ORIGINAL ARTICLE

Analysis of Differentially Expressed Genes and Molecular Pathways in Spontaneous Preterm Birth: A Bioinformatics Approach

Noor Haliza Mohamed Ibrahim[1,2], Zurina Zainudin[3], Amilia Afzan Mohd Jamil[4], Norshariza Nordin[1,2,5], Habibah Abdul Hamid[4], Karuppiah Thilakavathy[1,2,5]

[1] Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.
[2] Genetics & Regenerative Medicine Research Group, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.
[3] Department of Paediatrics, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.
[4] Department of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.
[5] Malaysian Research Institute on Ageing (MyAgeing), Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

ABSTRACT

Introduction: Preterm birth (PTB) accounts for 35% of neonatal mortality worldwide within 28 days of life. Limited understanding of its molecular mechanisms hampers the development of effective preventive strategies. Thus, this study aims to investigate dysregulated genes in the maternal blood and placenta of PTB women to unveil potential underlying mechanisms. **Materials and Methods**: A meta-analysis of differentially expressed genes (DEGs) in PTB was conducted using the limma R package and p-value combination method. Seven maternal blood and four placenta studies was chosen for this analysis. Accordingly, protein-protein interaction networks were constructed using STRING and visualised in Cytoscape. Network centrality was assessed using cytoHubba and CytoNCA tools in Cytoscape to rank key genes. MCODE was then used to identify highly interconnected gene clusters. Functional enrichment analysis was performed to obtain significant ontologies and pathways. **Results**: This study discovered 20 blood and 18 placenta hub genes, with IKBKG, HSP90AA1, and EGR1 being the most significantly enriched and central in maternal blood, and IGF1, FN1, and LOX genes for placenta. These DEGs mainly involve cytokine-mediated signalling pathways (maternal blood) and extracellular matrix organisation (placenta) biological pathways. Comparably, KEGG revealed them to be enriched in pathways of IL-17 signalling and ECM-receptor interaction, respectively. **Conclusion**: These DEGs have the potential to precisely identify PTB-risk individuals by revealing the underlying pathophysiology, allowing for more targeted and effective interventions. Notably, maternal blood DEGs are of particular interest, as they may serve as early, non-invasive biomarkers for PTB risk, enabling earlier detection and preventive care during pregnancy.
Malaysian Journal of Medicine and Health Sciences (2026) 22(3): 1-10.doi:10.47836/mjmhs.v22i3.1578

Keywords: Premature Birth, Gene Expression Profiling, Biomarkers, Inflammation, Extracellular Matrix

Corresponding Author:
Karuppiah Thilakavathy, PhD
Email: thilathy@upm.edu.my
Tel: +603-97692652

INTRODUCTION

Preterm birth (PTB), i.e., initiation of labour before 37 gestational weeks, is one of the major contributors to perinatal mortality and morbidity. Approximately 35% of preterm infants succumb within 28 days of life, while the survivors suffer through physical and neurodevelopmental impairments such as cerebral palsy, intellectual disabilities, and impairments in vision and hearing (1). In the US, caring for these infants costs an estimated $26 billion per year (2), posing a socioeconomic challenge. However, the precise molecular triggers for PTB remain elusive, presenting a significant challenge to preventive initiatives (1). In addressing this challenge, gene studies, particularly differentially expressed gene (DEG) analyses, offer a promising avenue for predicting health outcomes and unravelling the underlying molecular mechanisms of PTB, as evidenced by studies conducted by (3,4). These target genes could be utilised as biomarker in prevention of PTB.

The Preterm Birth International Collaborative has catalogued 116 potential biomarkers for PTB across diverse samples (5,6). However, their efforts encountered a challenge in identifying a singular molecule that reliably predicts PTB, especially in asymptomatic women. Furthermore, existing gene regulation studies have proposed various genes as potential screening targets for PTB, yet a definitive and singular effective biomarker remains elusive (3,7). These inconsistencies may be due to several factors, including platform heterogeneity, variations in sample processing and handling, cohort size, ethnicity, and differences in analytical methods or study quality. Additionally, DEG studies often utilise different tissue types such as maternal peripheral blood, foetal-derived cord blood, decidual or placental tissue, which also influences gene expression profiles, adding biological complexity to the interpretation of results (5). To address these challenges, an integrated bioinformatics approach that combines public datasets could significantly improve the accuracy and reliability of analyses by reducing dataset-specific bias and increasing statistical power, potentially revealing key genes implicated in spontaneous PTB (sPTB).

Our objective is to consolidate studies from the Gene Expression Omnibus of the National Center for Biotechnology Information (GEO NCBI) (https://www.ncbi.nlm.nih.gov/geo/), focusing on gene regulation in the placenta and maternal blood in sPTB compared to term birth (TB). It's widely recognised that gene expression varies between tissues or cell types during the early stages of disease (8). Given the heterogeneity of PTB and its complex aetiology, combining results from both maternal blood and placenta enhances the ability to detect shared and tissue-specific molecular signatures. Therefore, identifying gene regulation across different tissues is crucial for understanding PTB's causes. The identification of significant DEGs and their associated biological functions may lead to more precise and reliable therapeutic targets for early PTB diagnosis and improved treatment efficacy.

MATERIALS AND METHODS

Data collection

This study conducted a cross-study meta-analysis using publicly available independent datasets to pinpoint DEGs in two distinct conditions: (i) mothers experiencing sPTB and (ii) mothers undergoing TB. We accessed PTB datasets from the GEO NCBI database and selected studies meeting specific criteria outlined in Table I.

Table I: Inclusion and exclusion criteria of the gene regulation studies.

Inclusion criteria	Aspect	Exclusion criteria
Focuses on coding or messenger RNAs	**Type of dataset**	Focuses on other than messenger RNA such as long non-coding RNA
Microarray and high-throughput sequencing	**Type of dataset generation**	Other than microarray or high-throughput sequencing
Total number of samples is more or equal to 10	**Sample size**	Total number of samples is less than 10
Mothers of sPTB and term birth (TB) (control)	**Study population**	Other than sPTB or TB
Include samples collected before or after delivery	**Samples**	No samples were collected or studied on
Either maternal blood, placenta, or umbilical cord blood	**Type of tissue utilised in the study**	Tissues other than stated tissues
On the causality of sPTB	**Study focus**	Focusing on other medical conditions, such as obesity, intrauterine growth restriction, etc.

Note: RNA – Ribonucleic acid; sPTB – spontaneous preterm birth; TB – term birth

Identification of DEGs

DEG analysis was performed separately for maternal blood and placental samples, according to the steps depicted in Fig. 1. The coding utilised in this study is provided in Supplementary 1. Since datasets were derived from different sequencing platforms and studies, each dataset was analysed independently to control for within-dataset variability. For RNA-seq data, we transformed raw data into counts per million (CPM) to normalise for library size differences (9). Subsequently, quality control checks were conducted which included removing genes with negative average log2 CPM using edgeR's aveLogCPM function and focusing on genes with an average count of five or more per sample. A multi-dimensional scaling (MDS) plot was generated to assess sample differences and the impact of specific conditions on DEG identification. We accounted for variation-causing factors as covariates (9) and eliminated composition bias between libraries using the trimmed mean of M-values (TMM) method, through calcNormFactors function in edgeR (10). Limma-voom (11), a modified version of limma-trend (12), was employed for RNA-seq data analysis to discern differences between PTB and TB conditions. Outliers within studies were managed using the arrayWeights function from the limma package (13). Finally, DEGs were screened based on the criteria: (i) p-value < 0.05 and (ii) log2FoldChange |logFC| ≥ 1. This two-fold change criterion was selected as it represents a commonly accepted threshold for biologically meaningful expression differences and helps prioritise genes with potentially stronger functional impact. A more stringent cutoff was not used to avoid missing subtle but important regulatory changes, given the complex and heterogeneous nature of PTB.

