creative commons
COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

연세대학교
YONSEI UNIVERSITY

Identification of a novel gene signature in second-trimester amniotic fluid for the prediction of preterm birth

Min-A Kim

Department of Medicine

The Graduate School, Yonsei University

연세대학교
YONSEI UNIVERSITY

연세대학교
YONSEI UNIVERSITY

Identification of a novel gene signature in second-trimester amniotic fluid for the prediction of preterm birth

Min-A Kim

Department of Medicine

The Graduate School, Yonsei University

연세대학교
YONSEI UNIVERSITY

Identification of a novel gene signature in second-trimester amniotic fluid for the prediction of preterm birth

Directed by Professor Young Han Kim

The Doctoral Dissertation submitted to the Department of Medicine, the Graduate School of Yonsei University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Min-A Kim

June 2020

연세대학교
YONSEI UNIVERSITY

This certifies that the Doctoral Dissertation of Min-A Kim is approved.

Thesis Supervisor : Young Han Kim

Thesis Committee Member#1 : Jae-Hoon Kim

Thesis Committee Member#2 : Jong Rak Choi

Thesis Committee Member#3: Hyoung-Pyo Kim

Thesis Committee Member#4: Hae Sung Won

The Graduate School
Yonsei University

June 2020

연세대학교
YONSEI UNIVERSITY

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor Prof. Kim, Young Han for giving me the opportunity to carry out my PhD thesis. I appreciate his contributions of time and ideas to make my work productive and stimulating. Without his incredible patience and timely wisdom, I would not have been able complete this thesis. A particular debt of gratitude goes also to my other supervisor, Prof. Jae-Hoon Kim. His guidance and supervision have been essential during dissertation phase of the present work, and his patience and constructive comments during my laboratory work made this thesis complete. In addition, I express my appreciation to Prof. Jong Rak Choi for providing feedback on my research. His thoughtful questions and comments were valued greatly. I would also like to thank Prof. Hyoung-Pyo Kim for dedicated guidance and support of my coursework and academic research. In addition to this, my appreciation also goes to Prof. Hye Sung Won for her intellectual influence, constructive feedback, and sharp supervision of my study. This research would not have been possible without the support and encouragement of my family. I would like to especially thank my parents for providing emotional support and continuous love in having conducted this research.

연세대학교
YONSEI UNIVERSITY

<TABLE OF CONTENTS>

ABSTRACT ····· 1

I. INTRODUCTION ····· 3
II. MATERIALS AND METHODS ····· 7
1. Patients and amniotic fluid collection ····· 7
2. Cell free circulating nucleic acid extraction ····· 7
3. Sequencing analysis ····· 8
4. Sequencing read mapping and gene expression signature analysis ···· 8
5. Gene viability assay ····· 9
6. Lipopolysaccharide treatment and glucose derivation ····· 9
7. Quantitative Real-Time Polymerase Chain Reaction ····· 10
8. Statistical analysis ····· 10
III. RESULTS ····· 12
1. Clinical characteristics of the study population ····· 12
2. Identification of differentially expressed genes between preterm birth and term birth by sequencing analysis ····· 13
3. qRT-PCR based validation of differentially expressed genes in the extended samples ····· 17
4. Functional GO analysis and KEGG analysis of differentially expressed genes ····· 19
5. Validation of the candidate genes in H_2O_2 treated human trophoblast cell line HTR-8/SVneo and JEG-3 using qRT-PCR ····· 21
6. Effect of liposaccharide and glucose depriviation on the expression of candidate genes ····· 23
IV. DISCUSSION ····· 24
V. CONCLUSION ····· 28

연세대학교
YONSEI UNIVERSITY

REFERENCES ····· 29
ABSTRACT(IN KOREAN) ····· 35

연세대학교
YONSEI UNIVERSITY

LIST OF FIGURES

Figure 1. Flow chart of normalization procedures and differentially expressed genes in each normalization method …… 14

Figure 2. Venn diagram and heat-map comparison of differentially expressed transcripts from non-normalization, quantile, RPM, RPKM normalization methods …… 15

Figure 3. qRT-PCR validation of the expression of 5 chosen differentially expressed genes in the extended samples ····· 17

Figure 4. Volcano plot of sequencing data …… 19

Figure 5. DAVID functional GO analysis and KEGG analysis of differentially expressed genes…… 20

Figure 6. Time- and dose-dependent cytotoxic activity in the HTR-8/SVneo and JEG-3 cell lines …… 21

Figure 7. Validation of the candidate genes in H_2O_2 treated human trophoblast cell line HTR-8/SVneo and JEG-3 using qRT-PCR …… 22

