RNA-Sequencing of Umbilical Cord Blood to Investigate Spontaneous Preterm Birth: A Pilot Study

Neeta L. Vora, MD1 Joel S. Parker, PhD2 Piotr A. Mieckowski, PhD2 Lisa Smeester, MS3 Rebecca C. Fry, PhD3 Kim A. Boggess, MD1

1 Department of Obstetrics and Gynecology, University of North Carolina School of Medicine, Chapel Hill, North Carolina
2 Department of Genetics, Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, North Carolina
3 Department of Environmental Sciences and Engineering, University of North Carolina Gillings School of Global Public Health, Chapel Hill, North Carolina

Address for correspondence Neeta L. Vora, MD, Division of Maternal-Fetal Medicine, University of North Carolina at Chapel Hill, 3010 Old Clinic Building, CB # 7516, Chapel Hill, NC 27599-7516 (e-mail: nvora@med.unc.edu).

In the United States, preterm birth (PTB) accounts for 9.6% of live births and is the leading cause of infant morbidity and mortality in nonanomalous infants. PTB poses a significant economic burden of up to $26 billion to care for these infants in the United States. Primary prevention strategies such as antioxidant supplementation or screening and treatment of maternal infections have failed to reduce or eliminate spontaneous PTB (sPTB). To date, much of the focus in understanding PTB has focused on maternal factors that incite preterm labor, such as inflammation, infection, and maternal decidual or endometrial factors, and not on possible contributing fetal factors. Little is known about the fetal

Abstract

Objective To analyze the transcriptomic gene expression of umbilical cord blood leukocytes using RNA-sequencing from preterm birth (PTB) and term birth (TB).

Study Design Eight women with spontaneous PTB (sPTB) and eight women with unlabored TB were enrolled prospectively. The sPTB and TB cohorts were matched for maternal age, race, mode of delivery, and fetal sex. Cord blood RNA was extracted and a globin depletion protocol was applied, then sequenced on the Illumina HiSeq 4000. Raw read counts were analyzed with DESeq2 to test for gene expression differences between sPTB and TB.

Results 148 genes had significant differential expression ($q < 0.01$). Cell cycle/metabolism gene expression was significantly higher and immune/inflammatory signaling gene expression significantly lower in the sPTB cohort compared with term. In African American (AA) infants, 18 genes specific to cell signaling, neutrophil activity, and major histocompatibility complex type 1 had lower expression in preterm compared with term cohort; the opposite pattern was seen in non-Hispanic Whites (NHWs).

Conclusion Compared with term, preterm fetuses have higher cell cycle/metabolism gene expression, suggesting metabolic focus on growth and development. Immune function gene expression in this pilot study is lower in the sPTB group compared with term and differs in AA compared with NHW infants.
A new and emerging high throughput next-generation genomic technologies have led to the ability to sequence messenger RNA, permitting interrogation of the entire fetal transcriptome in umbilical cord blood. The transcriptome is the sum total of messenger RNA expressed in a tissue. Transcriptome analysis captures a snapshot of cellular activity that reflects the response to genetic, environmental, and epigenetic changes in a biological system. Knowledge of the transcriptome allows for the quantification and assessment of genes that may be active in different disease processes and at various stages of development. This technology has been applied to cancer therapeutics and diagnostics leading to new insights and therapeutics. While transcriptomics has been studied in pregnancy conditions such as sPTB, preeclampsia, and obesity among others, this field is still emerging in terms of shedding light on the molecular underpinnings of these complex pregnancy disorders.

The objective of the study was to measure fetal gene expression from umbilical cord blood at the time of delivery in term and preterm pregnancies to identify differentially regulated genes related to common PTB pathways, such as inflammation, immune function, and oxidative stress. Our second objective was to evaluate differences in gene expression in preterm compared with term fetuses to gain insight into fetal development. We hypothesized that fetal genes are differentially expressed in common PTB pathways following sPTB compared to term birth (TB). These findings have the potential to increase our understanding of the fetal molecular contribution to sPTB, and will lay the foundation to improve diagnosis, prognosis, and therapeutic strategies in obstetrics and pediatrics.

### Methods

The study was approved by the University of North Carolina at Chapel Hill Institutional Review Board. This prospective case-control study included eight women who delivered via idiopathic sPTB (<34 weeks) and eight women who delivered at term (>37 weeks) with singleton fetuses who delivered at the University of North Carolina at Chapel Hill. Preterm labor was defined as the presence of regular uterine contractions and documented cervical effacement and/or dilatation in patients <37 weeks’ gestational age (GA). The preterm patients were admitted to the antepartum service and presented with preterm contractions. TB was defined as delivery at greater than 37 weeks gestation with no labor. Preterm premature rupture of membranes (PPROM) was confirmed by vaginal “pooling,” and positive “nitrazine” or “ferning” tests.