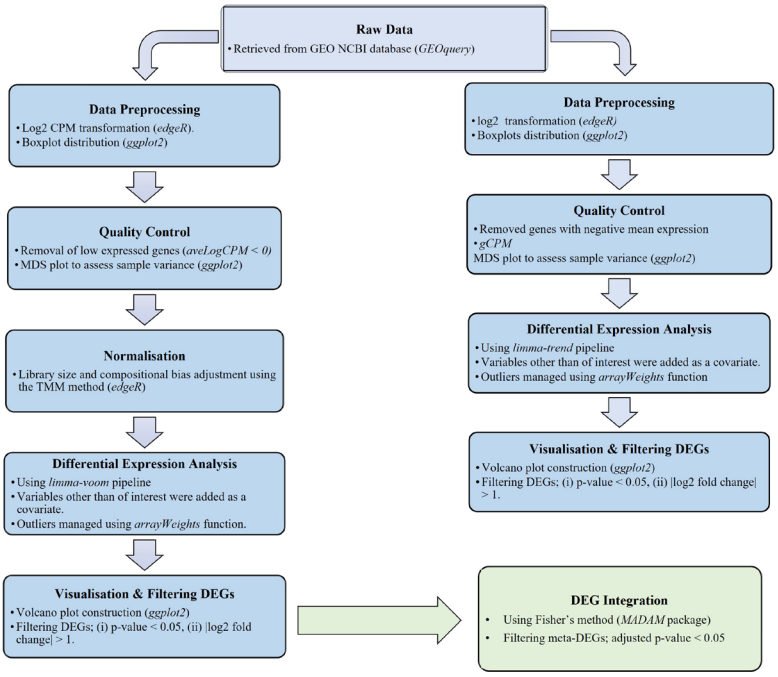

Fig. 1: Method framework of DEG analysis used in RNA-seq and microarray studies.

For microarray studies, we normalised the data by converting the data into a log2 scale. Genes with a negative mean expression value were filtered out. Next, we separately analysed the expression level differences between cases and control samples for each group (PTB vs. TB) using the limma R package (12), in line with the previously mentioned steps.

Batch effects between studies were not corrected directly. Instead, to integrate findings across datasets and minimise cross-platform and inter-study batch effect concerns, we applied Fisher's combined probability test on the filtered DEG lists. This approach allows for the identification of consistently significant genes across multiple datasets without merging raw expression matrices, which could introduce artificial variability due to platform-specific differences.

Integration of DEGs

Meta-analyses of the studies were performed using Fisher's p-value combination method which combines the p-values from multiple hypothesis tests into a single overall test statistic. This step was performed using the fisher.method function in the MADAM package (14). The approach, as described by Fisher, combines the p-values into a statistic;

$$S = -2 \sum^k \log p$$

which follows a χ^2 distribution with 2k degrees of freedom.

Genes with a false discovery rate (FDR)-adjusted p-value < 0.05 were chosen for subsequent analysis and denoted as 'meta-DEGs'. The meta-logFC for these genes was computed by obtaining the mean of repeated fold change values in the gene expression dataset (15). Both p-value combination and meta-logFC calculation were carried out independently for maternal blood and placenta samples.

Functional Enrichment Analysis (FEA)

FEA was performed to identify overrepresented Gene Ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) terms associated with maternal blood and placental meta-DEGs, for both upregulated and downregulated genes. The enrichment analysis was conducted through the Enrichr tool (https://maayanlab.cloud/Enrichr/), which provides the biological process (BP), (ii) cellular component (CC), and (iii) molecular function (MF) GOs and significant KEGG biological signalling pathways. Significant overrepresentation of GO terms and biological pathways in the gene list was determined by a p-value < 0.05.

Protein-protein interaction (PPI) network analysis

We created PPI networks for the placenta and maternal blood datasets using the Search Tool for the Retrieval of Interacting Genes (STRING; https://string-db.org/), with a confidence score ≥ 0.4 and visualised them in Cytoscape (16). To identify highly interconnected nodes within the network, we applied several Cytoscape plug-ins.

Firstly, cytoHubba plug-in, which ranks nodes based on their network features, was utilised to identify the key nodes of the network (17). Four approaches of cytoHubba; maximal clique centrality (MCC), edge percolated component (EPC), maximum neighbourhood component (MNC), and Stress, was chosen and the top 100 nodes from each approach were obtained.

Next, CytoNCA plug-in, which also ranks a network's nodes based on the nodes' centrality was utilised (18). The top 100 ranked genes from 4 selected approaches of CytoNCA, i.e., (i) betweenness centrality (BC), (ii) closeness centrality (LC), (iii) degree centrality (DC), and (iv) eigenvector centrality (EC), were acquired. Nodes that were consistently identified by all cytoHubba and CytoNCA approaches were chosen for further analysis.

Finally, the Molecular Complex Detection (MCODE), plug-in was applied to extract densely connected modules from the PPI network, with a degree cut-off = 2, node score cut-off = 0.2, K-score = 5, and max depth = 100 (19). The key nodes identified by both cytoHubba and CytoNCA were then compared with the MCODE clusters to pinpoint highly interconnected genes sharing similar biological pathways. The resulting list of specific nodes were analysed further as hub genes.

Verification of the hub genes

The potential role of hub genes in PTB was assessed using GENEVESTIGATOR (https://nebion.com/gv/start/start.jsp), a web-based software designed for molecular expression meta-analysis and data mining (20). Specifically, we utilised the Signature tool within the Similarity Search toolset, enabling the comparison of a research gene signature data with in-store available gene signature from previous experiments. The Signature tool correlates the entered signature; the logFC values of the genes, with previously curated expression data and identifies conditions that portray a similar gene signature as our data. The results are obtained based on Pearson correlation (PC) values, where a higher PC value indicates greater similarity. We extracted the top 50 conditions with similar DEG magnitudes to infer the potential impact of the hub genes in various biological conditions.

Receiver operating characteristics (ROC) analysis

The diagnostic accuracy of the DEGs was assessed through the ROC curves, generated using the software Epitools (https://epitools.ausvet.com.au/roccurves). Log10 transformed expression data of the targeted genes were inserted in the software. The overall diagnostic accuracy was assessed through the Area under the ROC curve (AUC) where a value of 0 indicates high inaccuracy,

and a value of 1 indicates high accuracy (21). In general, an AUC of 0.5 suggests no discrimination in the ability to diagnose patients with and without the condition, 0.7 to 0.8 is considered acceptable, 0.8 to 0.9 is considered excellent, and > 0.9 is considered outstanding.

RESULTS

Identification of DEGs between maternal blood and placenta samples

A total of 24 datasets were selected from the NCBI-GEO database where only ten datasets were included for further analysis. Five of the datasets were excluded as they focused on non-coding RNAs, while another nine of the datasets were removed as they involved sample types that did not meet the specific criteria of this study, such as myometrium and cervix. Supplementary 2 shows the list of selected datasets for the DEG analysis. A total of seven datasets comprising maternal blood samples and four datasets comprising placenta samples were identified in the GEO NCBI. Certain datasets, such as GSE150167 and GSE120480, also investigated conditions like preterm premature rupture of membrane (PPROM) or elective caesarean delivery, alongside sPTB. To ensure a more precise focus on sPTB, samples associated with these additional conditions were excluded from the DEG analysis.

The background correction and normalisation of the studies were analysed through box plots (Supplementary 3). Each box plot displayed a comparable median expression level across samples, indicating that correction and normalisation processes had been carried out effectively. An MDS plot was drawn to illustrate the spatial distribution of the samples (Supplementary 4), and the batch correction was conducted accordingly. The total DEGs obtained (p-value < 0.05, |logFC > 1|) were visualised through a volcano plot (Supplementary 5) and categorised in Table II according to the sample utilised in the study.

Table II: The number of DEGs obtained, up and downregulated genes, for each study sorted according to the sample utilised. (CONT.)

Maternal Blood				
No	GEO Accession	No. of DEGs		
		Total	Upregulated	Downregulated
1	GSE28387	9	4	5
2	GSE46510	2	2	0
3	GSE73685	1	0	1
4	GSE96083	1226	757	469
5	GSE142974	6	3	3
6	GSE148402	70	40	30
7	GSE150167	2	1	1

CONTINUE

Table II: The number of DEGs obtained, up and downregulated genes, for each study sorted according to the sample utilised.