Figure 8. Effect of liposaccharide and glucose depriviation on the expression of the candidate genes in HTR-8/SVneo cells 24

연세대학교
YONSEI UNIVERSITY

LIST OF TABLES

Table 1. Primer sequences for Quantitative Real-Time Polymerase Chain Reaction ····· 11
Table 2. Clinical characteristics of the study population ····· 12
Table 3. The candidate gene markers for the preterm birth ··· 16
Table 4. Validation data of selected differentially expressed genes by qRT-PCR ····· 18

연세대학교
YONSEI UNIVERSITY

ABSTRACT

Identification of a novel gene signature in second-trimester amniotic fluid for the prediction of preterm birth

Min-A Kim

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Young Han Kim)

Preterm birth (PB) affects approximately 5% to 7% of live births worldwide and is the leading cause of neonatal morbidity and mortality. Amniotic fluid supernatant (AFS) contains abundant cell-free nucleic acids (cf NAs) that can provide genetic information associated with pregnancy complications. In the current study, cf NAs of AFS in early second-trimester before the onset of symptoms of preterm birth were analyzed and we compared the differentially expressed genes (DEGs) between PB (n=5) and term birth (TB, n=5) groups. Sequencing analysis was performed on Illumina Nextseq 500 platform. We performed qRT-PCR based validation of the candidate genes using 21 second-trimester AFS samples from the preterm birth group and 40 from the term birth group. To explore the association between oxidative stress and preterm birth, we further compared the mRNA expression of candidate gene in H2O2-treated human trophoblast cell line HTR-8/SVneo and choriocarcinoma JEG-3. As a result, differential expression analyses detected 43 genes with increased and 11 genes with decreased expression in

연세대학교
YONSEI UNIVERSITY

the future PB group compared to normal term birth. Among them, upregulated expressions of RDH14, ZNF572, VOPP1, SERPINA12, and TCF15 were validated in an extended sample by quantitative PCR (PB, n=21; TB, n=40). To further characterize whether the 5 selected candidate genes were associated with the oxidative mechanism as a detrimental factor in preterm birth pathology, we quantified their mRNA expression levels in HTR-8/SVneo and JEG-3 cell lines with H_2O_2 treatment. Five candidate genes displayed a significant increase in mRNA expression in HTR-8/SVneo cell with H_2O_2 treatment, whereas not in JEG-cell resistant to H_2O_2 treatment. Also, the expression of five candidate genes was increased to more than 2-fold by pretreatment with lipopolysaccharide in HTR-8/SVneo cells. Alteration of gene expression in early second-trimester related to oxidative stress may be helpful for early identification of women at risk of having PB. Identification of differences in gene expression between PB and TB before the onset of clinical symptoms will lay the groundwork for future prospective studies to delineate mechanisms leading to PB.

Key words : preterm birth, cell-free nucleic acid, nucleic acid sequencing, oxidative stress

연세대학교
YONSEI UNIVERSITY

Identification of a novel gene signature in second-trimester amniotic fluid for the prediction of preterm birth

Min-A Kim

Department of Medicine

The Graduate School, Yonsei University

(Directed by Professor Young Han Kim)

I. INTRODUCTION

Preterm birth, defined as birth before the completion of the 37 weeks of gestation, is the leading cause of neonatal morbidity and mortality. Despite recent advances in neonatal care, the burden of disease in preterm birth remains significant, increasing the risk of developmental and medical disabilities.[1] However, the genetic and molecular characteristics of preterm birth remain elusive and preventive strategies have not yet been established. One reason is the wide array of etiologies, coupled with diverse molecular pathways, which makes it difficult to reproduce the clinically

연세대학교
YONSEI UNIVERSITY

similar phenotype of preterm birth. Preterm birth is the result of an interaction of both genetic and environmental factors, and it constitutes a complex and multi-factorial phenotype accompanied by numerous gestational tissues for the induction of parturition.[2-5]