The sPTB and TB cohorts were matched for factors that could affect fetal gene expression including: maternal age, race, fetal sex, medication exposure except for glucocorticoids (all women in the sPTB received betamethasone as per American College of Obstetricians and Gynecologists’ guidelines), and mode of delivery (all women had cesarean sections). Pertinent clinical information about the patients is provided in Table 1. Term patients were enrolled prior to elective cesarean delivery at term (390/7 to 396/7 weeks). All preterm infants received steroids prior to delivery with the median number of days from administration of steroids to delivery of 2.18 (range: 1–19 days). Four of the preterm fetuses were born via sPTB after PPROM with the median number of days from rupture to delivery of 16 (range: 3–38 days). The phenotype of sPTB was limited to women who had spontaneous labor with or without PPROM. Patients with abortion, diabetes, clinical chorioamnionitis, fetal anomalies, and medication exposures were excluded because these factors can all affect neonatal gene expression. Clinical chorioamnionitis was defined by fever of ≥38.0°C in addition to at least two other signs of chorioamnionitis, including uterine fundal tenderness, maternal tachycardia (>100 per minute), fetal tachycardia (>160 per minute), and purulent or foul-smelling amniotic fluid.

Power studies of gene expression using DESeq2 have shown that eight cases and controls are a sufficient number to have adequate power to detect significant differential gene expression between dichotomous groups (term/preterm). It was not possible to match by GA or glucocorticoid exposure except for glucocorticoids (all women in the sPTB received betamethasone as per American College of Obstetricians and Gynecologists’ guidelines), and mode of delivery (all women had cesarean sections). Pertinent clinical information about the patients is provided in Table 1. Term patients were enrolled prior to elective cesarean delivery at term (390/7 to 396/7 weeks). All preterm infants received steroids prior to delivery with the median number of days from administration of steroids to delivery of 2.18 (range: 1–19 days). Four of the preterm fetuses were born via sPTB after PPROM with the median number of days from rupture to delivery of 16 (range: 3–38 days). The phenotype of sPTB was limited to women who had spontaneous labor with or without PPROM. Patients with abortion, diabetes, clinical chorioamnionitis, fetal anomalies, and medication exposures were excluded because these factors can all affect neonatal gene expression. Clinical chorioamnionitis was defined by fever of ≥38.0°C in addition to at least two other signs of chorioamnionitis, including uterine fundal tenderness, maternal tachycardia (>100 per minute), fetal tachycardia (>160 per minute), and purulent or foul-smelling amniotic fluid.

Power studies of gene expression using DESeq2 have shown that eight cases and controls are a sufficient number to have adequate power to detect significant differential gene expression between dichotomous groups (term/preterm). It was not possible to match by GA or glucocorticoid exposure except for glucocorticoids (all women in the sPTB received betamethasone as per American College of Obstetricians and Gynecologists’ guidelines), and mode of delivery (all women had cesarean sections). Pertinent clinical information about the patients is provided in Table 1. Term patients were enrolled prior to elective cesarean delivery at term (390/7 to 396/7 weeks). All preterm infants received steroids prior to delivery with the median number of days from administration of steroids to delivery of 2.18 (range: 1–19 days). Four of the preterm fetuses were born via sPTB after PPROM with the median number of days from rupture to delivery of 16 (range: 3–38 days). The phenotype of sPTB was limited to women who had spontaneous labor with or without PPROM. Patients with abortion, diabetes, clinical chorioamnionitis, fetal anomalies, and medication exposures were excluded because these factors can all affect neonatal gene expression. Clinical chorioamnionitis was defined by fever of ≥38.0°C in addition to at least two other signs of chorioamnionitis, including uterine fundal tenderness, maternal tachycardia (>100 per minute), fetal tachycardia (>160 per minute), and purulent or foul-smelling amniotic fluid.