Placenta				
No	GEO Accession	No. of DEGs		
		Total	Upregulated	Downregulated
1	GSE73685	67	57	10
2	GSE118442	433	134	299
3	GSE120480	299	207	92
4	GSE174415	36	4	32

Note: DEG – Differentially Expressed Genes; GEO – Gene Expression Omnibus

Meta-analysis of DEGs

To increase the reliability of the DEGs, we integrated the DEGs (referred as meta-DEGs) through p-value combination in both tissues, respectively. We found a total of 1113 maternal blood meta-DEGs; 455 downregulated and 658 upregulated genes, and 768 placenta meta-DEGs; 397 downregulated and 371 upregulated genes. Interestingly, this study also identified 32 common meta-DEG in this study, expressed in both maternal blood and placenta (Supplementary 6). Among these, 19 genes, including GNG11, CD36, SH3BGRL, LRRN3, TYROBP, S100A12, CD33, LYZ, VSIG4, CCL2, MMRN1, TDRD9, S100A8, HLA-A, SLC4A1, OLFM1, CYP1B1, AHSP, and CD177, exhibited contrasting regulation between the two tissues. The remaining thirteen genes consistently showed either downregulation (LYVE1, LTF, OLFM4, and CLEC12A) or upregulation (NEXN, FLNA, ATP884, ALAS2, CE51, OSBP2, PPBP, SIGLEC11, and ITGA2B) in both tissues.

FEA of meta-DEGs

FEA of meta-DEGs in maternal blood

The gene ontologies for both up and downregulated genes in maternal blood were found to be associated with similar BPs (refer to Fig. 2. A and B, respectively). The genes were significantly implicated in biological processes such as neutrophil degranulation (GO:0043312) and neutrophil activation involved in immune response (GO:0002283), suggesting that gene dysregulation is mainly attributable to infection or inflammation. Comparably, upregulated blood genes are involved in KEGG pathways such as Shigellosis and Salmonella infection (Fig. 2. C), while downregulated genes are involved in T cell receptor and VEGF signalling pathway (Fig. 2. D), which are critical pathways in infections and cell survival. In terms of molecular functions, upregulated genes are enriched in protease binding (GO:0002020) and RNA binding (GO:0003723), while downregulated genes are associated with RNA (GO:0003723) and protein kinase binding (GO:0019901). For cellular components, upregulated genes are linked to secretory granule membrane (GO:0030667) and focal adhesion (GO:0005925), whereas downregulated genes are associated with intracellular membrane-bounded organelle (GO:0043231) and the nucleus (GO:0006834).

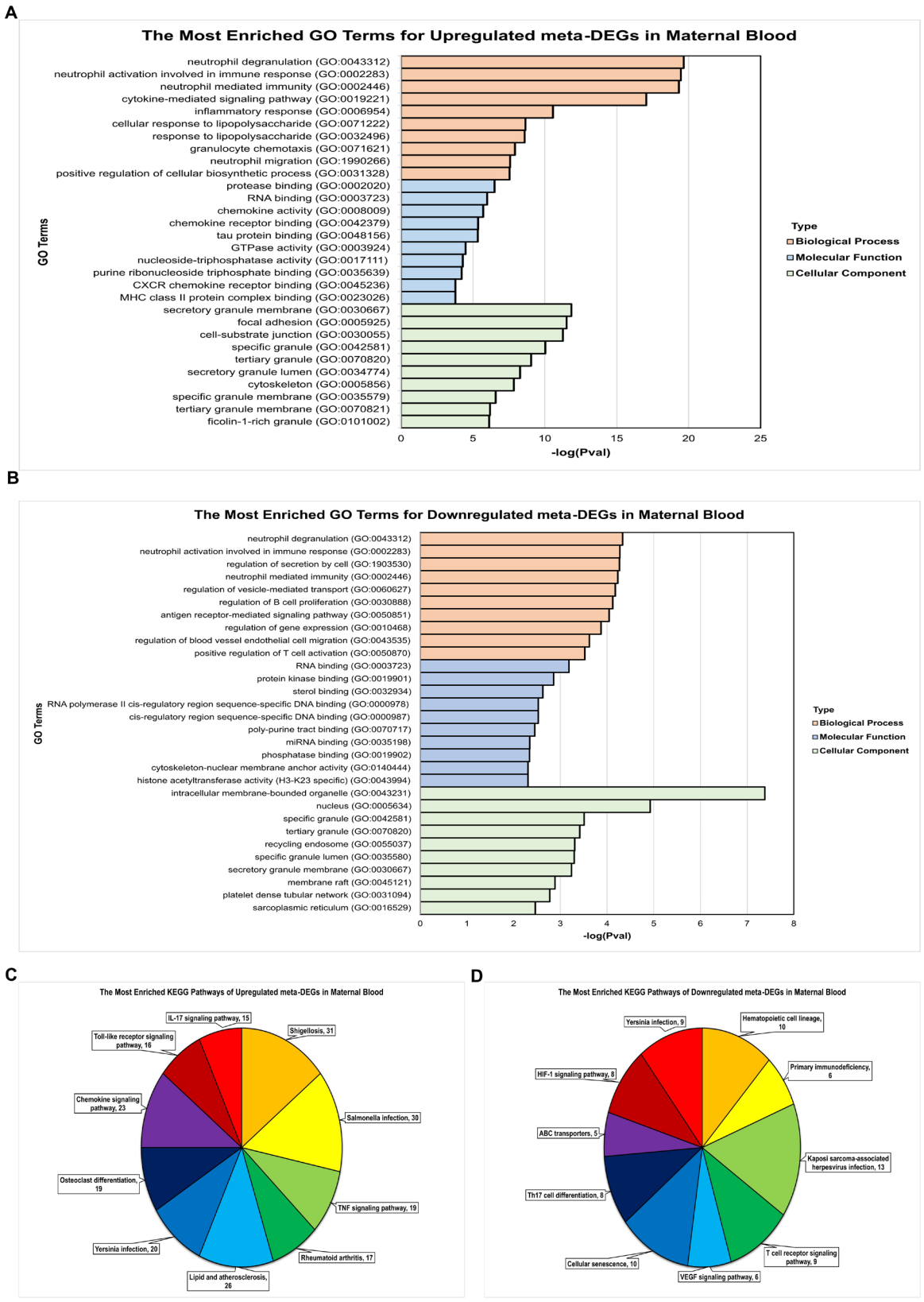

Fig. 2: GO enrichment analysis of the maternal blood meta-DEGs. This figure shows the top ten most significantly ($p < 0.05$) enriched GO terms in biological process, molecular function and, cellular component for the (A) upregulated meta-DEGs, (B) downregulated meta-DEGs, while the top ten most significantly ($p < 0.05$) enriched KEGG pathways is showed at (C) upregulated meta-DEGs and (D) downregulated meta-DEGs. All the statistically significant values of the terms were negative 10-base log-transformed. Abbreviations: (i) DEGs - differentially expressed genes, (ii) GO - gene ontology.

FEA of meta-DEGs in placenta

Similar to the FEA of maternal blood, the gene ontologies for up and downregulated genes in placenta are comparable (refer to Fig. 3. A and B, respectively). Both sets of genes are enriched in biological processes related to extracellular structure (GO:0043062) and extracellular matrix (ECM) organization (GO:0030198). This finding aligns with the KEGG analysis, which highlights ECM-receptor interaction pathways (Fig. 3. C and D). Moreover, the meta-DEGs are also enriched in molecular functions of platelet-derived growth factor (GO:0048407) and heme binding (GO:0020037) for upregulated genes, while for downregulated genes are more enriched in receptor ligand activity (GO:0048018) and BMP receptor binding (GO:0070700). For cellular components, the upregulated meta-DEGs are enriched in collagen-containing extracellular matrix (GO:0062023) and endoplasmic reticulum lumen (GO:0005788), while the downregulated meta-DEGs are enriched in the integral component of plasma membrane (GO:0005887) and the collagen-containing extracellular matrix (GO:0062023).