Amniotic fluid (AF) is a complex and dynamic environment that is essential for fetal well-being. AF contains nutrients and growth factors for fetal development and is a source of cells for the prenatal diagnosis of chromosomal abnormalities and fetal infections, as well as for the determination of fetal lung maturity.[6,7] While the cells present in the amniotic fluid are utilized in these diagnostic methods, the remaining AF supernatant (AFS) is usually discarded. However, the AFS contains abundant cell-free nucleic acids (cf NAs), a uniquely dispersed form of genetic material. Cf NAs may be derived from active cellular secretions via exosomes and shedding vesicles, or from microparticles through apoptosis or necroptosis.[8,9] Apoptosis results in systematic cleavage of chromosomal DNA into fragments of 160–180 base pairs (bp), leading to the extracellular presence of mono- (approximately 166 bp) and poly-nucleosomes (332 bp, 498 bp). Necrosis causes clumping of nuclear chromatin and random digestion, producing cf DNA fragments larger than 10,000 bp. Cf DNA derived from active cellular secretions has been shown to range from 1000 to 3000 bp, whereas cf DNA derived from extrachromosomal circular DNA ranges from 30 to 20,000 bp.[10-14]

Cell-free fetal nucleic acids (cff NAs) have become the subject of intense interest in prenatal diagnosis since Lo et al. found their presence in maternal blood in 1997.[15] Cff NAs in maternal plasma are now established as molecular diagnostic materials for noninvasive prenatal testing and have been routinely used in clinical care. The AF-

연세대학교
YONSEI UNIVERSITY

derived cf NAs originate from the fetus, and the AFS is not contaminated by maternal nucleic acids due to unidirectional fetal–maternal circulation.[16-19] Moreover, the AF directly contacts the placenta and may also include cf NAs originating from placental tissues.[20] Cf NAs can serve as intermediate messengers to convey genetic information and cellular signals, which affect various cellular responses such as oxidative stress and immune response.[9] Recently, molecular studies using next-generation sequencing have suggested that cf NAs obtained from AFS could serve as a good diagnostic tool for the detection of abnormal fetal conditions such as fetal aneuploidies,[21-24] fetal growth restriction, twin to twin transfusion syndrome (TTTS)[25] and altered molecular pathways. However, despite substantial research on cf NAs in fetal development, further studies are needed on AFS-derived cf NAs in relation to preterm birth.

Through advancements in sequencing technology and large scale genomic profiling in both clinical and research fields, simple methods to generate comprehensive and accurate whole-genome and transcriptome sequencing data will become increasingly valuable. Until now, RNA-sequencing analysis provides detailed information of the transcriptome while enabling a novel RNA transcript variation to be detected. Recently, the integration of DNA and RNA analysis has shown promising results in the detection of fusion transcripts and alternative transcripts. These sequencing data can provide useful information to explore genetic variation underlying human disease and normal phenotypic variability.[26]

Nucleic acid sequencing is an advanced technology for genome analysis that enables a comprehensive characterization of gene expression. Most gene expression data regarding human pregnancy are based on microarrays,[27-29] and most of these are focused

연세대학교
YONSEI UNIVERSITY

on the transcriptome of preeclampsia, not on preterm birth. In addition, published studies on human labor using next-generation sequencing (NGS) have been confined to the placenta or blood. Until now, many studies have focused on maternal factors, such as inflammation, infection, and maternal gestational tissues, in studying the causes of preterm birth; however, little is known about the fetal contribution to preterm birth.[30-33] Although maternal factors are known to contribute to preterm birth risk, there are emerging data that variations in the fetal gene expression, not the maternal, may induce preterm labor. If sequencing technologies can analyze the differentially expressed genes (DEGs) in preterm birth to identify unique gene signatures at the fetal molecular level, it will help to understand more about the fetal contribution to pregnancy-specific diseases such as premature birth. Further, it will lay the groundwork for preventive, diagnostic, and therapeutic strategies in obstetrics and pediatrics. Overall, both maternal and fetal genomes can simultaneously or separately contribute to the occurrence of preterm birth, influenced by the environmental factors.

The aim of this study was to characterize the gene expression signatures from AFS during second-trimester pregnancy destined to conclude with a preterm birth and to reveal the biological information about genes involved in the pathogenesis of preterm birth by performing sequencing analysis using cf NAs present in the AFS. We hypothesized that the AFS is a more diverse source of cf NAs that might be able to provide real-time information on preterm birth along with nucleic acid from amniocytes.

연세대학교
YONSEI UNIVERSITY

II. MATERIALS AND METHODS

1. Patients and amniotic fluid collection

AF samples were obtained from 74 pregnant women who underwent amniocentesis for fetal karyotyping and genetic diagnosis at 16 to 19 weeks of gestation in the Department of Obstetrics and Gynecology, Gangnam Severance Hospital (Seoul, Korea) between March 2011 and May 2017. This study was reviewed and approved by the Institutional Review Board of Gangnam Severance Hospital. All participants provided written informed consent to participate in the study before AF was obtained. For all subjects, gestational age was assessed during early gestation using the crown-rump length measurement by transvaginal ultrasonography. We included only cases of spontaneous preterm birth and excluded all pregnancies complicated by preeclampsia, major fetal anomalies, abnormal karyotypes, and gestational diabetes mellitus. We defined control subjects as those who had normal term births, defined as a live birth from 37 to 42 weeks of gestation.