Table 1 Demographic characteristics of the sPTB and TB cohorts

<table>
<thead>
<tr>
<th>Demographic characteristics</th>
<th>sPTB</th>
<th>TB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age, mean (SD) [range] (y)</td>
<td>30 (8.3) [22–44]</td>
<td>35.2 (3.92) [30–42]</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>African American</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Gestational age at delivery, [range] (wk)</td>
<td>29 [241/3 to 341/7]</td>
<td>39 [370/7 to 404/7]</td>
</tr>
<tr>
<td>Fetal sex (number of males, number of females)</td>
<td>6, 2</td>
<td>6, 2</td>
</tr>
<tr>
<td>Cesarean sections</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>
intracellular RNA and in processing the tubes, intact cells are pelletted and RNA extracted. Total RNA was extracted from whole blood using a PAXgene Blood miRNA kit on the QIAcube. A globin depletion protocol using custom Human Globin AnyDeplete (NuGen) probes enriched with fetal globin gene primers was applied to reduce uninformative reads, enabling more RNA transcripts to be detected. Quality control was performed for all samples using the LabChip GX system. All samples were handled in a uniform way by one technician. Universal Plus mRNA-seq library prep kit (NuGen) was used to generate mRNA-seq libraries from 500-ng umbilical cord blood total RNA. Cord blood (coding and noncoding) was sequenced on the Illumina HiSeq 4000 PE 2 × 50. Sequencing reads were analyzed as follows: chastity filtered reads were aligned to hg38 using single pass STAR\(^{14}\) with default parameters and transcript abundance estimates for each sample were estimated with Salmon\(^{15,16}\) to quantify the transcriptome defined by Gencode v22. Gene level counts were summed across isoforms and the resulting gene count estimates were analyzed with DESeq2\(^{17}\) to test for differences between sPTB and TB adjusted for race. Race was modeled as an additive covariate, and sPTB versus TB effects were tested using the DESeq2 version 1.4.5 likelihood ratio test comparing the bivariate model of race and birth to a reduced model of just race. Significant differences in gene expression were defined by a Benjamini–Hochberg FDR (q-value) <0.1 for the primary outcome (GA).\(^{18}\) Individual cellular fractions by race and GA were estimated by xCell followed by association testing with a linear model.\(^{19}\) Gene set analysis was performed to identify differentially expressed pathways.\(^{20}\) The Gene Set Analysis in R (GSAR) is an open-source R/Bioconductor software package for gene set analysis. A gene set is a group of genes that shares pathways, functions, chromosomal localization, or other features.

Hierarchical clustering and principal component analysis. Two-dimensional unsupervised hierarchical clustering, dendrogram generation, and heat-map plotting were performed using the pheatmap R package version 0.7.7. A complete linkage algorithm with a Manhattan distance function was applied to normalized, log2transformed fold change data for statistically significant differentially expressed RNAs. Principal component analysis for differentially expressed transcripts was accomplished using the ‘prcomp’ function in R version 3.0.3, in which a singular value decomposition algorithm was applied to a centered and scaled correlation matrix of normalized, log2transformed fold change RNA-seq data. Three-dimensional scatterplots of the first three principal components (PC1, PC2, PC3) were generated using SigmaPlot version 12.5 (Systat Software, San Jose, CA; – Supplementary Fig. 1).

Results

RNA-seq was performed on cord blood collected from TB (\(n = 8\)) and sPTB (\(n = 8\)) events. The median RNA integrity number (RIN) value of the cord blood samples was 8.5 (range: 7–10); there was no statistical difference in RIN values between cases and controls. Fetal sex and maternal race displayed only minor effects on transcript abundance as fetal sex was associated with only two genes (XIST and TSIX) and race with none at the specified criteria.

One hundred and forty eight genes were differentially expressed in the setting of sPTB compared with TB. Two pathways, cell cycle/metabolism and immune/inflammatory signaling, represent the predominant pathways with significant differential expression in the preterm versus term cohorts. Cell cycle/metabolism (e.g., electron transport chain) gene expression was significantly higher (\(z\)-score = 18–24), and immune/inflammatory signaling gene expression (cytokines, interleukins, and T-helper cells) was significantly lower (\(z\)-score = –41 to –26) in the sPTB cord blood compared with term (\(\sim\) Figs. 1 and 2).

Exploratory analysis was performed to investigate interactions of GA with race. The results yielded 18 genes decreased in preterm African American (AA) infants and increased in term AA infants (\(q < 0.1\)). However, the opposite pattern was seen in non-Hispanic White (NHW) preterm/term infants (\(\sim\) Fig. 3). These included genes associated with cell signaling, neutrophil activity, and major histocompatibility complex type 1: DmX-like protein 2 (DMXL2), ubiquitin-like modifier ISG15 (ISG15), and butyrophilin subfamily 3 member A2 (BTN3A2).

