramovici A, Showalter L, et al. In utero tobacco exposure epigenetically modifies placental CYP1A1 expression. Metabolism 2010;59:1481–90.

36. Barrozo ER, Seferovic MD, Castro ECC, et al. SARS-CoV-2 niches in human placenta revealed by spatial transcriptomics. Med 2023;4:612–34.e4.

37. Tsang JCH, Vong JSL, Ji L, et al. Integrative single-cell and cell-free plasma RNA transcriptomics elucidates placental cellular dynamics. Proc Natl Acad Sci U S A 2017;114:E7786–95.

38. Pique-Regi R, Romero R, Tarca AL, et al. Single cell transcriptional signatures of the human placenta in term and preterm parturition. Elife 2019;8:e52004.

39. Pique-Regi R, Romero R, Tarca AL, et al. Does the human placenta express the canonical cell entry mediators for SARS-CoV-2? Elife 2020;9:e58716.

40. Garcia-Flores V, Romero R, Xu Y, et al. Maternal-fetal immune responses in pregnant women infected with SARS-CoV-2. Nat Commun 2022;13:320.

41. Sbrana E, Suter MA, Abramovici AR, et al. Maternal tobacco use is associated with increased markers of oxidative stress in the placenta. Am J Obstet Gynecol 2011;205:246.e1-7.

42. Li Q, Kappil MA, Li A, et al. Exploring the associations between microRNA expression profiles and environmental pollutants in human placenta from the National Children's Study (NCS). Epigenetics 2015;10:793–802.
43. Antony KM, Ma J, Mitchell KB, Racusin DA, Versalovic J, Aagaard K. The preterm placental microbiome varies in association with excess maternal gestational weight gain. Am J Obstet Gynecol 2015;212:653.e1-e16.
44. Aagaard-Tillery K. The placenta harbors a unique microbiome. Washington, DC: Science Translational Medicine; 2014.
45. Mendez-Figueroa H, Shchelochkov OA, Shaibani A, Aagaard-Tillery K, Shinawi MS. Clinical and biochemical improvement of very long-chain acyl-CoA dehydrogenase deficiency in pregnancy. J Perinatol 2010;30:558–62.
46. Sarno L, Neola D, Carbone L, et al. Use of artificial intelligence in obstetrics: not quite ready for prime time. Am J Obstet Gynecol MFM 2023;5:100792.
47. Pappa KI, Gazouli M, Anastasiou E, Iliodromiti Z, Antsaklis A, Anagnou NP. Circadian clock gene expression is impaired in gestational diabetes mellitus. Gynecol Endocrinol 2013;29:331–5.
48. Saleh NM, Raj SM, Smyth DJ, et al. Genetic association analyses of atopic illness and proinflammatory cytokine genes with type 1 diabetes. Diabetes Metab Res Rev 2011;27:838–43.
49. Adachi T, Inoue M, Hara H, Maehata E, Suzuki S. Relationship of plasma extracellular-superoxide dismutase level with insulin resistance in type 2 diabetic patients. J Endocrinol 2004;181:413–7.
50. El-Sayed Moustafa JS, Eleftherohorinou H, de Smith AJ, et al. Novel association approach for variable number tandem repeats (VNTRs) identifies DOCK5 as a susceptibility gene for severe obesity. Hum Mol Genet 2012;21:3727–38.
51. Mannik J, Vaas P, Rull K, Teesalu P, Laan M. Differential placental expression profile of human growth hormone/chorionic somatomammotropin genes in pregnancies with pre-eclampsia and gestational diabetes mellitus. Mol Cell Endocrinol 2012;355:180–7.
52. Li J, Song L, Zhou L, et al. A MicroRNA signature in gestational diabetes mellitus associated with risk of macrosomia. Cell Physiol Biochem 2015;37:243–52.
53. Du Z, Uversky VN. A comprehensive survey of the roles of highly disordered proteins in type 2 diabetes. Int J Mol Sci 2017;18:2010.
54. Wojcik M, Mac-Marcjanek K, Wozniak LA, Nadel I, Lewinski A, Cypryk K. The association of leukocyte phosphatidylinositol 3-kinase delta overexpression with gestational diabetes mellitus (GDM). Endokrynol Pol 2014;65:17–24.
55. Xu Y, Jin B, Sun L, Yang H, Cao X, Zhang G. The expression of FoxO1 in placenta and omental adipose tissue of gestational diabetes mellitus. Exp Clin Endocrinol Diabetes 2014;122:287–94.
56. Ciaccio M. Introduction of glycated albumin in clinical practice. J Lab Precis Med 2019;4:28.
57. Fang M, Daya N, Coresh J, Christenson RH, Selvin E. Glycated albumin for the diagnosis of diabetes in US adults. Clin Chem 2022;68:413–21.
58. Chume FC, Kieling MH, Correa Freitas PA, Cavagnolli G, Camargo JL. Glycated albumin as a diagnostic tool in diabetes: an alternative or an additional test? PLoS One 2019;14:e0227065.
59. Xiong JY, Wang JM, Zhao XL, et al. Glycated albumin as a biomarker for diagnosis of diabetes mellitus: a systematic review and meta-analysis. World J Clin Cases 2021;9:9520–34.
60. Yang H, Wang J, Zhang Z, et al. Sp1-Induced lncRNA rmrp promotes mesangial cell proliferation and fibrosis in diabetic nephropathy by modulating the miR-1a-3p/JunD pathway. Front Endocrinol (Lausanne) 2021;12:690784.
61. Pendeloski KP, Mattar R, Torloni MR, Gomes CP, Alexandre SM, Daher S. Immunoregulatory molecules in patients with gestational diabetes mellitus. Endocrine 2015;50:99–109.
62. Lu S, Wang J, Kakongoma N, et al. DNA methylation and expression profiles of placenta and umbilical cord blood reveal the characteristics of gestational diabetes mellitus patients and offspring. Clin Epigenetics 2022;14:69.
63. Del Vecchio G, Li Q, Li W, et al. Cell-free DNA methylation and transcriptomic signature prediction of pregnancies with adverse outcomes. Epigenetics 2021;16:642–61.
64. Pervjakova N, Moen GH, Borges MC, et al. Multi-ancestry genome-wide association study of gestational diabetes mellitus highlights genetic links with type 2 diabetes. Hum Mol Genet 2022;31:3377–91.
65. Elliott A, Walters RK, Pirinen M, et al. Distinct and shared genetic architectures of gestational diabetes mellitus and type 2 diabetes mellitus. Nat Genet 2024;56:377–82.
66. Xiang AH, Kjos SL, Takayanagi M, Trigo E, Buchanan TA. Detailed physiological characterization of the development of type 2 diabetes in Hispanic women with prior gestational diabetes mellitus. Diabetes 2010;59:2625–30.
67. Stasenko M, Cheng YW, McLean T, Jelin AC, Rand L, Caughey AB. Postpartum follow-up for women with gestational diabetes mellitus. Am J Perinatol 2010;27:737–42.
68. Kjos SL, Peters RK, Xiang A, Henry OA, Montoro M, Buchanan TA. Predicting future diabetes in Latino women with gestational diabetes. Utility of early postpartum glucose tolerance testing. Diabetes 1995;44:586–91.