2. Cell-free circulating nucleic acid extraction

AF samples were centrifuged at 350 x *g* for 10 min at 4°C. The supernatant samples were spun at 1500 rpm for 10 min at 10°C to remove residual vernix and then stored at -80°C. Cf NAs were extracted from 1 mL of AFS using the QIAamp Circulating Nucleic Acid (Qiagen, Germany) kit according to the manufacturer's protocol and eluted with RNase-free water. The concentration of cf NA in each sample was measured using a NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies Inc., Wilmington, DE, USA). Samples with an A260/A280 ratio greater than 2.0 were stored at -70°C until further analyses.

연세대학교
YONSEI UNIVERSITY

3. Sequencing analysis

The quantity and quality of the extracted nucleic acids were assessed using an Agilent 2100 Bioanalyzer RNA chip (Agilent Technologies, Santa Clara, CA, USA) according to the manufacturer's instructions. Sequencing libraries were prepared using the SMARTer Stranded Total RNA-Seq kit-v2 Pico Input Mammalian (Takara Bio USA, Mountain View, CA, USA) and validated using a DNA 1000 chip on an Agilent 2100 Bioanalyzer (Agilent Technologies) according to the manufacturer's instructions. Nucleic acid sequencing was performed using the Illumina NextSeq 500 platform (Illumina, Inc., San Diego, CA, USA) with a 75-nucleotide paired-end indexed run.

4. Sequencing read mapping and gene expression signature analysis

Nucleic acid sequencing reads were mapped to the human genome reference (Homo sapiens GRCh38 dbSNP150genome template, obtained from DNASTAR, Inc., Madison, USA). The read counts for each gene were calculated and analyzed using three different normalizations, that is, quantile, reads per million (RPM), and reads per kilobase million (RPKM), and non-normalization methods for differential expression detection. Gene transcript numbers comparing preterm birth versus term birth with a false discovery rate (FDR) P-value <0.1 or adjusted P-value <0.05 were considered to be significantly different. Differentially expressed genes (DEGs) were retained if they were detected by at least two of the methods used. All analyses were performed using DNASTAR Lasergene 15 software. Common markers were identified using a heat map with hierarchical clustering and Venn Diagram analysis. Functional enrichment analysis

연세대학교
YONSEI UNIVERSITY

of the DEGs was performed based on GO terms and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway annotations. GO term analyses were performed using the DAVID database (https://david.ncifcrf.gov/) according to P-value (-log10), which is an essential tool for successful high-throughput gene function analysis.

5. Cell viability assay

JEG-3 and HTR-8/Svneo cell lines were plated onto 24-well plates at densities of approximately 8.0×10^4 cells and 10.0×10^4 cells per well, respectively, in 0.5 ml of growth media. After 24 h, the medium was replaced with DMEM containing 10% FBS. Meanwhile, H_2O_2 (0, 1, 10, 25, 50, 100, and 250 μM) was added to the corresponding wells and incubated at 37°C for 24 h or 48 h. For the crystal violet staining assay, cells were fixed with fixation solution (10% acetic acid, 10% methanol, 80% H_2O),stained with 0.5% crystal violet staining solution with 20% methanol for 1 h, photographed, and extracted using 1% SDS solution. The absorbance was measured at 595 nm using a VersaMax microplate reader (Molecular Devices, Sunnyvale, CA). All experiments were performed in triplicates.

6. Lipopolysaccharide treatment and glucose deprivation

HTR-8/SVneo cells were seeded at 4×10^5 cells in 6 well plates and treated with 10 ng/ml and 100 ng/ml lipopolysaccharide (LPS, Sigma-Aldrich, St. Louis, MO, cat. L4391) for 24 h. For glucose deprivation, the complete DMEM medium in culture plates was replaced with glucose-free DMEM medium (Thermo Fisher Scientific, Rockford, IL, USA, cat. 11966-025) and cells were cultured for 24 h.