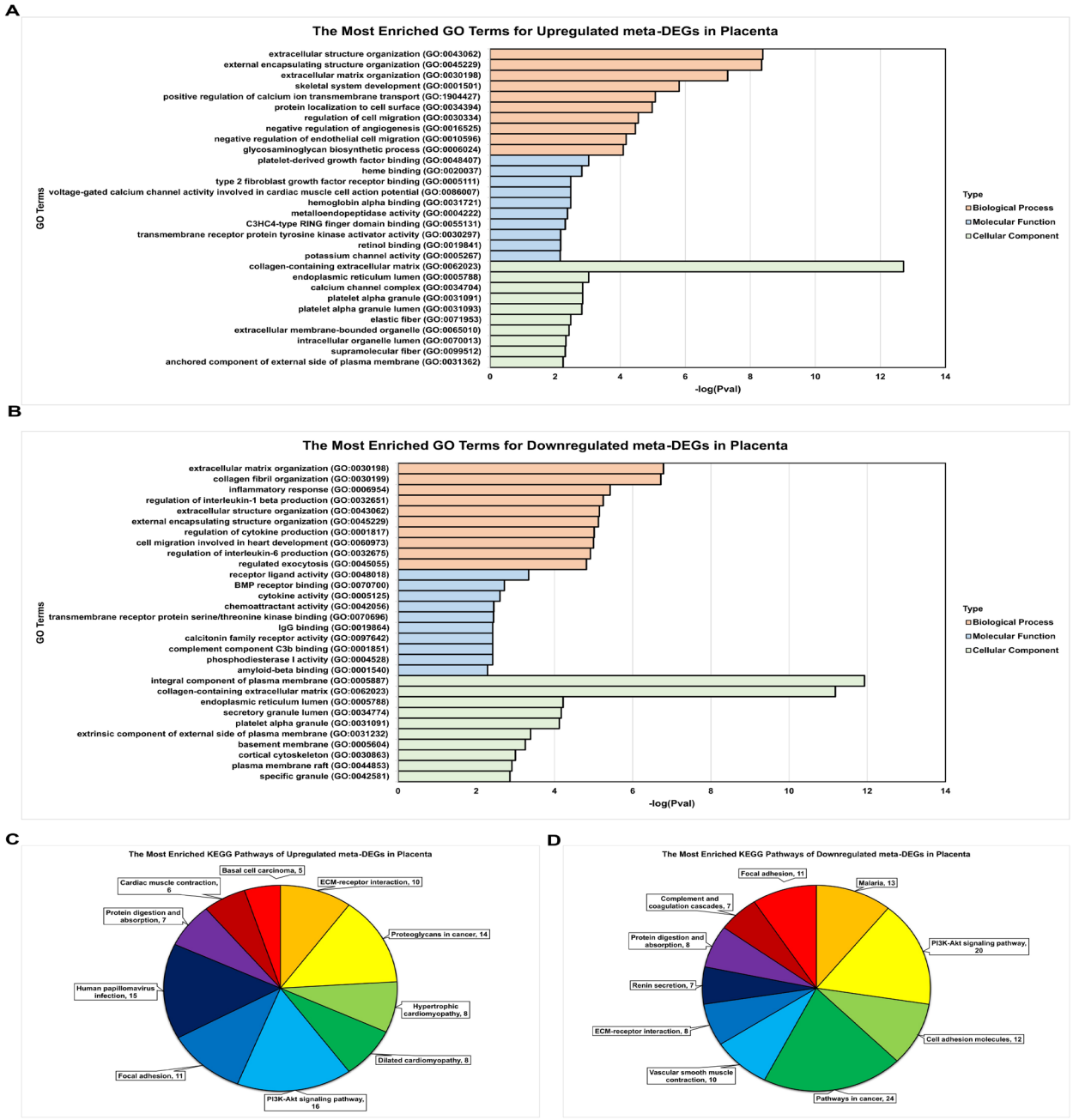

Fig. 3: GO enrichment analysis of the placental meta-DEGs. This figure shows the top ten most significantly ($p < 0.05$) enriched GO terms in biological process, molecular function and, cellular component for the (A) upregulated meta-DEGs, (B) downregulated meta-DEGs, while the top ten most significantly ($p < 0.05$) enriched KEGG pathways is showed at (C) upregulated meta-DEGs and (D) downregulated meta-DEGs. All the statistically significant values of the terms were negative 10-base log-transformed. Abbreviations: (i) DEGs - differentially expressed genes, (ii) GO - gene ontology.

FEA of common meta-DEGs among maternal blood and placenta

Meanwhile, GO analysis of the common meta-DEGs found between blood and placenta, for both similar or different magnitudes of expression (Supplementary 7), showed them to be significantly involved in biological processes such as neutrophil degranulation (GO:0043312) and neutrophil mediated immunity (GO:0002446). This result coincides with the GO analysis of blood meta-DEGs. Common meta-DEGs with different magnitudes of expression were found to be enriched in KEGG pathways of malaria and cytomegalovirus infection, while those with expressions of a similar magnitude were found to be enriched in KEGG pathways of glycine, serine, and threonine metabolism.

PPI network construction and cluster analysis

PPI networks were constructed for both maternal blood and placenta meta-DEGs. We ran cytoHubba and CytoNCA to obtain the network's most influential nodes or edges. The top 100 nodes ranked from cytoHubba (EPC, MCC, MNC, and Stress) were compared with the top 100 nodes obtained from cytoNCA (DC, EC, BC, and CC). Overlapping analysis revealed 29 key blood (Supplementary 8) and 46 key placenta genes (Supplementary 9).

Furthermore, MCODE was applied to identify clusters in each PPI network, mainly to screen densely connected regions that may represent molecular complexes and be part of a pathway (31). In the maternal blood PPI network, the main cluster formed had a score of 15.194 and comprised of 68 nodes and 509 edges, while the main placenta PPI network cluster (score = 14.667) had 28 nodes and 198 edges (Fig. 4). Key genes from previous analyses were compared with the MCODE clusters to identify interconnected genes in biological pathways. This step resulted in the selection of 20 maternal blood and 18 placenta hub genes. In the MCODE cluster, the three most centrally located maternal blood genes with high cluster scores are IKBKG (12.27), HSP90AA1 (12.55), and EGR1 (12.68). Meanwhile, IGF1 (12.97), FN1 (12.66), and LOX (12.66), were highly interconnected for the placenta. The cluster score of the hub genes and their regulation can be seen in Supplementary 10.

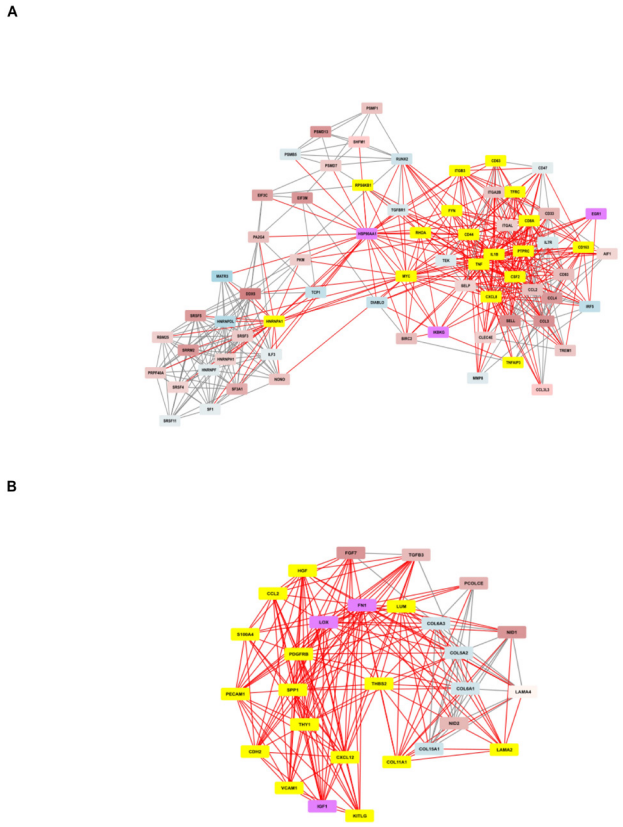

Fig. 4: MCODE clusters. (A) MCODE cluster for maternal blood PPI network and (B) MCODE cluster for placenta PPI network. The nodes colored in yellow are the key genes obtained after the cytoNCA and CytoHubba comparison. The nodes colored in purple are the highly interconnected genes with high cluster scores among the cluster.

