Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Fetal sex confounds placental transcriptomic signatures of gestational diabetes: a six-cohort audit of public data

Jinhua Tian
Renmin Hospital of Wuhan University
Suting Li
Renmin Hospital of Wuhan University
Zhi Wang
Renmin Hospital of Wuhan University
Hongming Li
Renmin Hospital of Wuhan University
Li Hong
Renmin Hospital of Wuhan University
Songming He
francishsm@qq.com

Renmin Hospital of Wuhan University

Research Article

Keywords: fetal sex, placenta, sexual dimorphism, gestational diabetes mellitus, transcriptomics, confounding, reproducibility

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11583612/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

The placenta is sexually dimorphic: fetal sex shapes placental gene expression, growth and immunity. Public placental transcriptomes are increasingly reused to define gestational diabetes mellitus (GDM) signatures; whether fetal sex can be separated from maternal hyperglycemia and fetal overgrowth in these datasets has not been tested.

Methods

Six public cohorts (bulk and single-cell RNA-seq, spatial transcriptomics, expression array) were reanalyzed under an a priori-defined Plan–Do–Check–Act audit (C1–C17). Fetal sex was inferred per sample from seven Y-linked markers and XIST; sex-linked, copy-number and segmental-duplication genes were excluded; X enrichment was tested rather than assumed; signatures were tested for cross-cohort transfer; compartment mapping used within-sample correlations in four control villi; cis-Mendelian randomization was exploratory.

Results

In the only cohort with a growth contrast, sex separated the groups completely: both cases were female (Y-marker counts 24 and 12; XIST 14,129 and 18,191), whereas the control and the insulin-treated sample were male (Y 29,124 and 26,300). Twelve of the 20 largest uncorrected effects were Y-linked or copy-number artifacts. Under the corrected specification the overgrowth-associated set fell from 2,930 to 1,872 genes, and the growth-discordant set (1,014 mapped genes) was unstable across leave-one-out (1,376 vs 665; Jaccard 0.55–0.70). The overgrowth-associated set showed no X enrichment (1.02×, P = 0.89). Transferability was limited: subtype signatures did not transfer to an independent hyperglycemia cohort (rho − 0.03 to − 0.06) and replicated only weakly on an array (rho 0.105–0.132); the shared-direction set reversed in both external cohorts; mitochondrial-RNA down-regulation tracked preeclampsia (13/13) rather than GDM (8/13). In control villi the overgrowth-associated set localized to cytotrophoblast (r = 0.42) and the growth-discordant set to fibroblast (r = 0.40).

Conclusions

In public placental transcriptomes fetal sex is not a covariate that can be adjusted away: it is confounded with, and in the critical arm identical to, the exposure contrast. GDM signature and sex-difference claims derived from single-sex arms are non-identifiable. Cohorts should report fetal sex and growth per sample, balance sex within arm, and pre-specify sex-aware filters.

Highlights

• In the only public cohort with a fetal-growth contrast, fetal sex separates cases from controls completely: two female-fetus cases versus one male-fetus control.

• Fetal sex is therefore collinear with the exposure; gestational-diabetes-signature and sex-difference claims from single-sex arms are non-identifiable.

• Y-linked and copy-number genes drive 12 of the 20 largest effects yet form only 0.6% of the gene set.

• Gestational-diabetes placental signatures do not transfer across cohorts (rho − 0.03 to − 0.06); mitochondrial-RNA changes track preeclampsia rather than gestational diabetes.

• We recommend reporting fetal sex with marker counts for every sample, balancing sex within arm by design, and pre-specifying sex-aware gene filters.

Background

The placenta is a sexually dimorphic organ. Sex chromosomes, escape from X-chromosome inactivation and sex-steroid signaling shape placental transcript abundance, growth trajectory, vascular resistance and immune signaling, and these differences are detectable in bulk and single-cell placental data [1, 2]. Fetal sex is therefore not a nuisance covariate to be statistically removed; it is part of the biology that placental transcriptomic studies set out to describe.

Gestational diabetes mellitus (GDM), defined as glucose intolerance first recognized during pregnancy, is accompanied by reproducible histological and functional changes in the placenta and by elevated long-term metabolic risk in the offspring. A large literature has therefore sought placental transcriptional markers of GDM exposure, typically by bulk differential expression, co-expression networks or machine-learning classifiers [3, 5, 9, 22]. The paradigm is now saturated: successive studies return largely non-overlapping gene lists, and diagnostic models rarely replicate outside their discovery cohort.

Two structural reasons for this inconsistency are rarely addressed. First, GDM is a composite exposure: maternal hyperglycemia and accelerated fetal growth (large for gestational age, LGA; macrosomia) travel together but are not the same biological exposure [3, 9]. Second, most cohorts do not ascertain fetal sex at sequencing, even though sex is the strongest single axis of placental transcriptional variation after cell composition. A signature derived from such a cohort can reflect any mixture of hyperglycemia, overgrowth, treatment, adiposity, cell composition and fetal sex, and cannot be interpreted causally.

The problem has acquired new urgency because public cohorts are increasingly reused. Secondary analyses chain several Gene Expression Omnibus (GEO) datasets into multi-omic narratives, but the audits that would make such chaining trustworthy — normalization correctness, sample-metadata fidelity, cell-composition confounding and sex balance — are seldom reported [5, 8, 9]. A secondary analysis that does not know fetal sex cannot distinguish a GDM effect from a sex effect, and sex-linked

genes are a well-known source of spurious differential expression when groups differ in sex composition. Reanalysis-only studies thus risk compounding rather than resolving the inconsistency of the primary literature.

We therefore asked three questions. (Q1) In public placental transcriptomes, can fetal sex be separated from the GDM exposure contrast — and how often is it even recorded? (Q2) Do circulating GDM signatures transfer across cohorts once sex-linked and copy-number artifacts are removed? (Q3) What survives a full audit, and what should GDM and sex-difference studies report differently? To answer them we assembled six public cohorts spanning bulk RNA-seq, single-cell RNA-seq, spatial transcriptomics and an expression array, together with genome-wide association and expression-quantitative-trait-locus resources [17–20, 23]. Every module was governed by an a priori-defined Plan–Do–Check–Act chain whose checkpoints (C1–C17) are released with the analysis, including three corrections (normalization units, cohort group assignment, spatial sample composition) that materially changed our own conclusions.

Methods

The analytical workflow is described below step by step; all parameters, decision records and corrections are archived in the quality-control chain (records C1–C17) released with the code.

The full audit trail (records C1–C17) is provided as Additional file 1; supplementary tables are provided as Additional file 2; code and environment details are in Additional file 3; the complete code and QC archive is Additional file 5. This study is reported according to the STROBE statement (Additional file 4).

2.1 Data acquisition and inclusion/exclusion criteria

Six GEO cohorts were used (Table 1; Fig. 1). GSE249311 (bulk RNA-seq, Illumina HiSeq 2500 (GPL11154)) contains 54 placentas annotated as GDMA1 (diet-treated, n = 5), GDMA2 (insulin-treated, n = 9), T2DM (type 2 diabetes mellitus; n = 5), controls (n = 11) and 24 samples with unassigned subtype; only the 30 annotated samples entered subtype-level analyses [3].

Table 1
Public datasets used in this study.