Author and article information

From the Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, Baylor College of Medicine and Texas Children's Hospital, Houston, TX (Drs Barrozo, Racusin, Jochum, Mr Garcia, Drs Suter, Delbeccaro, Ms Shope, Drs Antony, and Aagaard); Medical Scientist Training Program, Baylor College of Medicine, Houston, TX (Mr Garcia); and Genetics & Genomics Graduate Program, Baylor College of Medicine, Houston, TX (Mr Garcia).

Received Dec. 28, 2023; revised May 6, 2024; accepted May 11, 2024.

The authors report no conflict of interest.

This study was supported in part by NIH grants to K.M.A. (R24DK090964, R01NR014792, R01DK089201, and R01HD091731) and the March of Dimes Prematurity Initiative. E.R.B. was supported by a National Science Foundation Postdoctoral Fellowship (Award #2208903) and a Career Development Award from the American Society of Gene & Cell Therapy. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors acknowledge the Submitting Investigators of dbGaP accessions phs001886.v1.p1, phs001886.v2.p1, and phs001886.v3, their primary funding organizations (NICHD and Wayne State University), and dbGaP.

Presented as Oral Abstract #87 at the Society for Maternal Fetal Medicine's 43rd Annual Pregnancy Meeting, Feb. 6th—11th, 2023.

Corresponding author: Kjersti M. Aagaard, MD, PhD. aagaardkm@outlook.com