Verification of hub genes

The signature tool of GENEVESTIGATOR was utilised to provide the top 50 studies with similar expression patterns as to the 20 maternal blood and 18 placenta gene signatures; refer to Supplementary 11 for maternal blood and Supplementary 12 for placental key genes. The studies are ranked based on their relative similarity, PC values (indicated on the right). For maternal blood key genes, the studies or conditions that share similar gene expression signatures are (i) irritable bowel syndrome (IBS) (PC: 2.998), (ii) lung cancer (PC: 2.786), (iii) Parkinson's disease (PC: 2.751), and (iv) ovary cancer (PC: 2.388), among others. On the other hand, placenta key genes share similar expression patterns

with studies of (i) endometrial cancer (PC: 3.899), (ii) lung adenocarcinoma (PC: 3.573), (iii) ovary cancer (PC: 3.040), and (iv) IBS (PC: 2.097), among others.

ROC curve analysis

ROC curve analyses were conducted for individual hub genes to determine the relative accuracy of each candidate for discrimination across sPTB and TB groups. The diagnostic performance of IKBKG, HSP90AA1, and EGR-1 in maternal blood and IGF1, FN1 and LOX in placental samples was evaluated using their respective AUC, sensitivity, and specificity values (Table III). In maternal blood, EGR1 showed the most promising diagnostic potential with an AUC of 0.81 (95% CI: 0.644–0.976), indicating good discriminative ability. It demonstrated perfect specificity (100%), suggesting its strong potential in correctly identifying non-cases, although the sensitivity was moderate at 58.3%. HSP90AA1 also showed acceptable diagnostic value (AUC = 0.737), with very high specificity (94.7%) but lower sensitivity (46.2%), indicating it may be more useful in ruling out false positives. In contrast, IKBKG had the lowest AUC of 0.6, with balanced but relatively modest sensitivity (63.6%) and specificity (57.9%), suggesting limited diagnostic reliability.

Table III: AUC, Sensitivity, and Specificity of Hub Genes in Maternal Blood and Placental Tissue Samples

Tissue Type	Hub Genes	AUC (95% CI)	Sensitivity (%)	Specificity (%)
Maternal Blood	*IKBKG*	0.6 (0.385 – 0.816)	63.6	57.9
	HSP90AA1	0.737 (0.557 – 0.917)	46.2	94.7
	EGR-1	0.81 (0.644 – 0.976)	58.3	100.0
Placenta	*IGF1*	0.75 (0.485 – 1.000)	87.5	66.7
	FN1	0.505 (0.285 – 0.724)	92.9	20.0
	LOX	0.762 (0.575 – 0.95)	93.3	12.5

Note: AUC – The area under the receiver operating characteristics curve

In the placental tissue, IGF1 and LOX exhibited good diagnostic utility, with AUCs of 0.75 and 0.762, respectively. IGF1 had a strong sensitivity (87.5%) and moderate specificity (66.7%), while LOX showed very high sensitivity (93.3%) but poor specificity (12.5%), indicating it may be prone to false positives. FN1, however, showed an AUC of 0.505 and a very low specificity (20.0%) despite a high sensitivity (92.9%), making it unsuitable as a standalone diagnostic marker.

DISCUSSION

In this study, we aimed to address the perplexing mechanisms underlying the initiation of PTB. Given the limited number of large-scale studies and the small sample sizes in existing research, we implemented a meta-analysis approach to increase the sample size and identify reliable genes associated with spontaneous PTB. Our primary objective was to distinguish the gene expression signatures in two different tissues, maternal blood, and placenta, to better understand their roles in PTB progression. This study identified a PTB DEG signature comprising 1113 maternal blood and 768 placenta DEGs, with 32 genes commonly expressed. This signature, derived from multiple studies encompassing various risk factors, can serve as a valuable reference for a more diverse population of women experiencing spontaneous PTB, transcending any single dataset.

Remarkably, GENEVESTIGATOR showed that both maternal blood and placental meta-DEG in this study sharing similar expression signature as other known PTB-inducing conditions; IBS (22), ovarian and endometrial cancer (23). Thus, these findings underscore the relevance and potential clinical implications of the identified meta-DEGs in understanding and addressing PTB risk factors.

DEGs in maternal blood

The maternal blood meta-DEGs are primarily associated with immunity processes such as neutrophil degranulation. In cases of infection, the local inflammatory response and pro-inflammatory cytokine production is triggered, increasing the production of circulating neutrophils (24). Simultaneously, neutrophils emit reactive oxygen species (ROS) and more cytokines into the environment, which in turn, initiates the synthesis of labour mediators such as prostaglandins and matrix metalloproteinases (MMPs) (25). This aligns with the identified enriched KEGG pathways, including bacterial infections such as shigellosis and salmonellosis, known to induce PTB (26, 27), as well as pathways of TNF or T cell receptor signalling pathway that produces neutrophils (28, 29). In this study, 20 hub maternal DEGs were identified from network clustering analysis, with IKBKG, HSP90AA1, and EGR1 emerging as key upregulated genes. However, their precise mechanistic roles in sPTB remain incompletely characterised.

IKBKG

The Inhibitor of nuclear factor kappa B kinase subunit gamma (IKBKG) is responsible for the expression of the I-kappa B Kinase (IKK) protein complex, that regulates the activity of nuclear factor-kappa-B (NF-κB). Under normal conditions, NF-κB remains inactive while bound to the IKK complex (24). Upon specific signals of infection or inflammation, the IKK complex releases NF-κB into the cytoplasm, facilitating its translocation to the nucleus, where it binds to κB motifs in the promoters of NF-κB-regulated genes. These motifs are prevalent in genes encoding pro-inflammatory mediators such as IL-1B, IL-6, TNFa and IL-8 (30).

Implantation initially requires a highly inflammatory environment, necessitating NF-κB activation. Subsequent foetal development requires an anti-inflammatory phenotype, prompting the downregulation of the NF-κB activity (30). Toward the end of pregnancy, this balance shifts back to a pro-inflammatory phenotype, marked by increased IL-8 and COX2 expression, leading to contractions and rupture of foetal membranes. PTB or

even abortions are promoted when the profile change occurs prematurely. High levels of IKBKG and persistent NF-κB expression indirectly contribute to the initiation of human labour (24). However, the regulation of this gene during pregnancy or labour has not been clearly defined in previous studies.

HSP90AA1

Heat Shock Protein 90 Alpha Family Class A Member 1 (HSP90AA1) plays a crucial role in the oestrogen signalling pathway, essential for myometrium maturation and the initiation of the contractile response that causes parturition at term (31). In the cytoplasm, oestrogen receptors require binding to HSP90 proteins to maintain its conformation until they interact with 17β-oestradiol (32). Upon binding with this ligand, a conformational change is induced, leading to receptor translocation to the nucleus and binding with oestrogen response elements, thereby regulating the expression of target genes. Increased expression of HSP90 proteins, potentially may lead to PTB, through disrupting oestrogen signalling and also by activating the immune response, evident during stress (32).

EGR-1

The early growth response gene (EGR-1) encodes a zinc finger transcription factor, crucial for regulating cell growth, differentiation, and survival. Abnormal expression of EGR-1 relates to inflammation, atherosclerosis, and carcinogenesis (33). Previous study also associated EGR-1 expression with obese pregnant women, in which exposure to lipotoxic stimuli can lead to a transcriptionally tolerant chromatin profile of its promoter and allows the expression of pro-inflammatory cytokines (34), resulting in cervical ripening and the onset of PTB.