Accession	Platform / design	Samples	Role in this study
GSE249311	Bulk RNA-seq (HiSeq 2500, GPL11154)	54 placentas: GDMA1 (5), GDMA2 (9), T2DM (5), controls (11), unclassified (24)	Discovery set: subtype DEG, mtRNA anchor, PCA (C1/C2/C11)
GSE263483	Bulk RNA-seq (author-annotated count table)	16: 8 GDM vs. 8 controls	Hyperglycemia arm; cross-cohort discordance diagnosis (C6)
GSE267340	scRNA-seq (10x)	48,021 cells profiled (37,448 after QC) from 4 samples: GDMM1 (female fetus), GDMM2 (female), GDMMINS (male), control (male)	Fetal-overgrowth arm (descriptive); single-cell atlas; sex inference
GSE173193	scRNA-seq (sample-level pseudobulk)	8 samples in 4 groups of 2: control, GDM, preeclampsia, advanced age	Directional concordance, group assignment corrected (C15)
GSE222987	Spatial transcriptomics (Visium)	16 samples available; 4 control villous samples (7,041 spots) used	Compartment mapping in control tissue (C16)
GSE70493	Affymetrix HTA-2.0 array (RMA series matrix)	55 independent samples: 30 GDM vs. 25 controls	Second external bulk cohort (C13)

GSE263483 (bulk RNA-seq; 8 GDM v 8 control samples, with an author-annotated gene-level count table) served as the hyperglycemia arm and for the cross-cohort discordance diagnosis. GSE267340 (10x scRNA-seq; GDMM1, GDMM2, GDMMINS and one control) served as the fetal-overgrowth arm and single-cell atlas [7, 24].

GSE173193 (10x scRNA-seq) contains eight placental samples in four groups of two — control, GDM, preeclampsia (PE) and advanced age (GEO series description: “the elderly group”); group assignments were taken exclusively from the GEO sample metadata, and the advanced-age group was kept separate throughout after an initial script error silently merged it into GDM (C15) [5, 6].

GSE222987 (Visium spatial transcriptomics) contains placental samples from healthy, COVID-19 and intrauterine-fetal-death backgrounds; only the four control villous samples (NC1a, NC2–NC4) were used for compartment mapping. GSE70493 (Affymetrix HTA-2.0 array, RMA-normalized series matrix; 30 GDM v 25 controls after averaging four triplet sets) served as a second external bulk cohort [8, 9].

GSE103552 was excluded because it contains diabetic-pregnancy endothelial cells without healthy controls.

2.2 Bulk differential expression and reproduction anchor

Counts were filtered (genes retained with counts ≥ 10 in at least two samples; 20,346 genes) and duplicate gene symbols aggregated by summation (20,331 analyzed); differential expression computed with PyDESeq2 under a ~condition design [12, 13].

Significance was declared at |log2FC| ≥ 0.58 and adjusted P < 0.05. For principal component analysis (PCA), counts were converted to CPM (counts per million, i.e. counts divided by library size), transformed as log2(CPM + 1), and the top 2,000 variable genes were selected; this specification is used consistently in the text and figures.

As a reproduction anchor we used the 13 protein-coding mitochondrial respiratory-chain genes, which an independent reanalysis of the same dataset reported as down-regulated in GDM (log2FC ≈ − 1.4). Our pipeline reproduced the direction for all 39 subtype-by-gene tests (39/39; log2FC − 1.0 to − 1.7). Sensitivity analysis excluding one low-library control (1.6 million reads) preserved the anchor (13/13; mean log2FC − 1.37) with a differentially expressed gene (DEG) Jaccard index of 0.838 (C11) [4].

Because mitochondrial RNA (mtRNA) abundance is sensitive to RNA integrity, rRNA-depletion protocol and library chemistry, the anchor is interpreted as a pipeline-correctness benchmark only, not as a GDM biomarker.

2.3 Cross-cohort external validation and discordance diagnosis

GSE263483 raw counts were extracted from the author-annotated table and reprocessed independently; recomputed log2FC values reproduced the published estimates (Spearman rho = 1.000 over 31,334 genes), confirming faithful extraction. Genome-wide rank correlation (Spearman) between cohort-level log2FC values was assessed with permutation P values (1,000 permutations). Discordance was diagnosed in three steps: author-value verification, mtRNA-direction check, and inspection of candidate genes.

2.4 Two-arm comparison (descriptive) and correction audit

The two-arm comparison contrasted a hyperglycemia arm — GSE263483 GDM v control, with valid inferential statistics — against a fetal-overgrowth arm — GSE267340 GDMM1 + GDMM2 pseudobulk v a single control, which is descriptive only (n = 1 control). The overgrowth arm carries no inferential statistics, and all gene sets derived from it are labeled descriptive.

An initial implementation was rejected during audit (C14) because it (i) normalized counts by a relative size factor without dividing by library size, producing values up to 2.5×10^{12} mislabeled as CPM; (ii) used a + 1 pseudocount on that scale, which is numerically inert and produced |log2FC| up to 31.9 from zero-versus-low counts; and (iii) filtered on a single arm only.

The corrected implementation uses CPM = counts/library-size × 10^6 (column sums = 10^6), a + 1-CPM pseudocount, and a two-arm filter of CPM ≥ 1 and raw count ≥ 10 in both arms (36,591 → 15,575 genes),

intersected with the hyperglycemia differential-expression table (13,940), after exclusion of the 29 artifact genes passing this filter (13,911), and after requiring baseMean ≥ 30 in the hyperglycemia arm (12,401 genes retained).

Y-linked genes (1,507 annotated to chromosome Y in NCBI Gene, of which 88 are present in the pseudobulk matrix; overall, 29 artifact genes passed the two-arm expression filter and were removed), a curated common-copy-number list (e.g. GSTM1, GSTT1, C4A/B, FCGR3B, DEFA3, KIR, AMY) and segmental-duplication families (e.g. NOTCH2NL, TBC1D3, AGAP) were excluded from all gene sets.

Gene sets were labeled shared-direction (hyperglycemia-significant, same direction in the overgrowth arm, |log2FC| ≥ 0.29; n = 439), growth-discordant (hyperglycemia-significant, opposite direction; n = 1,027; names indicate direction, not validated reversal) and overgrowth-associated (not hyperglycemia-significant, |log2FC| ≥ 0.58 in the overgrowth arm; n = 1,872).

Sensitivity analyses comprised threshold variation (|log2FC| 0.29/0.58/1.0), leave-one-out over the two cases, chromosome-composition enrichment (Fisher exact test on chromosome-X membership), and fetal-sex inference from Y-chromosome marker and XIST counts. Gene-set enrichment used Enrichr (GO Biological Process 2023) with Benjamini–Hochberg adjustment [16].

2.5 Fetal sex inference and sex-aware analysis

Fetal sex was inferred per sample from pseudobulk counts summed over seven Y-linked markers (DDX3Y, EIF1AY, KDM5D, RPS4Y1, UTY, USP9Y, ZFY) and XIST. Samples with Y-marker counts below 100 and XIST counts above 5,000 were called female; the observed separation spanned more than two orders of magnitude, so no sample fell near the threshold (Table 2). For the single-cell cohort the same markers were aggregated per sample from cell-level counts and gave the same assignment.

Table 2
Fetal sex inference in the overgrowth cohort (GSE267340).

Sample	Group	Y-marker counts	XIST counts	Inferred fetal sex	In two-arm comparison
GDMM1	GDM, macrosomia	24	14,129	Female	Yes (case)
GDMM2	GDM, macrosomia	12	18,191	Female	Yes (case)
CONTROL	Control	29,124	10,731	Male	Yes (control)
GDMMINS	GDM, macrosomia, insulin-treated	26,300	8,209	Male	No (exploratory only)
Figure legends are provided in the separate file Figure_Legends_BoSD.docx.					

For the four cohorts in which fetal sex is not recorded in the deposited metadata (GSE249311, GSE263483, GSE173193, GSE70493), sex could not be modeled directly. Three safeguards were applied instead: (i) Y-linked genes and curated copy-number/segmental-duplication loci were excluded from

every two-arm contrast; (ii) X-chromosome enrichment of differentially expressed genes was tested rather than assumed; (iii) any arm in which sex was perfectly aligned with case–control status was reported as non-identifiable rather than ‘adjusted’ for sex, because adjustment is impossible when a covariate is collinear with the exposure. Throughout, ‘sex’ denotes fetal sex, the biologically relevant variable in placental tissue; maternal sex is invariant in this setting and sex is not conflated with gender.