연세대학교
YONSEI UNIVERSITY

7. Quantitative Real-Time Polymerase Chain Reaction

Total RNA (50 ng) from each sample was reverse-transcribed into cDNA using Maxima First Strand cDNA Synthesis Kit (Thermo Scientific, Waltham, MA) according to the manufacturer's protocol. Real-time polymerase chain reaction (PCR) was performed to quantify mRNA expression using SYBR Green PCR Master Mix (Enzynomics, Daejeon, Republic of Korea) and an ABI PRISM 7300 real-time PCR system (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. Relative mRNA expression was quantified using the comparative Ct (ΔCt) method and expressed as 2-ΔΔCt, where ΔΔCt =ΔE–ΔC,ΔE = CtE target–CtE GAPDH, and ΔC = CtC target–CtC GAPDH (E = experimental result and C = controls). Each assay was performed in triplicate and expressed as the mean±standard error (SE). Serial dilutions were prepared from a stock solution of total RNA to generate a standard curve and determine reaction efficiencies. The primers used for PCR are listed in Table 1.

8. Statistical analysis

Experimental results were statistically evaluated with a two-tailed paired Student's t-test using GraphPad Prism 7. All tests of significance were set at $P < 0.05$.

연세대학교
YONSEI UNIVERSITY

Table 1. Primer sequences for Quantitative Real-Time Polymerase Chain Reaction

Marker name	Gene bank	PCR product size	Forward sequence	Reverse sequence
RDH14	NC_000002.12	70bp	TCTTCTGAACAAATGCCCTCTGA	TCCGTTGAGATTGAGGTAGAAGAGT
SERPINA12	NC_000014.9	70bp	TCCATCTTCAGCTCAGCCTTGT	AGGAGCTTGGCAGACCTTGA
VOPP1	NC_000007.14	70bp	GAGTCTACGGGCCACGATTTAG	CACAGCGACCACCTGTTGTT
ZNF572	NC_000008.11	77bp	ATTCAGCAGCAGCTCTCACCTTA	CGCAGACAGAACATTCATATGGTT
TCF15	NC_000020.11	70bp	AGGTGATCCCATCTCTGATGCT	CCTTCAGGCAGGTAGTTTTTCTG
LOC105374432	NC_000004.12	80bp	CTCGTCTGTTCTCATGCTGCTAA	TGTGAGTCCATTAAACCTCTTTCCT
GAPDH	NM_001256799	292bp	CCTGACCTGCCGTCTAGAAA	GGTGGTCCAGGGGTCTTACT

연세대학교
YONSEI UNIVERSITY

III. RESULTS

1. Clinical characteristics of the study population.

We obtained 74 AF samples from pregnant women at 16 to 19 weeks of gestation. Among them, 24 women had a preterm delivery and 50 delivered at term. The average amounts of total nucleic acid in the preterm birth and term birth groups were 219.2±102.1 ng and 151.3±69.27 ng, respectively, which were significantly different. To pursue sequencing analysis, we paired cases and controls based on the amount of total nucleic acid and selected 5 of the 21 preterm birth and 5 of the 40 term birth samples. The demographic and clinical characteristics of the study population are presented in Table 2.

Table 2. Clinical characteristics of the study population

Sample	Age (years)	GA at amniocentesis (weeks)	Indication of amniocentesis	GA at delivery (weeks)
P1	39	16+3	AMA	35+0
P2	37	17+2	AMA	35+4
P3	33	16+5	Positive screening test for Down syndrome	36+6
P4	35	17+3	Positive screening test for Down syndrome	36+3
P5	32	18+4	Positive screening test for Down syndrome	36+4
N1	31	16+2	Positive screening test for Down syndrome	40+4
N2	35	17+2	Positive screening test for Down syndrome	40+1
N3	39	16+0	AMA	40+5
N4	41	18+2	AMA	40+0
N5	39	17+4	Positive screening test for Down syndrome	40+0

GA, gestational age; AMA, advanced maternal age; NTD, neural tube defect

In sample names, the prefix P is used for preterm birth and N for normal term birth.

연세대학교
YONSEI UNIVERSITY

There were no significant differences in maternal age (35.2±2.9 vs 37.0±4.0), body mass index (22.8±3.4 vs 19.7±2.1), or gestational age at amniocentesis (17.3±0.8 vs 17.1±0.9) between preterm birth and term birth groups, whereas a significant difference was observed between both groups for gestational age at delivery (35.5±1.3 vs 40.3±0.3).

2. Identification of differentially expressed genes between preterm birth and term birth by sequencing analysis

To examine the gene expression signatures in women destined to have a preterm birth, we performed nucleic acid sequencing on second-trimester AF samples and compared the expression patterns of DEGs between the preterm and term birth groups. Because sequencing technologies often lead to unexpected experimental errors in the data, normalization methods are commonly used to remove systematic experimental bias and technical variation while preserving biological variation. However, although various normalization approaches have been proposed, there is no standardized manner in the biofluid. Therefore, we compared the performance of three normalization methods, i.e., Quantile, RPM, and RPKM normalization, in addition to the non-normalization method. Figure 1 shows the flow chart of normalization procedures for each of the five AFS samples in preterm birth and term birth groups, and the number of DEGs in each normalization method. It showed quite similar results of four overlapping DEGs (with FDR, $P < 0.1$) in the different normalization methods. We generated a heat map of the DEGs for preterm birth samples with respect to term birth samples.