Gene expression data were evaluated using xCell to quantify a variety of cellular phenotypes. There were no differences between leukocyte/lymphocyte compositions of the preterm and term fetal cord blood; however, we were underpowered to detect a difference. Preterm infants had significantly lower macrophage gene expression (\(p = 0.001\)) compared with term infants. This cellular composition analysis testing for race-dependent effects showed that preterm and term AA infants had lower native CD4 T-cell gene expression compared with NHW infants (\(p = 0.01; \sim\) Fig. 4).
Half of the preterm placentas and cords showed evidence of chorionitis and funisitis even though none of the patients had clinical chorioamnionitis. None of the neonates had sepsis. All of the term infants transitioned to room air without difficulty, had Apgar scores at 1 and 5 minutes of >8, and did well clinically.

Discussion

Our findings that immune and inflammatory gene expressions were higher in term compared with preterm infants provide novel insights into fetal development and fetal gene expression in sPTB compared with TB. While other authors studied PTB by focusing on gene expression changes in maternal whole blood, we chose to focus on umbilical cord blood samples as they are representative of the in utero environment of a neonate shortly before birth. While other investigators have examined gene expression via microarray in amniotic fluid supernatant, a few have performed whole transcriptome analyses using RNA-seq from cord blood, and none to date have done this on preterm infants following sPTB. Our findings suggest that despite the inflammatory milieu so commonly seen in sPTB, preterm infants born due to sPTB had lower gene expression of immune/inflammatory signaling function compared with term infants. This finding alone is not surprising as the immune system is immature in preterm infants and not even fully mature at the time of TB as the adaptive immune system must develop specificity and memory, which does not occur until the early childhood years.

We found fetal genes associated with inflammation were down-regulated in the infants born preterm compared with term despite previously published clinical and experimental data showing a measurable fetal inflammatory response following preterm labor and other inflammatory processes.

In fact, our group found that even prior to the onset of symptoms of sPTB, there is evidence of a fetal inflammatory response. In a previous study, we extracted cell-free fetal (cff)
RNA from amniotic fluid supernatant and used a customized nanostring panel containing genes related to common PTB pathways (oxidative stress and inflammation) and compared gene expression in fetuses who delivered preterm and those who delivered at term. Compared with fetuses who ultimately delivered at term, fetuses who delivered preterm had significantly increased gene expression of β-2 microglobulin in amniotic fluid, a marker of the fetal inflammatory response.25

The question then arises regarding why other data including data from our laboratory show evidence of a fetal inflammatory response and this cohort does not. We hypothesize several possible reasons. Previous studies including our own measured gene expression via cff RNA in amniotic fluid;21,25 this study uses RNA extracted from cord blood. Gene expression in two different biofluids and from two different forms of RNA (cff RNA vs. RNA) can differ. The gene expression from RNA in amniotic fluid supernatant is a reflection of gene expression in various fetal secretions including from urine, saliva, and tracheal secretions compared with gene expression found in umbilical cord blood. Prior studies using amniotic fluid showed that the regulation of gene expression varies extensively among fetal tissues with skeletal and muscular system development and function, tissue development, and hematological system development and function having higher gene expression compared with other organ systems.26 In our study using umbilical cord blood, we also noted differential gene expression in genes related to tissue development and hematological system development. Previous studies using amniotic fluid supernatant made use of customized nanostring panels or microarrays that had limited gene coverage. Our data using RNA-seq suggest that preterm infants in utero have lower gene expression specific to immune/inflammatory response at the time of sPTB. The question then arises whether this results in an inability to mount an immune response or contributes to PTB, which cannot be ascertained by our study.

There are multiple strengths to using RNA-seq from umbilical cord blood compared with using cff RNA from amniotic fluid supernatant to understand fetal disease processes. These include (1) less degradation of RNA compared with amniotic fluid supernatant and higher quality of cellular RNA compared with cff RNA; (2) ability to identify alternative splicing; and (3) identification of specific biological pathways from comprehensive interrogation of the genome. Given the ability to interrogate the whole transcriptome robustly with RNA-seq of umbilical cord blood, RNA-seq methods have the potential to identify dysregulated genes specific to certain pregnancy conditions as described in the current study with more sensitive detection of genes, transcripts, and differential expression compared with previously published methods using cff RNA and microarrays.