2.6 Single-cell preprocessing, integration and annotation

GSE267340 matrices were QC-filtered per cell (gene counts, total counts, mitochondrial fraction; median genes per sample 1,101–1,879; mitochondrial fraction 1.3–2.7%), yielding 37,448 of 48,021 cells. Highly variable genes (n = 3,000; batch key = sample), PCA and Harmony integration were followed by Leiden clustering (resolution 0.6; 28 clusters) and uniform manifold approximation and projection (UMAP) [14, 15].

Automated score-based annotation with canonical placental markers was manually corrected at cluster level, resolving systematic misassignment caused by ambient red-blood-cell RNA; corrected identities include neutrophils (clusters 7/8/18; CSF3R/PADI4/MNDA) and B cells (cluster 23; CD79A/MS4A1).

Fetal sex was inferred per sample as described in 2.5 and matched the pseudobulk assignment (Table 2).

2.7 Independent-cohort directional concordance

GSE173193 samples were aggregated to sample-level pseudobulk (eight samples, four groups), normalized as CPM, and each non-control group compared with controls by log2FC. Concordance between the gene sets of section 2.4 and each cohort's log2FC profile was quantified by Spearman rank correlation.

The initial implementation assigned groups by the first letter of the sample name, which silently merged the two advanced-age samples (GL1/GL2) into GDM; this was corrected against the GEO series metadata (C15), and all concordance statistics reported here use the corrected four-group design.

For GSE70493, probe-to-gene mapping used the Bioconductor hta20transcriptcluster annotation (transcript-cluster to Entrez; 30,300 of 67,528 TC probes mapped) and NCBI Gene for symbols (23,862 genes); four triplet sets were averaged to 55 independent samples. Genome-wide log2FC rank correlations with the GSE249311 subtype signatures used permutation tests (1,000 permutations), and module scores used z-scored marker means with Mann–Whitney U tests and 5,000 permutations (C13).

2.8 Spatial compartment mapping

GSE222987 Visium h5 matrices were read after stripping the GRCh38 vendor prefix from gene names; because the deposit lacks tissue-positions files, array coordinates were decoded from the 10x barcode geometry and validated empirically. Only the four control villous samples were used. Spots whose decidual marker score (IGFBP1, PRL, DKK1, HAND2, LEFTY2) exceeded the villous marker score

(syncytiotrophoblast, cytotrophoblast and extravillous trophoblast markers; SCT, CTB, EVT) were excluded (97–100% retained per sample) [23].

Signature scores were computed per spot with scanpy score_genes for three signatures (S1 shared-direction, S2 growth-discordant, S3 overgrowth-associated; top 50 genes each). Compartment association was quantified as Pearson correlation between signature and compartment marker scores within each sample, pooled across samples by inverse-variance-weighted Fisher z, with between-sample heterogeneity (I^2) reported.

Because spatial autocorrelation was not modeled, no P values are reported for these correlations, and because signatures derive from bulk differential expression their correlation with cell-type markers is partly circular; the mapping is therefore interpreted as expression-compartment attribution in control tissue, not as disease localization.

2.9 Exploratory two-sample Mendelian randomization

Candidate genes comprised the top eight shared-direction and top eight growth-discordant genes by adjusted P; the overgrowth-associated set was excluded from Mendelian randomization (MR) on artifact grounds (C10) — an exclusion that, under the corrected audit, extends to treating the whole overgrowth arm as non-identifiable.

Exposure instruments were significant whole-blood cis-expression quantitative trait loci (eQTLs) from GTEx v8 (effect sizes derived from nes with standard errors recovered from nes and its P value); outcome data were FinnGen R11 GEST_DIABETES (16,802 cases, 237,816 controls). Instruments were harmonized by allele matching with palindromic removal and distance pruning (> 1 Mb, ≤ 5 single-nucleotide polymorphisms (SNPs)); single-SNP Wald and inverse-variance-weighted (IVW) multi-SNP estimates are reported as odds ratios per 1-SD genetically predicted expression [17–20].

Sixteen primary tests allow a Bonferroni threshold of 0.0031; no LD-based clumping or colocalization was performed.

2.10 Quality-control framework

All modules were governed by a Plan–Do–Check–Act (PDCA) cycle: a priori-defined anchors and thresholds (Plan), execution (Do), quantitative checks (Check; records C1–C17), and corrective action (Act). Three corrections materially changed conclusions during this cycle: normalization units and pseudocounts (C14), the GSE173193 group assignment (C15), and the spatial sample composition (C16). The full chain, including the failed v1 analyses, is released for auditability.

Results

Results follow the order of the questions. We first ask whether fetal sex can be separated from the GDM contrast in the one cohort that contains a fetal-growth arm (3.1–3.3), then test whether the remaining

signatures transfer across cohorts (3.4), then place the surviving signals in single-cell and spatial context (3.5–3.6) and report exploratory causal estimates (3.7).

3.1 Fetal sex separates cases from controls in the only cohort with a growth contrast

GSE267340 contains four samples: two GDM–macrosomia cases, one control and one insulin-treated GDM–macrosomia sample. Sex inference was unambiguous (Table 2): the two cases carried 24 and 12 Y-marker counts against 14,129 and 18,191 XIST counts (female fetuses), whereas the sole control carried 29,124 Y-marker counts against 10,731 XIST counts (male fetus); the insulin-treated sample was also male. Fetal sex is therefore not approximately but completely aligned with case–control status in this arm, and fetal sex is absent from the deposited metadata.

Because the two cases are female and the control is male, fetal sex is collinear with disease status: the arm contains no male case and no female control. Any case–control difference is therefore a comparison of a female GDM–macrosomia placenta with a male control placenta, and no statistical adjustment can separate the two.

3.2 Sex-linked and copy-number genes dominate the largest effects

The extreme effect sizes that dominated the uncorrected analysis were artifacts: 12 of the 20 largest |log2FC| values in the overgrowth arm were Y-linked genes or copy-number/segmental-duplication loci (e.g. DDX3Y − 31.9, EIF1AY − 31.5, KDM5D − 29.1, GSTM1 + 25.5 under the erroneous units), and after correction the dynamic range compressed from ± 32 to ± 10 (Fig. 3A).

By membership in the original set obtained under the erroneous units, artifacts were rare (19 of 2,930 genes, 0.6%; 8 Y-linked and 11 copy-number/segmental-duplication), but they concentrated precisely where the eye and the enrichment narrative were drawn.

The overgrowth-associated set was not enriched for chromosome-X genes (1.02×, Fisher P = 0.89), arguing against a pure X-dosage explanation, but sex can act through cell composition and individual effects, and with a one-sample control arm no statistical adjustment can separate them.

3.3 Sex-aware filtering destabilises the overgrowth signature

The corrected two-arm comparison (Fig. 3) retained 12,401 genes after two-arm filtering and artifact exclusion.

Even after this filtering, any case-versus-control contrast in this arm remains non-identifiable. The growth-discordant set, in contrast to the overgrowth-associated set, showed a mild chromosome-X

excess (51 of 1,027 (5.0%) vs 3.3% in the remaining analyzed genes; 1.50×, Fisher P = 0.0034), consistent with residual sex-related composition rather than a purely autosomal program.

The gene sets are also unstable. Leave-one-out over the two cases gave 1,376 versus 665 growth-discordant genes (Jaccard 0.70 and 0.55 against the pooled set), and set sizes varied three-fold across |log2FC| thresholds 0.29/0.58/1.0 (growth-discordant: 1,397/1,027/558). The genome-wide direction-concordance rate was 45.0% (binomial $P = 1.1\times10^{-28}$ against 50%) — a statistically detectable but small bias that cannot be attributed to biology given the confounding above.

We therefore report the two-arm comparison as descriptive and hypothesis-generating only.