연세대학교
YONSEI UNIVERSITY

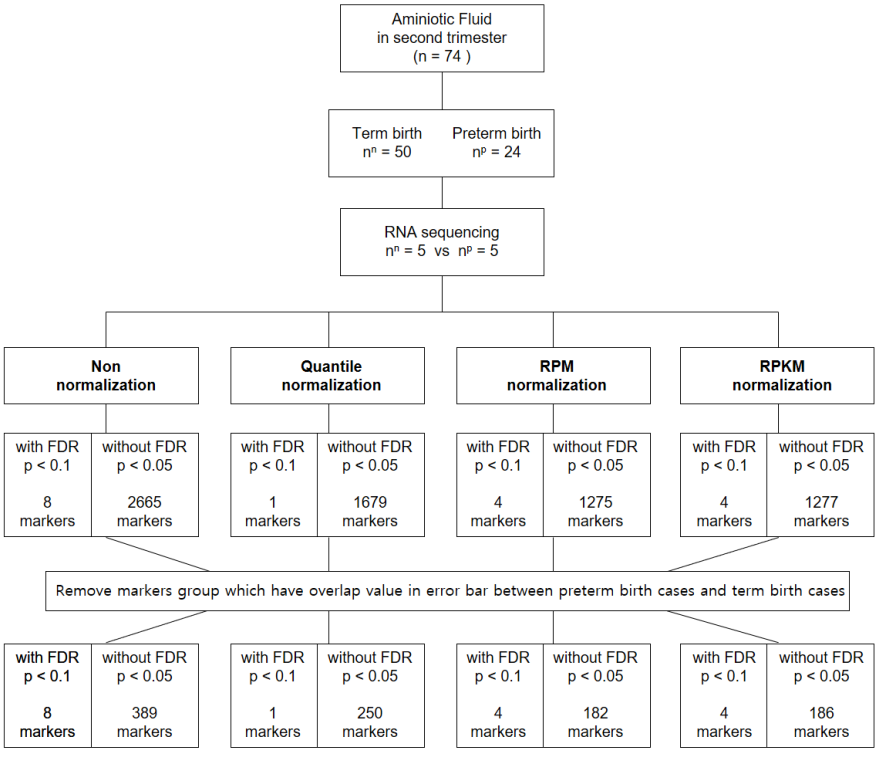

Figure 1. Flow chart of normalization procedures and differentially expressed genes in each normalization method

We found unique gene expression profiles from each of these four different normalization methods by pairwise comparisons. The overlap in the results from each normalization method is illustrated in Figure 2. Gene expression was considered significant if the FDR adjusted P-value was <0.01 or P-value <0.05. From the nucleic acid sequencing analysis, we obtained altogether 54 genes that were significantly differentially expressed (DE) between the preterm birth and term birth groups. Among

연세대학교
YONSEI UNIVERSITY

these genes, a total of 4 genes, MIR6801, SNORA108, MIR4749, and LOC105374432 were most commonly DE according to the 4 methods applied with FDR P < 0.1, and all 4 genes were upregulated. A total of 50 genes were DE with an adjusted P < 0.05. Among the 50 DE genes, 39 were upregulated and 11 were downregulated in the preterm birth relative to the term birth groups.