This study also sheds light on in utero developmental factors. Compared with term, preterm infants have higher cell cycle and metabolism gene expression, suggesting a metabolic focus on growth and development. Other studies confirmed differential gene expression related to GA with term fetuses showing enrichment of salivary gland, tracheal, and renal transcripts as compared with brain and embryonic neural cells in the second trimester.27 Functional analysis of genes upregulated at term in amniotic fluid supernatant revealed pathways that were highly specific for postnatal adaptation such as immune function, digestion, respiration, carbohydrate metabolism, and adipogenesis.21 Similar to the current study using cord blood, inflammation, a key process involved in normal labor, has been also shown to increase in studies using term amniotic fluid.21

There are other intriguing findings from this study to highlight. We did not identify any increase in gene expression related to the neonatal hypothalamic pituitary adrenal (HPA) axis gene expression in the sPTB cohort. Corticotropin-releasing hormone (CRH), a hypothalamic neuropeptide that regulates HPA axis activity, is integral to the regulation of labor. There is a premature surge of CRH levels in the amniotic fluid, placenta, and fetal membranes among pregnancies of women who deliver preterm. Fetal HPA axis activation is a potential therapeutic target pathway for sPTB prevention. However, we found no alteration in gene expression in HPA gene expression in preterm versus term neonates. All of the preterm infants received betamethasone prior to delivery, which has also been shown to suppress fetal HPA axis function.28–30 We also found racial/ethnic differences in gene expression related to immune-mediated cell signaling. Racial disparities exist in neonatal sepsis, with infants of AA women having higher rates of neonatal sepsis at term and preterm compared with infants of NHW mothers.31,32 Cell composition analysis showed that white blood cell gene expression also differed in both term and preterm AA infants compared with NHW infants. These findings warrant further investigation and we hypothesize that epigenetic mechanisms may be responsible for this discrepant pattern.

There are limitations of the current study. The small sample size allowed us enough power to detect differential gene expression between dichotomous groups but we were not powered to detect differences by race or fetal sex. Despite this, we showed evidence of differential gene expression by race. However, it is possible that this difference between races is a false-positive result (despite the use of statistical methods to minimize the chance of FDRs) due to normal variations in the human population. A larger sample size with validation of results by reverse transcription-polymerase chain reaction (RT-qPCR) is needed to further study this finding. Because we recruited patients in real clinical situations, fetuses born were exposed to antenatal corticosteroids which can change fetal gene expression.33 However, the data on fetal gene expression changes with antenatal corticosteroids are limited to murine and in vivo studies with minimal data on humans.33 It is also possible that chorioamnionitis was present despite no clinical signs of chorioamnionitis. Half of the preterm placenta showed evidence of chorioitis and funisitis. However, this is neither a sensitive nor specific screening test for chorioamnionitis. To the best of our ability, we limited variability in the cohorts which could alter gene expression. However, a limitation inherent in the study design was differences in GAs. This may contribute to some of the changes that were observed in this study, which
may be caused by different developmental stages. However, for practical and ethical reasons, relative matching was considered to be the most reasonable alternative for our study. Other authors have overcome this limitation by comparing the human placental transcriptome to the macaque placental transcriptome at equivalent GAs because macaques are closely related to humans. Another limitation was that we included women who had premature rupture of membranes (PROM) prior to onset of labor. Half of the preterm cohort had PROM and the other half had spontaneous labor. Because we enrolled patients who were in spontaneous labor with or without PROM with 50% of the sPTB cohort delivering after PROM, there was heterogeneity in the cohort. Further work to study differences may be accomplished with use of noninvasively obtained cff RNA in a larger cohort of patients.

In summary, RNA-seq of umbilical cord blood leukocytes can identify novel insights in neonates born via sPTB and TB. PTB prevention remains elusive because the mechanisms driving spontaneous preterm labor are poorly understood. Identification of differences in gene expression between preterm and term fetuses will lay the groundwork for future prospective studies to delineate mechanisms leading to sPTB, ultimately allowing development of novel prevention strategies. For example, if decreased immune function in preterm fetuses is a contributing factor to sPTB, novel strategies can be developed to bolster the fetal immune system. These novel pharmacologic interventions can be tested in animal models. This is the first of many steps necessary to better understand the mechanism of sPTB and ultimately lead to effective preventive strategies.

Statement of Financial Support
BIRCWH K12 HD001441 funded by NICHD (PI: Boggess); NICHD R03 HD080788 (PI: Vora); NICHD K23 HD088742 (PI: Vora); NCATS NIH UL1TR001111. The National Institutes of Health had no involvement in the study design, collection of specimens, analysis, interpretation of the data, writing of manuscript, or decision to submit the article for publication.

Conflict of Interest
The authors declare that there is no conflict of interest.

Acknowledgments
None.

References
11 Practice bulletins No. 139: premature rupture of membranes. Obstet Gynecol 2013;122(04):918–930
13 Bi R, Liu P. Sample size calculation while controlling false discovery rate for differential expression analysis with RNA-sequencing experiments. BMC Bioinformatics 2016;17:146
34 Eidem HR, Rinker DC, Ackerman WE IV, et al. Comparing human and macaque placental transcriptomes to disentangle preterm birth pathology from gestational age effects. Placenta 2016;41:74–82