3.4 Signature transferability and what survives correction

Thirty analyzable GSE249311 samples (GDMA1 n = 5, GDMA2 n = 9, T2DM n = 5, control n = 11) yielded 3,950 DEGs for GDMA1 (1,368 up / 2,582 down), 1,306 for GDMA2 (666 / 640) and 3,717 for T2DM (1,586 / 2,131). Principal-component analysis on log2(CPM + 1) of the top 2,000 variable genes separated diabetic from control samples along PC1 (30.8%) (Fig. 2).

All 39 subtype-by-gene mtRNA tests were concordantly down-regulated (log2FC − 1.0 to − 1.7), validating the pipeline against an independent reanalysis of the same data. This anchor did not extend to the hyperglycemia arm (see below), and is interpreted as a correctness benchmark rather than a disease finding.

Contrary to the assumption underlying most secondary analyses, the GSE249311 subtype signature did not transfer to GSE263483: genome-wide log2FC rank correlations were − 0.03 to − 0.06 (permutation P = 0.001), and genes significant in both cohorts agreed in direction only 38–46% of the time. Both pipelines were independently verified, excluding a technical explanation.

The discordance was not explained by pipeline implementation, but it cannot be attributed to biology rather than protocol or RNA quality with these data: the 13 mitochondrial RNAs that are strongly down-regulated in GSE249311 (mean log2FC ≈ − 1.4) show no signal in GSE263483 (mean − 0.01; 8/13 down).

In the corrected GSE173193 analysis the same pattern held — GDM − 0.01 (8/13), whereas preeclampsia showed 13/13 down-regulation (mean − 0.27) and the advanced-age group showed a positive mean (+ 0.47; 1/13 down).

Mitochondrial-RNA down-regulation is therefore not a general GDM feature; it is cohort-associated and better interpreted as a preeclampsia-associated or mixed-cohort phenomenon.

In the second external platform (GSE70493, 55 independent samples), subtype signatures retained only weak directional consistency (genome-wide rho + 0.132 / +0.107 / +0.105 for A1/A2/T2DM; permutation $P \le 0.001$), with no genes passing conventional DE thresholds ($|\log2FC| \ge 0.58$), and the shared-direction immune module was again down-regulated in GDM ($\Delta = -0.107$, permutation P = 0.0078), consistent with GSE173193 [9].

We conclude that cohort composition — fetal-growth outcome, treatment, ancestry, gestational age and protocol — dominates the reported signatures, and that the widely circulated “GDM placental signature” is not a portable entity.

Three observations survived the full audit. First, the growth-discordant set (n = 1,014 mapped genes) followed the hyperglycemia-arm direction in the corrected GSE173193 GDM contrast (rho = + 0.331, P = 2.5×10^{-27}) — stronger than the uncorrected estimate (rho = + 0.230), whose dilution was partly caused by the mis-assigned advanced-age samples.

Because the overgrowth arm itself is non-identifiable, this concordance supports directional repeatability of the hyperglycemia-associated extracellular matrix (ECM) program across cohorts, not a macrosomia-specific reversal. In the same data, preeclampsia showed the opposite relationship (rho = − 0.189), underscoring that GDM and PE placental transcription are directionally distinct.

Second, enrichment findings were recomputed from the corrected sets (the v1 manuscript reported enrichments for which no results file existed): the overgrowth-associated set for regulation of cell migration (adjusted P = 3.7×10^{-7}) and receptor tyrosine kinase (RTK) signaling. The growth-discordant set was enriched for extracellular-matrix organization (2.7×10^{-9}) and collagen-fibril organization (1.3×10^{-8}).

The shared-direction set was enriched for defense response to virus (1.6×10^{-5}) and type-II-interferon response (2.9×10^{-3}) — an antiviral/interferon signature rather than a generic immune-activation program.

The previously reported DNA-replication and axon-guidance enrichments did not reproduce and are withdrawn.

Third, the shared-direction set failed directional replication in both external cohorts (GSE173193 GDM rho = − 0.247, PE rho = − 0.499; GSE70493 Δ = −0.107, P = 0.0078) and the growth-discordant module was neutral in GSE70493 (Δ = −0.057, P = 0.13), which contains no fetal-growth annotation. These negative results are reported because they bound the claims that public data can support.

3.5 Single-cell atlas of the overgrowth cohort

After QC and Harmony integration, 37,448 cells resolved into 14 curated types, dominated by neutrophils (8,176 plus an interferon-stimulated gene (ISG)-high subset of 1,818), T cells (6,728), cytotrophoblast (4,554), syncytiotrophoblast (4,375), natural killer (NK) cells (3,787) and Hofbauer cells (3,051) (Fig. 4).

Sex inference from Y-marker and XIST counts identified the two GDMM cases as female fetuses and both male-annotated samples (GDMMINS, control) as male; the single-cell data recapitulate the pseudobulk inference. Composition differences are reported descriptively because group and sample are confounded: GDMMINS showed 60.9% neutrophils versus 3.4% in the control, fibroblasts were depleted in GDMM (1.9% versus 10.3%), and SCT proportions differed (14.3% versus 9.9%) [10, 11].

Any of these differences can generate bulk-level differential expression and none can be attributed to disease given n = 4.

3.6 Compartment mapping in control placental tissue

Projection onto the four control villous samples (7,041 spots after decidual-marker filtering) located the overgrowth-associated signature in cytotrophoblast markers (pooled r = + 0.42; positive in 4/4 samples, + 0.30 to + 0.54), the growth-discordant signature in fibroblast markers (r = + 0.40; 4/4 positive, + 0.11 to + 0.69; $I^2 \approx 100\%$), and the shared-direction signature weakly in Hofbauer markers (r = + 0.20) with anti-correlation to syncytiotrophoblast (r = − 0.25) (Fig. 5A) [21].

Notably, the overgrowth-associated signature was not endothelial (r = + 0.10, $I^2 = 0\%$): our earlier analysis had reported endothelial localization (r = + 0.37), which pooled all 16 samples — including eight COVID-19 and two intrauterine-fetal-death samples — across spots; the corrected within-sample analysis shows this to be a pooling artifact.

The fibroblast localization of the growth-discordant set is also a caution: it is equally consistent with the fibroblast-composition difference observed in section 3.5 as with a disease program.

3.7 Exploratory Mendelian randomization

Of 16 candidate genes, 8 had serviceable whole-blood cis-eQTL instruments. Genetically predicted IL1R2 expression was associated with higher GDM risk (OR 1.81, 95% CI 1.02–3.20; Wald P = 0.044; single SNP), and CRABP2 showed a protective point estimate (OR 0.93, P = 0.088); no estimate survived Bonferroni correction applied to the 16 planned tests (threshold 0.0031).

These are exploratory, hypothesis-generating estimates: they derive from whole blood rather than placenta, use single-SNP instruments without LD clumping or colocalization, and are reported here only to bound what the current evidence supports — a candidate for future work, not a causal claim.

Discussion

The principal findings of this reanalysis are negative and methodological, and they are more consequential for study design than any gene list we could have produced.

4.1 Principal findings

First, currently circulating GDM placental signatures are cohort-composition dependent. A signature that reproduces internally (39/39 mtRNA anchor) fails to transfer to a phenotypically adjacent cohort (rho − 0.03 to − 0.06), retains only weak directional consistency on an independent platform (rho ≈ 0.11–0.13 with no threshold-level DEGs), and its immune module reverses in both external cohorts.

Secondary analyses that chain public cohorts should therefore demonstrate transferability explicitly rather than assume it.

Second, the hyperglycemia-versus-overgrowth decomposition is not identifiable with existing public data. The only cohort annotated for fetal overgrowth pairs two female-fetus cases with one male-fetus control; after removing the artifact-driven extremes its gene sets are unstable (leave-one-out Jaccard 0.55–0.70), and its headline enrichments, although partly reproducible (cell migration), cannot be attributed to overgrowth rather than to sex or composition.

We initially reported this decomposition as three separable programs; the corrected audit withdraws that claim.