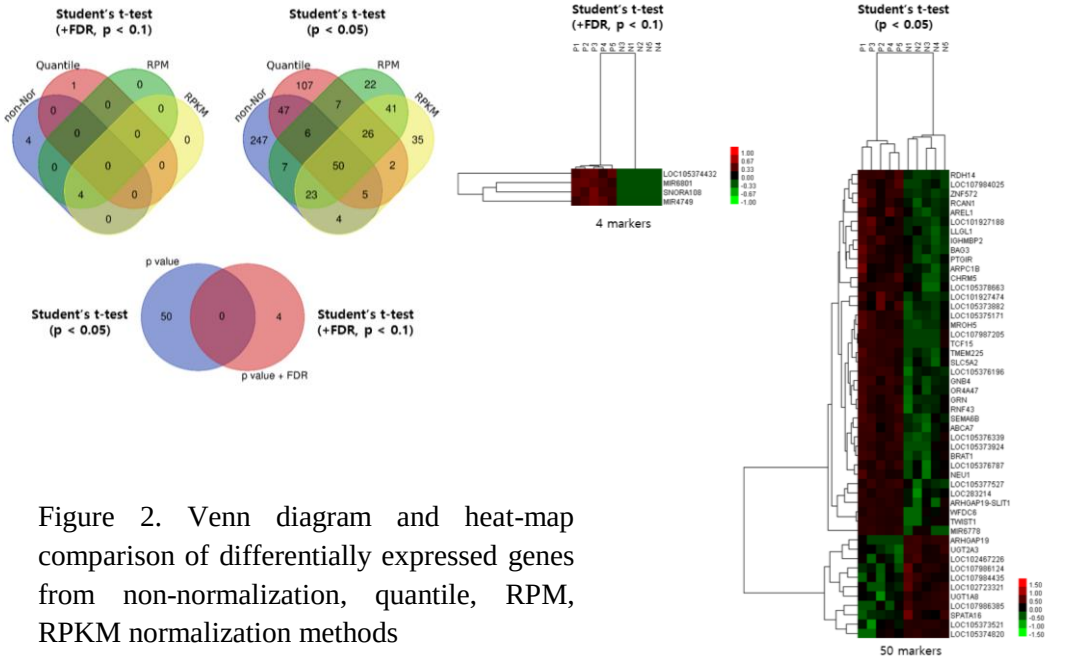

Figure 2. Venn diagram and heat-map comparison of differentially expressed genes from non-normalization, quantile, RPM, RPKM normalization methods

The full list of DE genes detected (P < 0.05) is shown in Table 3. The candidate markers were miRNAs, SNORA, and uncharacterized genes (LOC105374432), including the ordinary genes.