DEGs in Placenta

Placental meta-DEGs primarily involve in ECM organisation. The ECM plays a key role in maintaining uterine structural integrity, facilitating embryo adhesion, and regulating placental invasion into the endometrium to establish the maternal-foetal interface (35). Membrane-bound adhesion molecules, such as integrins and selectins, mediate cell attachment to the ECM which allows the cells to perceive environmental signals, such as pressure or substrate stiffness, and translate them into biochemical signals that guide cellular structure and behaviour. Disruption in ECM components expression can compromise ECM stability, resulting in adverse outcomes such as abnormal placentation, PPROM, cervical insufficiency, uterine rupture, and pelvic organ prolapse (36).

KEGG analysis further highlighted the involvement of placental meta-DEGs in malaria and PI3K-AKT signalling pathways. Malaria infection, particularly in the placenta before 24 weeks of gestation, is associated with dysregulation of essential regulators in angiogenesis and inflammation, elevating the risk of placental dysfunction and, ultimately contributing to PTB (37). Moreover, the PI3K/AKT pathway, vital for diverse cellular activities, has been recognised as a precursor of PTB. Desensitisation of the PI3K/AKT pathway increases the soluble endoglin, sEng (38). sEng hinders angiogenic processes such as capillary formation, inhibiting placental vascularisation necessary for the efficient growth of the developing foetus. Elevated maternal sEng levels also is associated with P. falciparum infection in pregnancy, linking both pathways in the initiation of PTB (38).

Eighteen placental hub genes were identified through network clustering analysis. The most prominent among them, based on cluster scores, include IGF1 (downregulated), FN1 (upregulated), and LOX (downregulated).

IGF1

Insulin-like growth factor (IGF1) controls the proliferation and survival of placental fibroblasts in early pregnancy, contributing to placental vascular development (39). Aberrations in vascular development can lead to oxidative stress in the maternal body, triggering systemic inflammation (40). A study by (41) have reported lower expression of IGF1 in the placenta of preterm babies compared to full-term babies, consistent with the findings of this study.

FN1

Fibronectin 1 (FN1) is one of the prominent fibrosis-related factors, and previous studies have indicated that its overexpression in the placenta engendered placental fibrosis (42). It has been proposed that placental fibroblasts, under hypoxic or ischemic conditions, stimulate ECM overproduction, inducing fibrosis – a notable feature of conditions such as pre-eclampsia (43). This finding aligns with the results of the current study.

LOX

Studies have demonstrated that lysil oxidase (LOX) expression is decreased in PTB, particularly in preeclamptic placentas (44), aligning with the findings of this study. Additionally, LOX downregulation enhances collagen activation in trophoblast cells through stimulation of the TGF-1/Smad3 pathway (44). Inadequate vascular remodelling at the maternal-foetal interface may arise from excessive collagen deposition, leading to pregnancy-induced hypertension (45).

Diagnostic potential of the hub genes

Although previous literature suggests that these genes may directly or indirectly play a prominent role in sPTB, it is important to acknowledge that their regulation and molecular mechanism in sPTB remains poorly defined. Therefore, further functional investigations are essential to clarify their exact role in the pathogenesis of sPTB.

The assessed diagnostic performance of these hub

genes shows promising potential in early screening of sPTB. Among these, EGR-1 demonstrated the strongest diagnostic potential in maternal blood, with an AUC of 0.81, followed by HSP90AA1 (AUC: 0.737), while IKBKG, with an AUC of 0.6 demonstrated the weakest performance. Notably, EGR-1, and HSP90AA1 exhibited relatively high specificity, 100% and 94.7% respectively, indicating their ability to correctly identify individuals not at risk of sPTB. However, these genes also showed low to moderate sensitivity, with EGR1 at 58.3%, and HSP90AA1 at 46.2%, suggesting a considerable number of true sPTB cases may go undetected if these markers are used individually.

In placental samples, IGF1 and LOX showed acceptable diagnostic performance with AUCs of 0.75 and 0.762, respectively. IGF1 had high sensitivity (87.5%) and moderate specificity (66.7%), indicating its ability to correctly identify most true sPTB cases. Similarly, LOX exhibited very high sensitivity (93.3%) but extremely low specificity (12.5%), implying a high rate of false positives. In contrast, FN1 performed poorly (AUC = 0.505), with low specificity (20.0%) despite high sensitivity (92.9%), rendering it unsuitable as a standalone marker.

Collectively, these findings suggest that while certain genes, particularly EGR-1, HSP90AA1, and IGF1, show promise in distinguishing between sPTB and term deliveries, none of the genes evaluated exhibit both high sensitivity and specificity when used individually. This imbalance between specificity and sensitivity underscores the limitation of relying on any single gene as a standalone diagnostic marker and supports the need for further studies to validate their performance in larger cohorts and evaluate their use in combinatorial biomarker panels for improved early detection of sPTB. A multi-marker panel incorporating genes with complementary diagnostic characteristics, such as pairing high-specificity genes (e.g., EGR-1, HSP90AA1) with high-sensitivity ones (IGF1, LOX)—could improve overall diagnostic accuracy and reliability. This combinatorial approach may be more effective for early detection and risk stratification of sPTB, especially in asymptomatic pregnancies. Future research should therefore focus on validating these markers in combination and integrating them with clinical data or other molecular indicators to enhance their predictive power.

Common meta-DEGs between Maternal Blood and Placenta

On the other hand, the 32 common meta-DEGs identified in this study, were involved in a similar biological process to the maternal blood DEGs and contributed to pathways involved in infection. Common genes expressed in different magnitudes were involved in infectious diseases such as malaria. In contrast, genes that shares a similar magnitude of expression involve glycine, serine, and threonine metabolism, a pathway prominently involved in immune response (46). Downregulation of this metabolomic pathway, reduces the metabolism of glutathione and causes the activation of ROS (47). This further leads to the overexpression of IL-6, IL-8, and IL-1B, eventually causing tissue injury and inflammation. As such, the differential expression of these genes in both placenta and maternal blood indicate involvement in infection and immune-related processes.

Nevertheless, no study has studied the magnitude of the differential expression of these common genes in both maternal blood and the placenta. This lack of investigation could be attributed to the placenta's role as a conduit between the mother and foetus, facilitating the transport of nutrients and waste excretion (46). Thus, it was hypothesised that gene expression levels might differ between the maternal and foetal sides of the placenta due to the maternal and foetal genomes, respectively (47). The uncertainty about which side of the placenta the tissue was collected from in the four studies included in the DEG analysis also suggests that the observed discrepancy might be due to sampling from different placental zones. Although plausible, no precise homeostatic mechanism has been identified linking the upregulation of a gene in one tissue (e.g., placenta) to its down-regulation in another tissue (e.g., maternal blood). Therefore, the genes noted above should be considered secondary candidate biomarkers and warrant further investigation. Validation using experimental method such as qPCR and functional analysis in independent maternal and placental tissue cohorts could help clarify tissue-specific expression patterns and support their potential as biomarkers.

Future prospects of the study

In overall, the identification of central and highly interconnected DEGs with good diagnostic potential such as EGR-1 and HSP90AA1 in maternal blood, as well as IGF1 in placental samples, not only deepens the understanding of the molecular mechanisms underlying sPTB but also holds translational potential. Their consistent differential expression across maternal and placental datasets, combined with centrality in protein interaction networks, suggests their stability and biological relevance. If validated in prospective cohorts, blood-based measurement of these genes could allow early identification of women at risk, enabling closer monitoring and timely intervention. Furthermore, several hub genes identified are involved in targetable pathways, such as cytokine signalling and oxidative stress responses. This raises the possibility of therapeutic targeting, such as using anti-inflammatory agents or modulators of oxidative stress in women at risk of sPTB.