Third, mitochondrial-RNA down-regulation — an anchor we and others used as a GDM feature — tracks preeclampsia (13/13 down) rather than GDM (8/13, mean − 0.01) in group-corrected data, and is absent in the hyperglycemia arm. Given the technical sensitivity of mtRNA abundance, this finding is best read as a methodological warning: mtRNA signals in placental bulk transcriptomics are more plausibly cohort-composition or protocol artifacts than GDM biology.

4.2 Fetal sex as a biological variable and as a confounder

Placental sex differences are real biology, and the argument of this paper is not that they should be discarded [1, 2]. The argument is that they cannot be separated from a disease contrast after the fact when the cohort was not designed to separate them. In the overgrowth arm analyzed here, every case was a female fetus and the only control was male; a ‘GDM signature’ computed from that comparison is equally a sex signature, and any ‘sex difference’ computed from it is equally a disease effect. No covariate adjustment can resolve this, because when sex is collinear with exposure there is no independent variation left to adjust with. The same logic applies to archival cohorts in which fetal sex was never recorded: absence of the variable is not evidence that sex is balanced.

Three practices would prevent most of the ambiguity documented here. First, report fetal sex for every sample, together with the inference method and the marker counts, and deposit it in the series metadata rather than in a supplementary table. Second, balance sex within arm by design — ideally a 2 × 2 (sex × exposure) arrangement, which also makes sex-by-exposure interaction testable; fetal sex is inexpensive to record and is often already known to the recruiting clinician. Third, pre-specify sex-aware gene filters (Y-linked genes, X-inactivation escapees, copy-number and segmental-duplication loci) and state how many genes they remove; in this dataset those loci were 0.6% of the overgrowth-associated set yet accounted for 12 of the 20 largest effects, the pattern most likely to be believed and pursued.

We also caution against over-reading the X chromosome. The growth-discordant set showed a modest X excess (51 of 1,027, 5.0%, versus 3.3% in the remaining analyzed genes; 1.50×, Fisher P = 0.0034), whereas the overgrowth-associated set did not (1.02×, P = 0.89). With two cases, neither estimate supports a dosage story, and X enrichment in a signature derived from bulk tissue may equally reflect cell composition. Finally, sex differences may themselves be modified by the maternal environment, including hyperglycemia; testing that interaction requires balanced designs and considerably larger

samples than are currently deposited, which is an argument for prospective collection rather than further reuse of existing cohorts.

4.3 Why signatures diverge

Cohorts differ in fetal-growth outcome, treatment, ancestry, gestational age, tissue sampling and library protocol; each is a plausible driver of the observed discordance, and none is consistently annotated. The shared-direction set's antiviral/interferon character further suggests that cohort-specific immune contexts (including subclinical inflammation) contribute to what earlier studies interpreted as GDM immune activation.

Rather than adding another signature to the pile, we suggest that future primary studies pre-register the exposure decomposition — hyperglycemia, overgrowth, treatment — and report the negative results of transferability tests alongside any positive findings.

4.4 Methodological lessons from the audit

Four audit findings generalize beyond this dataset. (i) Normalization: a relative size factor must be divided into library size; failure here produced 10^{12}-scale “CPM” values and ± 32 log2FC artifacts that survived an expression filter designed to prevent exactly that.

1. Metadata: group assignment must come from GEO characteristics fields, not sample-name prefixes; our script silently merged an advanced-age group into GDM and changed a headline statistic (rho + 0.230 → +0.331 after correction); that is, the misclassification had attenuated rather than inflated the concordance.
2. Pseudobulk with single-sample arms: a one-control arm produces effect sizes whose sign and magnitude are dominated by that individual; such arms should be labeled non-identifiable. (iv) Spatial pooling: concatenating spots across disease backgrounds and tissue compartments created an endothelial localization that within-sample analysis does not support; compartment mapping should be computed within sample and reported with heterogeneity.

4.5 Design recommendations

To make the hyperglycemia/overgrowth question answerable, future cohorts should be designed as a 2×2 design (GDM-AGA, GDM-LGA, non-GDM-AGA, non-GDM-LGA; AGA, appropriate for gestational age) with sex-balanced enrollment and n ≥ 10 per cell; cell-composition deconvolution should be reported and used as a covariate; fetal sex and growth outcome should be mandatory metadata fields; and spatial projections should use disease-matched controls with within-sample inference.

Until such data exist, the transferability failures documented here are the most informative result the public domain currently supports.

4.6 Limitations

This study has the limitations it documents. The overgrowth arm has two cases and one control with complete sex confounding; all derived sets are descriptive, and the growth-discordant set carries a mild chromosome-X excess (1.50×, Fisher P = 0.0034) that further precludes an overgrowth-specific interpretation.

The spatial data contain no GDM cases, so compartment mapping reflects control tissue only, is partly circular (bulk-derived signatures include cell-type markers), and does not model spatial autocorrelation.

The MR analysis uses whole-blood eQTLs, single-SNP instruments, no LD clumping and no colocalization, and its estimates are nominal. GSE173193 has two samples per group and unknown fetal-growth outcomes, so its concordance statistics are directional descriptors rather than confirmatory tests. GSE70493 effects are weak (no threshold-level DEGs) and its platform lacks mitochondrial-probe annotation, so the mtRNA anchor could not be tested there.

The GDMA2 signal in the discovery cohort is weak (1,306 DEGs at $|\log2FC| \geq 0.58$ and 383 at $|\log2FC| \geq 1.0$, adjusted $P < 0.05$). Finally, the AJOG qPCR validation of the discovery cohort showed direction disagreements between the original RNA-seq and its own qPCR follow-up for CSH1, PER1 and IL2RB (digitized from Fig. 1C of that study), reinforcing the cohort-level direction heterogeneity documented here.

Conclusions

In public placental transcriptomes, fetal sex is not a covariate that can be adjusted away: it is confounded with, and in the critical arm identical to, the exposure contrast. GDM placental signatures do not transfer across cohorts; the hyperglycemia-versus-overgrowth decomposition is not identifiable with existing data because fetal sex is a complete confounder; and mitochondrial-RNA down-regulation is better interpreted as a preeclampsia-associated or technical feature than as GDM biology.

What the field needs is not another reanalysis of the same cohorts but prospective cohorts that report fetal growth outcome and fetal sex per sample, balance sex within arm, treat single-sex or single-sample pseudobulk arms as non-identifiable, and publish their transferability failures.

Abbreviations

AGA
appropriate for gestational age
CPM
counts per million
CTB
cytotrophoblast
DEG
differentially expressed gene

eQTL
expression quantitative trait locus
EVT
extravillous trophoblast
GDM
gestational diabetes mellitus
GEO
Gene Expression Omnibus
GWAS
genome-wide association study
IVW
inverse-variance weighted
LFC
log2 fold change
LGA
large for gestational age
MR
Mendelian randomization
mtRNA
mitochondrial RNA
PE
preeclampsia
SCT
syncytiotrophoblast
SD
segmental duplication
T2DM
type 2 diabetes mellitus
UMAP
uniform manifold approximation and projection
XIST
X-inactive specific transcript.

Declarations

Ethics approval and consent to participate

Not applicable. This study used only publicly available, de-identified data; no new human participants or human material were involved.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

Author Contribution

S.H. and J.T. conceptualized the study. Z.W. and S.H. established the methodology. J.T. and S.H. performed software programming. J.T., H.L. and S.L. completed validation. H.L., S.H. and J.T. performed formal analysis. J.T. and S.L. conducted the investigation and data curation. J.T. and S.L. wrote the original draft. J.T. reviewed and edited the manuscript. Z.W. and L.H. created visualizations. S.H. supervised the research. S.L. obtained funding. J.T. and S.H. managed project administration. All authors read and approved the final manuscript.

Acknowledgement

The authors thank the investigators of the original cohorts (GSE249311, GSE263483, GSE267340, GSE173193, GSE222987 and GSE70493) and the GTEx and FinnGen consortia for making their data publicly available.