For instance, HSP90AA1 is already a known drug target in cancer and inflammatory disorders (48), with inhibitors like geldanamycin and its derivatives under investigation (49). Similarly, IGF1 plays a role in cell proliferation and has been explored in metabolic and

growth-related therapies, suggesting its pathways could be pharmacologically modulated (50). While genes such as EGR-1, LOX, and IKBKG are less established as drug targets, they participate in pathways (e.g., NF-κB signalling, extracellular matrix remodelling) that are amenable to modulation. These overlaps with druggable pathways warrant further exploration. Future pharmacogenomic and drug-repurposing studies are needed to evaluate whether modulating the expression or activity of these genes could contribute not only to the early detection but also to the prevention or mitigation of sPTB.

CONCLUSION

In conclusion, the study highlights key DEGs, notably IKBKG, HSP90AA1, EGR1, IGF1, FN1, and LOX, as potential contributors of PTB. These genes warrant further investigation and validation to unravel their precise roles in PTB pathology. Moreover, the common DEGs identified between maternal blood and placenta should be explored further, considering their potential collective impact on PTB occurrence.

While this study offers valuable insights by identifying critical pathways and hub genes that may serve as screening targets, it is not without limitations. The reliance on publicly available datasets introduces variability in sample sources, platforms, and clinical annotations, which may affect the generalizability of the findings. Moreover, the lack of clinical or experimental validation limits the immediate applicability of these DEGs as biomarkers.

Given its non-invasive nature, maternal blood is particularly suitable for early screening efforts, and the identified DEGs in maternal blood samples exhibit higher potential as biomarkers for early detection of PTB. To build upon these findings, a well-designed prospective study is recommended, such as evaluating the expression of identified DEGs in maternal blood samples collected longitudinally throughout pregnancy. This would not only support the development of early biomarkers for PTB but also strengthen the translational relevance of the identified gene candidates. Overall, this study lays a foundational framework for understanding gene dysregulation in PTB and represents a significant step toward the development of predictive tools and interventions aimed at improving maternal and neonatal health outcomes.

ACKNOWLEDGEMENTS

This study is supported by the FRGS/1/2020/SKK08/UPM/02/1 grant under the Ministry of Higher Education. The study was submitted as an abstract and presented as a poster at the 3rd Women in Bioinformatics & Data Science LA conference, held virtually on 29th – 30th August 2022.

REFERENCES

1. Di Renzo GC, Tosto V, Giardina I. The biological basis and prevention of preterm birth. Best Pract Res Clin Obstet Gynaecol. 2018;52:13–22. doi: 10.1016/j.bpobgyn.2018.01.022.
2. Menon R. Spontaneous preterm birth, a clinical dilemma: Etiologic, pathophysiologic and genetic heterogeneities and racial disparity. Acta Obstet Gynecol Scand. 2008;87(6):590–600. doi: 10.1080/00016340802005126.
3. Heng YJ, Pennell CE, McDonald SW, Vinturache AE, Xu J, Lee MWF, et al. Maternal whole blood gene expression at 18 and 28 weeks of gestation associated with spontaneous preterm birth in asymptomatic women. PLoS One. 2016;11(6):e0155191. doi: 10.1371/journal.pone.0155191
4. Paquette AG, Shynlova O, Kibschull M, Price ND, Lye SJ. Comparative analysis of gene expression in maternal peripheral blood and monocytes during spontaneous preterm labor. Am J Obstet Gynecol. 2018;218(3):345.e1–30. doi: 10.1016/j.ajog.2017.12.234.
5. Menon R, Torloni MR, Voltolini C, Torricelli M, Merialdi M, Betrбn AP, et al. Biomarkers of spontaneous preterm birth: an overview of the literature in the last four decades. Reprod Sci. 2011;18(11):1046–70. doi: 10.1177/1933719111415548.
6. Yoo JY, Hyeon DY, Shin Y, Kim SM, You YA, Kim DY, et al. Integrative analysis of transcriptomic data for identification of T-cell activation-related mRNA signatures indicative of preterm birth. Sci Rep. 2021;11(1):1–12. doi: 10.1038/s41598-021-81834-z.
7. Vora NL, Smeester L, Boggess K, Fry RC. Investigating the role of fetal gene expression in preterm birth. Reprod Sci. 2017;24(6):824–8. doi: 10.1177/1933719116670038.
8. Mele M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, et al. The human transcriptome across tissues and individuals. Science. 2015;348(6235):660–5. doi: 10.1126/science.aaa0355.
9. Law CW, Alhamdoosh M, Su S, Dong X, Tian L, Smyth GK, et al. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. F1000Res [Internet]. 2018 [cited 2020 Sep 23];5. doi: 10.12688/f1000research.9005.3.
10. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. 2010;11(3):R25. doi: 10.1186/gb-2010-11-3-r25.
11. Law CW, Chen Y, Shi W, Smyth GK. voom: precision weights unlock linear model analysis

tools for RNA-seq read counts. Genome Biol. 2014;15(2):R29. doi: 10.1186/gb-2014-15-2-r29.
12. Smyth G, Hu Y, Ritchie M, Silver J, Wettenhall J, McCarthy D, et al. limma: Linear Models for Microarray Data [Internet]. Bioconductor. 2021 [cited 2024 Jan 15]. Available from: https://bioconductor.org/packages/release/bioc/html/limma.html
13. Ritchie ME, Diyagama D, Neilson J, van Laar R, Dobrovic A, Holloway A, et al. Empirical array quality weights in the analysis of microarray data. BMC Bioinformatics. 2006;7(1):261. doi: 10.1186/1471-2105-7-261.
14. Kugler KG, Mueller LA, Graber A. MADAM – An open source meta-analysis toolbox for R and Bioconductor. Source Code Biol Med. 2010;5(1):3. doi: 10.1186/1751-0473-5-3.
15. Phung DM, Lee J, Hong S, Kim YE, Yoon J, Kim YJ. Meta-analysis of differentially expressed genes in the substantia nigra in Parkinson's disease supports phenotype-specific transcriptome changes. Front Neurosci. 2020;14:613. doi: 10.3389/fnins.2020.596105.
16. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. Nucleic Acids Res [Internet]. 2016 [cited 2024 Jan 15];45(D1):D362–8. doi: 10.1093/nar/gkw937.
17. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol [Internet]. 2014 [cited 2024 Jan 15];8(Suppl 4):S11. doi: 10.1186/1752-0509-8-S4-S11.
18. Tang Y, Li M, Wang J, Pan Y, Wu FX. CytoNCA: A Cytoscape plugin for centrality analysis and evaluation of protein interaction networks. Biosystems. 2015;127:67–72. doi: 10.1016/j.biosystems.2014.11.005.
19. Kumar S, Kumar D, Siva R, Priya G, Younes S, Younes N, et al. Dysregulation of signaling pathways due to differentially expressed genes from the B-cell transcriptomes of systemic lupus erythematosus patients – a bioinformatics approach. Front Bioeng Biotechnol. 2020;8:412. doi: 10.3389/fbioe.2020.00276.
20. Zimmermann P, Laule O, Schmitz J, Hruz T, Bleuler S, Gruissem W. Genevestigator transcriptome meta-analysis and biomarker search using rice and barley gene expression databases. Mol Plant. 2008;1(5):851–7. doi: 10.1093/mp/ssn048.
21. Mandrekar, JN. Receiver operating characteristic curve in diagnostic Test assessment. J. Thorac Oncol. 2010 Sep; 5(9), 1315–1316. doi: 10.1097/JTO.0b013e3181ec173d.
22. Elbaz G, Fich A, Levy A, Holcberg G, Sheiner E. Inflammatory bowel disease and preterm delivery. Int J Gynaecol Obstet. 2005;90(3):193–7. doi: 10.1016/j.ijgo.2005.06.003.
23. Huang W, Sundquist K, Sundquist J, Ji J. Risk of being born preterm in offspring of cancer survivors: a national cohort study. Front Oncol. 2020;10:1450. doi: 10.3389/fonc.2020.01352.
24. Wang F, Wang Y, Wang R, Qiu H, Chen L. Predictive value of maternal serum NF-κB p65 and sTREM-1 for subclinical chorioamnionitis in premature rupture of membranes. 2016 Aug 13. doi: 10.1111/aji.12543.
25. Salminen A, Paananen R, Vuolteenaho R, Metsola J, Ojaniemi M, Autio-Harmainen H, et al. Maternal endotoxin-induced preterm birth in mice: fetal responses in toll-like receptors, collectins, and cytokines. Pediatr Res. 2008;63(3):280–6. doi: 10.1203/PDR.0b013e318163a8b2.
26. Kaur R, Bhat P. A case of Salmonella typhi infection leading to miscarriage. J Lab Physicians. 2011;3(1):61–2. doi: 10.4103/0974-2727.78574
27. Rebarber A. Shigellosis complicating preterm premature rupture of membranes resulting in congenital infection and preterm delivery. Obstet Gynecol. 2002;100(5):1063–5. doi: 10.1016/s0029-7844(02)01992-0.
28. Appelberg R. CD4+ T cells are required for antigen-specific recruitment of neutrophils by BCG-immune spleen cells. Immunology [Internet]. 1992 [cited 2024 Jan 15];75(3):414–9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1384733/
29. Vieira S, Lemos H, Grespan R, Napimoga M, Dal-Secco D, Freitas A, et al. A crucial role for TNF-α in mediating neutrophil influx induced by endogenously generated or exogenous chemokines, KC/CXCL1 and LIX/CXCL5. Br J Pharmacol. 2009;158(3):779–89. doi: 10.1111/j.1476-5381.2009.00367.x.
30. Gymez-Ch6vez F, Correa D, Navarrete-Meneses P, Cancino-Diaz JC, Cancino-Diaz ME, Rodrнguez-Martнnez S. NF-κB and its regulators during pregnancy. Front Immunol. 2021;12:687. doi: 10.3389/fimmu.2021.679106.
31. Anamthathmakula P, Kyathanahalli C, Ingles J, Hassan SS, Condon JC, Jeyasuria P. Estrogen receptor alpha isoform ERdelta7 in myometrium modulates uterine quiescence during pregnancy. EBioMedicine. 2019;39:520 - 530. doi: 10.1016/j.ebiom.2018.11.038.
32. Huusko JM, Tiensuu H, Haapalainen AM, Pasanen A, Tissarinen P, Karjalainen MK, et al. Integrative genetic, genomic and transcriptomic analysis of heat shock protein and nuclear hormone receptor gene associations with spontaneous preterm birth. Sci Rep. 2021 Aug 24;11(1). doi: 10.1038/s41598-021-96374-9.
33. Wu M, Melichian DS, de la Garza M, Gruner K, Bhattacharyya S, Barr L, et al. Essential roles for early growth response transcription factor Egr-1 in tissue fibrosis and wound healing. Am J Pathol. 2009 Sep 1; 175(3):1041–55. doi: 10.2353/

ajpath.2009.090241.
34. Enquobahrie DA, Williams MA, Qiu C, Muhie S, Slentz-Kesler K, Ge Z, et al. Early pregnancy peripheral blood gene expression and risk of preterm delivery: a nested case control study. BMC Pregnancy Childbirth. 2009 Dec 1;9(1). doi: 10.1186/1471-2393-9-56.
35. O'Connor BB, Pope BD, Peters MM, Ris-Stalpers C, Parker KK. The role of extracellular matrix in normal and pathological pregnancy: future applications of microphysiological systems in reproductive medicine. Exp Biol Med (Maywood). 2020 Jul ;245(13):1163–74. doi: 10.1177/1535370220938741.
36. Strauss JF. Extracellular matrix dynamics and fetal membrane rupture. Reprod Sci. 2012 Jan 19;20(2):140–53. doi: 10.1177/1933719111424454.
37. Elphinstone RE, Weckman AM, McDonald CR, Tran V, Zhong K, Madanitsa M, et al. Early malaria infection, dysregulation of angiogenesis, metabolism and inflammation across pregnancy, and risk of preterm birth in Malawi: a cohort study. PLoS Med. 2019 Oct 1;16(10):e1002914. doi: 10.1371/journal.pmed.1002914.
38. Silver KL, Conroy AL, Leke RGF, Leke RJI, Gwanmesia P, Molyneux ME, et al. Circulating soluble endoglin levels in pregnant women in Cameroon and Malawi—associations with placental malaria and fetal growth restriction. PLoS One. 2011 Sep 22;6(9):e24985. doi: 10.1371/journal.pone.0024985.
39. Sferruzzi-Perri AN, Owens JA, Standen P, Taylor RL, Robinson JS, Roberts CT. Early pregnancy maternal endocrine insulin-like growth factor I programs the placenta for increased functional capacity throughout gestation. Endocrinology. 2007 Sep 1;148(9):4362–70. doi: 10.1210/en.2007-0411.
40. Roberts JM, Hubel CA. The two stage model of preeclampsia: variations on the theme. Placenta. 2009 Mar;30 Suppl A:S32–7. doi: 10.1016/j.placenta.2008.11.009.
41. Nawathe AR, Christian M, Kim SH, Johnson M, Savvidou MD, Terzidou V. Insulin-like growth factor axis in pregnancies affected by fetal growth disorders. Clin Epigenetics. 2016 Jan 27;8(1). doi: 10.1186/s13148-016-0178-5
42. Matsubara S, Asanoma K, Fujikawa M, Fujita Y, Yagi H, Onoyama I, et al. Fibrosis in preeclamptic placentas is associated with stromal fibroblasts activated by the transforming growth factor-β1 signaling pathway. Am J Pathol. 2018 Mar 1;188(3):683–95. doi: 10.1016/j.ajpath.2017.11.008.
43. Devisme L, Merlot B, Ego A, Houfflin-Debarge V, Deruelle P, Subtil D. A case-control study of placental lesions associated with pre-eclampsia. Int J Gynaecol Obstet. 2013 Feb 1;120(2):165–8. doi: 10.1016/j.ijgo.2012.08.023.
44. Xu XH, Jia Y, Zhou X, Xie D, Huang X, Jia L, et al. Downregulation of lysyl oxidase and lysyl oxidase-like protein 2 suppressed the migration and invasion of trophoblasts by activating the TGF-β/collagen pathway in preeclampsia. Exp Mol Med. 2019 Feb;51(2). doi: 10.1038/s12276-019-0211-9.
45. Shi JW, Lai ZZ, Yang HL, Yang SL, Wang CJ, Ao D, et al. Collagen at the maternal-fetal interface in human pregnancy. Int J Biol Sci. 2020;16(12):2220–34. doi: 10.7150/ijbs.45586.
46. Cheng Z, Guo C, Chen Z, Yang T, Zhang J, Wang J, et al. Glycine, serine and threonine metabolism confounds efficacy of complement-mediated killing. Nat Commun. 2019 Jul 25;10(1):3325. doi: 10.1038/s41467-019-11129-5.
47. Garnica AD, Chan WY. The role of the placenta in fetal nutrition and growth. J Am Coll Nutr. 1996 Jun 1;15(3):206–22. doi: 10.1080/07315724.1996.10718591.
48. Chen O, Fu L, Wang Y, Li J, Liu J, Wen Y. Targeting HSP90AA1 to overcome multiple drug resistance in breast cancer using magnetic nanoparticles loaded with salicylic acid. Int J Biol Macromol. 2025 Apr;298:139443. doi: 10.1016/j.ijbiomac.2024.139443.
49. Kitson RRA, Kitsonov6 D, Siegel D, Ross D, Moody CJ. Geldanamycin, a naturally occurring inhibitor of Hsp90 and a lead compound for medicinal chemistry. J Med Chem. 2024 Oct 3;67(20):17946–63. doi: 10.1021/acs.jmedchem.4c01048.
50. Rong P, Mu Y, Wang M, Chen L, Liu F, Jin Y, et al. Targeting IGF1 to alleviate obesity through regulating energy expenditure and fat deposition. Sci China Life Sci. 2025 Jan 21;68:1662–75. doi: 10.1007/s11427-024-2768-y.