Data Availability

https://doi.org/10.5281/zenodo.23857573

References

1. Gabory A, Roseboom TJ, Moore T, Moore LG, Junien C. Placental contribution to the origins of sexual dimorphism in health and diseases: sex chromosomes and epigenetics. Biol Sex Differ. 2013;4:5. 10.1186/2042-6410-4-5.
2. Rosenfeld CS. Sex-specific placental responses in fetal development. Endocrinology. 2015;156(10):3422–34. 10.1210/en.2015-1227.
3. Barrozo ER, Racusin DA, Jochum MD, Garcia BT, Suter MA, Delbeccaro M et al. Discrete placental gene expression signatures accompany diabetic disease classifications during pregnancy. Am J Obstet Gynecol. 2025;232(3):326.e1–326.e15. 10.1016/j.ajog.2024.05.014. (GSE249311).

4. Zhang X, Tang W, Guo L, Bao H, Chen X. Gestational diabetes mellitus placentas exhibit mitonuclear transcriptional imbalance co-occurring with m6A regulatory gene dysregulation: insights from bulk RNA sequencing. Research Square preprint, posted 2026-07-22. 10.21203/rs.3.rs-10060346/v1. (anchor source; peer-reviewed publication pending).
5. Yang Y, Guo F, Peng Y, Chen R, Zhou W, Wang H, et al. Transcriptomic profiling of human placenta in gestational diabetes mellitus at the single-cell level. Front Endocrinol (Lausanne). 2021;12:679582. 10.3389/fendo.2021.679582. (GSE173193).
6. Zhou W, Wang H, Yang Y, Guo F, Yu B, Su Z. Trophoblast cell subtypes and dysfunction in the placenta of individuals with preeclampsia revealed by single-cell RNA sequencing. Mol Cells. 2022;45(5):317–28. 10.14348/molcells.2021.0211.
7. GEO accession GSE263483 (no accompanying primary publication identified); secondary analysis: Liu S, Feng J, Ding Y, Liang W. Placental transcriptomics identifies a candidate HLA-DQA2-FGL2 immune signature in gestational diabetes mellitus. Placenta. 2026;183:82–93.
8. Barrozo ER, Seferovic MD, Castro ECC, Major AM, Moorshead DN, Jochum MD, et al. SARS-CoV-2 niches in human placenta revealed by spatial transcriptomics. Med (N Y). 2023;4(9):612–e6344. 10.1016/j.medj.2023.06.003. (GSE222987).
9. Binder AM, LaRocca J, Lesseur C, Marsit CJ, Michels KB. Epigenome-wide and transcriptome-wide analyses reveal gestational diabetes is associated with alterations in the human leukocyte antigen complex. Clin Epigenetics. 2015;7:79. 10.1186/s13148-015-0116-y. (GSE70493).
10. Wong SL, Demers M, Martinod K, Gallant M, Wang Y, Goldfine AB, et al. Diabetes primes neutrophils to undergo NETosis, which impairs wound healing. Nat Med. 2015;21(7):815–9. 10.1038/nm.3887.
11. Vokalova L, van Breda SV, Ye XL, Huhn EA, Than NG, Hasler P, et al. Excessive neutrophil activity in gestational diabetes mellitus: could it contribute to the development of preeclampsia? Front Endocrinol (Lausanne). 2018;9:542. 10.3389/fendo.2018.00542.
12. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550. 10.1186/s13059-014-0550-8.
13. Muzellec B, Teleńczuk M, Cabeli V, Andreux M. PyDESeq2: inference of differential gene expression using bulk and single-cell RNA-sequence data. Bioinformatics. 2023;39(9):btad547. 10.1093/bioinformatics/btad547.
14. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods. 2019;16:1289–96. 10.1038/s41592-019-0619-0.
15. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 2018;19:15. 10.1186/s13059-017-1382-0.
16. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR. Ma'ayan A. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. 2013;14:128. 10.1186/1471-2105-14-128.

17. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science. 2020;369:1318–30. 10.1126/science.aaz1776.
18. Kurki MI, Karjalainen J, Palta P, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. Nature. 2023;613(7944):508–18. 10.1038/s41586-022-05473-8.
19. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. eLife. 2018;7:e34408. 10.7554/eLife.34408.
20. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet. 2014;23(R1):R89–98. 10.1093/hmg/ddu328.
21. Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, et al. Single-cell reconstruction of the early maternal-fetal interface in humans. Nature. 2018;563:347–53. 10.1038/s41586-018-0698-6.
22. Barker DJP. The developmental origins of adult disease. J Am Coll Nutr. 2004;23(6):S588–95. 10.1080/07315724.2004.10719428.
23. 10x Genomics. Visium Spatial Gene Expression — barcode structure, image alignment and tissue positions documentation. https://www.10xgenomics.com/support/software/space-ranger/latest/analysis/outputs/spatial-outputs. Accessed 30 Oct 2026.
24. Gene Expression Omnibus. GSE267340: single-cell RNA sequencing of placenta in patients with gestational diabetes during term pregnancy. National Center for Biotechnology Information. 2023. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE267340. Accessed 30 Oct 2026.

Figures

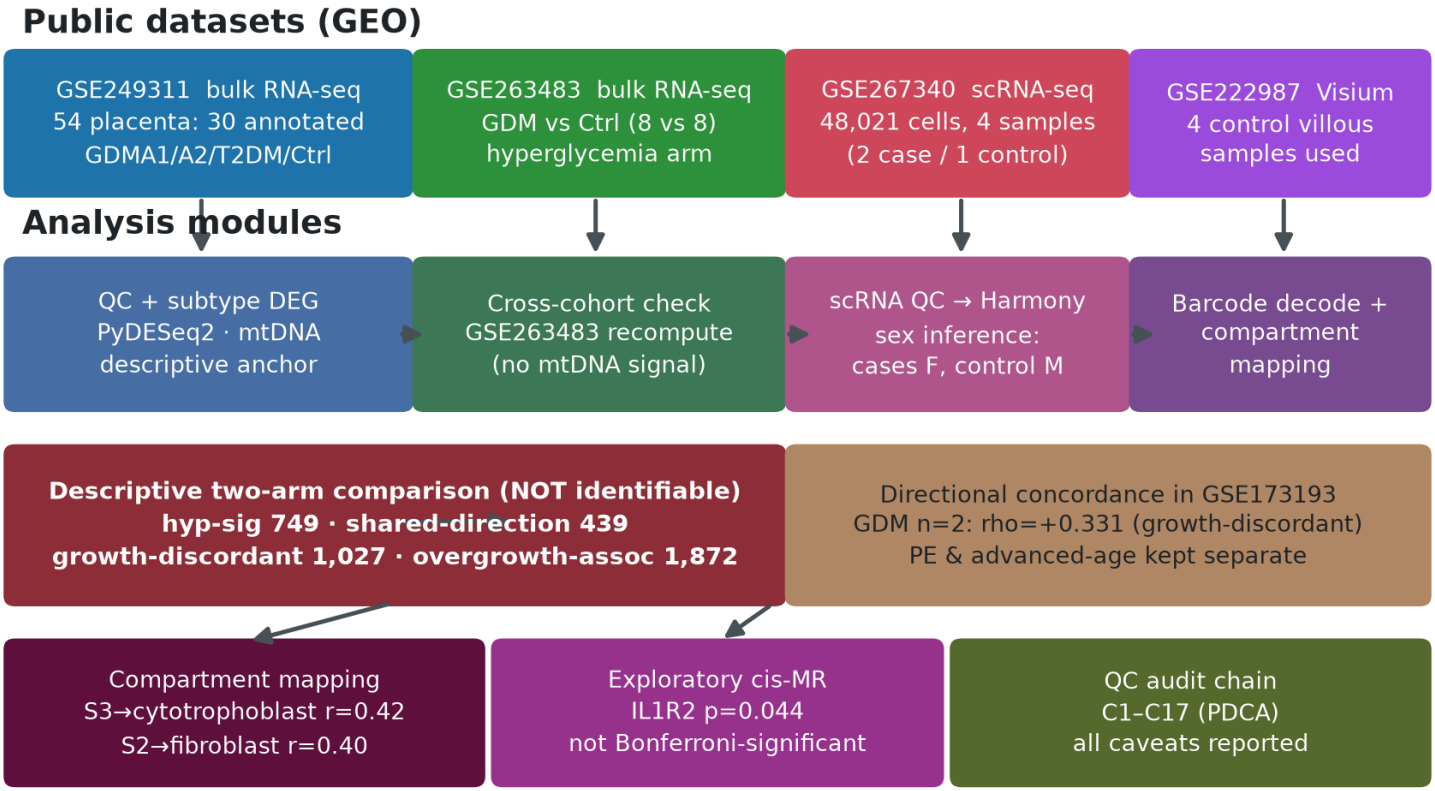


Figure 1

Legend not included with this version.

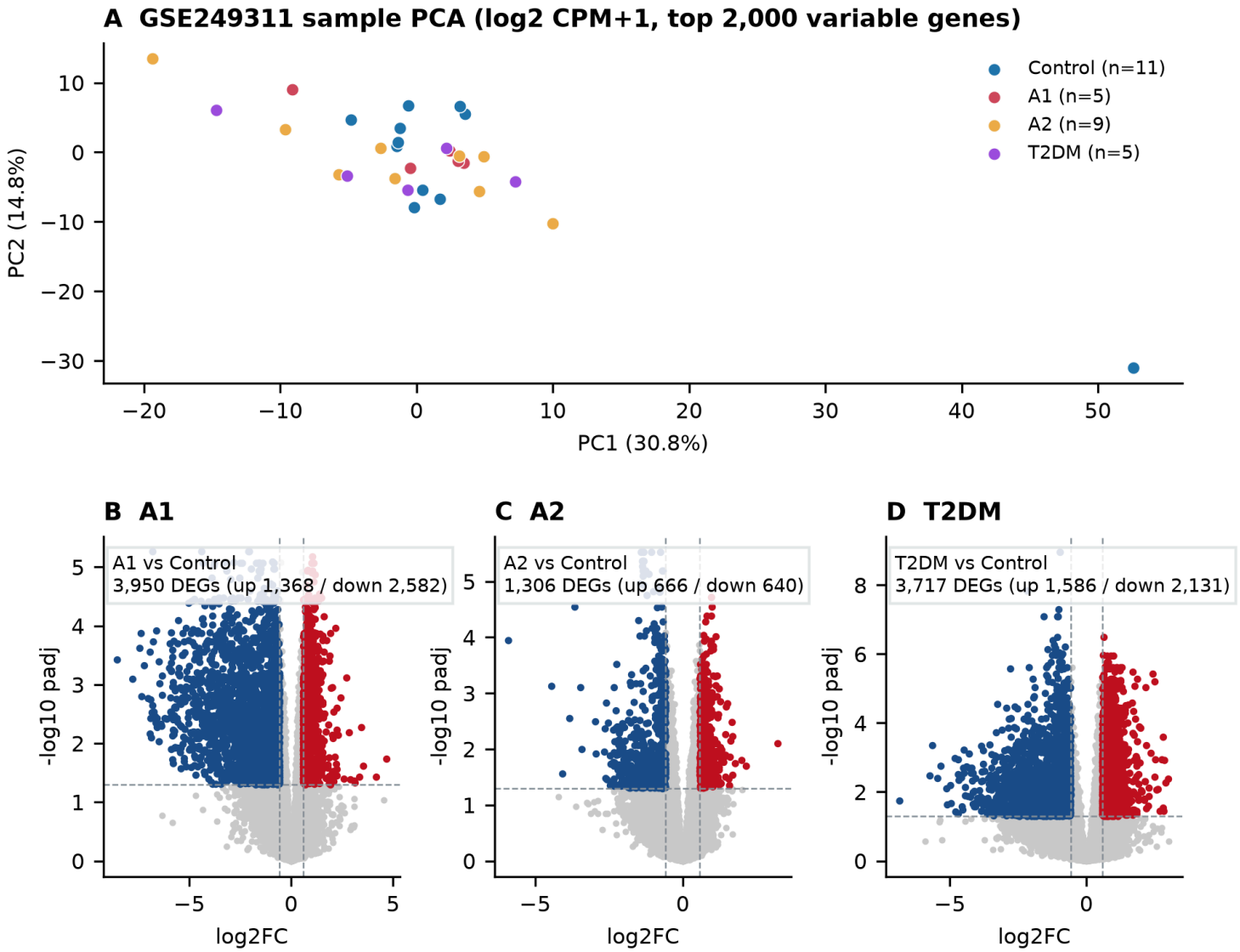


Figure 2

Legend not included with this version.

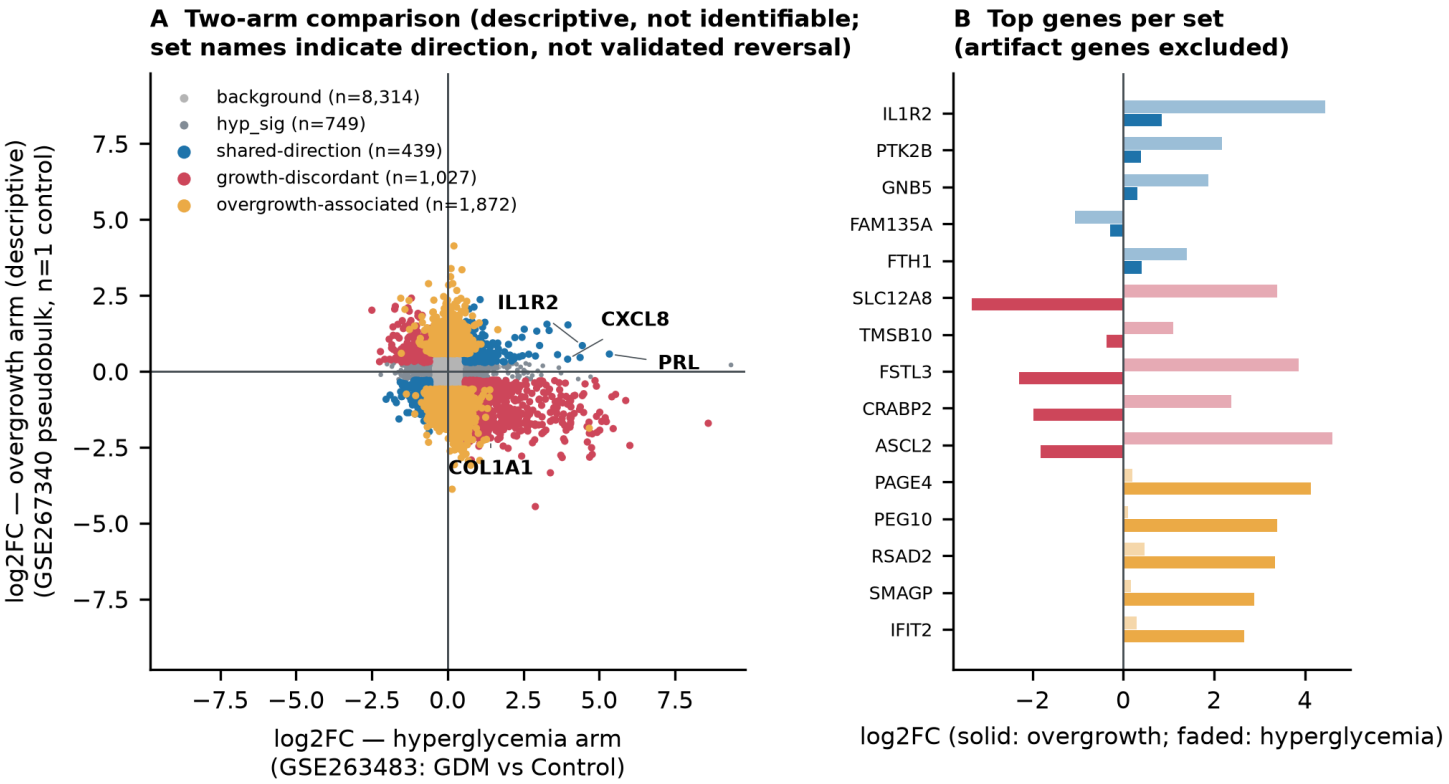


Figure 3

Legend not included with this version.

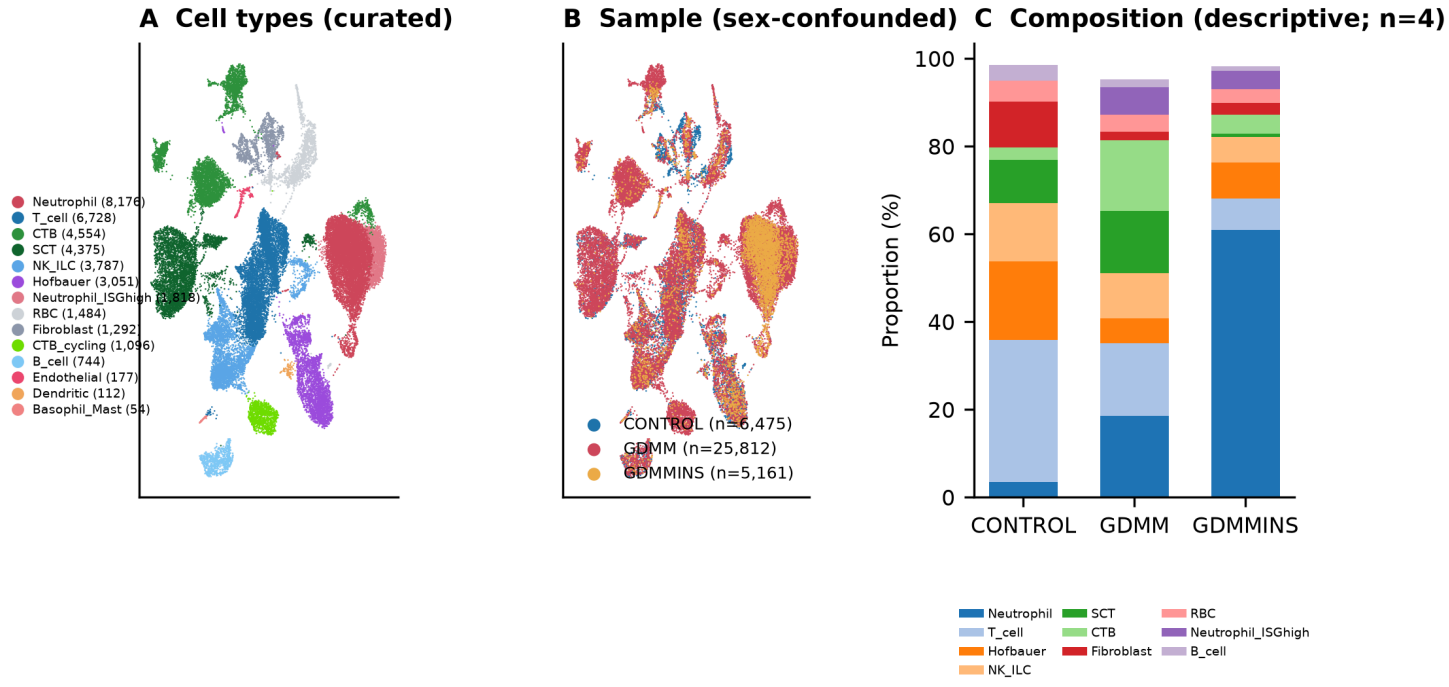


Figure 4

Legend not included with this version.

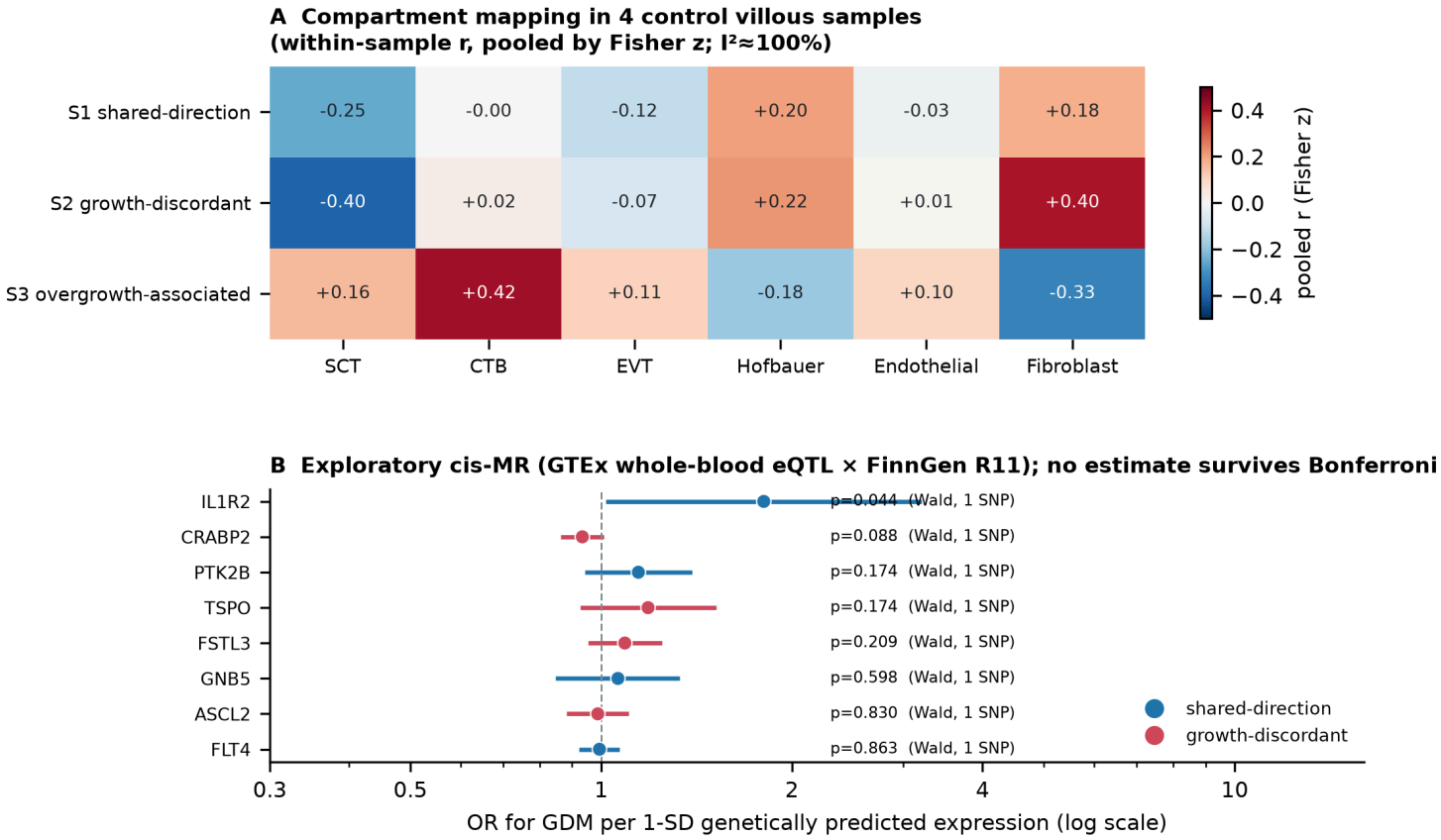


Figure 5

Legend not included with this version.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile3CodeandEnvironmentREADME.txt
- Additionalfile4STROBEchecklist.docx
- Additionalfile1QCrecordsC1C17.docx
- Additionalfile5CodeandQCarchive.zip
- Additionalfile2SupplementaryTablesS1S9.xlsx
- GraphicalAbstractBoSD.png