연세대학교
YONSEI UNIVERSITY

Table 3. The candidate gene markers for the preterm birth

Non-normlization						Quantile normalization						RPM normalization						RPKM normalization					
Name	Preterm Birth	Normal Birth	P value (+ FDR)	P value (- FDR)	Fold change	Name	Preterm Birth	Normal Birth	P value (+ FDR)	P value (- FDR)	Fold change	Name	Preterm Birth	Normal Birth	P value (+ FDR)	P value (- FDR)	Fold change	Name	Preterm Birth	Normal Birth	P value (+ FDR)	P value (- FDR)	Fold change
Sorting by p value (+FDR, p < 0.1)																							
MIR6801	2.605	0.021	0.00357	1.47E-07	125.048 up	MIR4420	0.084	1.326	0.00824	2.17E-07	15.836 down	MIR6801	1.399	0.009	0.00215	9.94E-08	163.579 up	MIR6801	62.539	0.024	0.000146	3.84E-09	2582.261 up
LOC105374432	2.759	0.021	0.00357	1.88E-07	132.454 up							SNORA108	0.814	0.009	0.00215	1.13E-07	95.172 up	SNORA108	6.119	0.024	0.000438	2.45E-08	252.649 up
SNORA108	1.516	0.021	0.00387	3.05E-07	72.754 up							LOC105374432	1.482	0.009	0.00383	3.02E-07	173.267 up	MIR4749	49.434	0.024	0.000438	3.45E-08	2041.176 up
MIR4749	1.974	0.021	0.00471	4.95E-07	94.768 up							MIR4749	1.06	0.009	0.00846	8.89E-07	123.969 up	LOC105374432	6.704	0.024	0.00146	1.54E-07	276.811 up
Sorting by p value (p < 0.05)																							
RDH14	29.913	10.005	0.026	3.41E-06	2.989 up	ARHGAP19	0.153	31.974	0.407	0.0000321	209.073 down	ARHGAP19	0.021	17.484	0.68	0.0000893	826.915 down	ARHGAP19	0.045	12.01	0.215	0.0000282	265.690 down
ZNF572	54.524	31.753	0.0784	0.0000145	1.717 up	LOC105375171	29.063	19.499	0.831	0.000127	1.490 up	RDH14	16.062	5.822	0.681	0.000108	2.758 up	RDH14	15.886	5.758	0.68	0.000108	2.759 up
ARHGAP19	0.052	30.044	0.39	0.000148	578.813 down	RDH14	27.736	11.153	0.856	0.000235	2.486 up	LOC102723321	17.767	40.544	0.874	0.000161	2.282 down	LOC102723321	13.923	31.767	0.876	0.000161	2.281 down
GNB4	21.803	10.027	0.39	0.000154	2.174 up	GNB4	19.689	11.808	0.856	0.000263	1.667 up	LOC107987205	2.724	0.02	0.953	0.00041	137.366 up	CHRM5	17.946	10.813	0.961	0.000786	1.659 up
LOC107984025	9.509	2.491	0.505	0.000239	3.816 up	ZNF572	51.559	34.46	0.856	0.000293	1.496 up	IGHMBP2	49.415	28.428	0.953	0.000505	1.738 up	IGHMBP2	17.229	10.172	0.961	0.000902	1.693 up
LOC105375171	31.405	17.36	0.571	0.000318	1.809 up	PTGIR	99.997	68.62	0.938	0.000385	1.457 up	CHRM5	28.698	17.294	0.953	0.000783	1.659 up	RCAN1	42.808	21.374	0.961	0.000934	2.002 up
LOC101927474	19.199	7.039	0.643	0.000703	2.727 up	LOC102723321	30.668	73.252	0.938	0.000561	2.388 down	TMEM225	15.52	6.289	0.953	0.000978	2.467 up	TCF15	4.747	0.058	0.961	0.00103	81.812 up
TMEM225	28.903	10.807	0.643	0.000719	2.674 up	GRN	39.523	19.507	0.938	0.000794	2.026 up	RCAN1	25.769	12.282	0.953	0.00105	2.098 up	BAG3	23.483	15.57	0.961	0.00106	1.508 up
IGHMBP2	92.028	48.848	0.643	0.00079	1.883 up	TMEM225	26.611	12.151	0.938	0.000906	2.189 up	BAG3	40.568	26.901	0.953	0.00106	1.508 up	TMEM225	29.893	11.961	0.961	0.00132	2.499 up
LOC107987205	6.128	0.052	0.643	0.000916	118.056 up	IGHMBP2	88.784	52.238	0.938	0.00098	1.699 up	TCF15	2.848	0.023	0.953	0.00118	125.042 up	LOC107984025	12.244	3.476	0.961	0.00157	3.522 up
LOC105376196	16.027	4.789	0.643	0.000936	3.346 up	SPATA16	11.873	34.209	0.938	0.00121	2.881 down	PTGIR	55.543	37.744	0.953	0.00139	1.471 up	ARPC1B	15.108	7.762	0.961	0.0018	1.946 up
TCF15	5.305	0.052	0.643	0.001	102.196 up	MROH5	54.677	32.883	0.938	0.00127	1.662 up	LOC107984025	5.106	1.45	0.953	0.00158	3.521 up	PTGIR	28.864	18.52	0.961	0.00194	1.558 up
LOC101927188	50.404	30.457	0.643	0.00104	1.654 up	TCF15	4.461	0.225	0.938	0.00139	19.786 up	UGT2A3	10.113	22.263	0.953	0.00177	2.201 down	LOC107986124	7.792	15.804	0.961	0.00197	2.028 down
SEMA6B	37.363	12.309	0.643	0.00116	3.035 up	LOC107986385	0.21	4.739	0.938	0.00141	22.540 down	ARPC1B	16.907	8.688	0.953	0.0018	1.946 up	LOC107987205	4.463	0.062	0.961	0.00215	71.689 up
MROH5	57.67	29.829	0.643	0.00116	1.933 up	LOC105376196	14.452	5.989	0.938	0.0015	2.413 up	LOC107986124	28.847	58.52	0.953	0.00197	2.028 down	LLGL1	10.155	5.962	0.961	0.0023	1.703 up
RCAN1	47.99	21.104	0.643	0.00119	2.274 up	LOC107986124	50.746	105.374	0.938	0.00157	2.076 down	LLGL1	33.056	19.442	0.953	0.00228	1.700 up	SEMA6B	8.134	2.96	0.961	0.00294	2.748 up
LLGL1	61.562	33.407	0.643	0.0015	1.842 up	UGT2A3	17.084	41.388	0.938	0.00194	2.422 down	AREL1	34.685	21.44	0.953	0.0029	1.617 up	ZNF572	18.412	11.619	0.961	0.00307	1.584 up
GRN	42.373	16.923	0.643	0.00182	2.503 up	LOC105373924	3.632	0.288	0.938	0.00209	12.600 up	ZNF572	29.277	18.479	0.953	0.00308	1.584 up	AREL1	14.816	9.156	0.961	0.0031	1.618 up
AREL1	64.596	36.84	0.643	0.00219	1.753 up	AREL1	61.691	39.648	0.938	0.00215	1.555 up	SEMA6B	20.062	7.163	0.953	0.00309	2.800 up	GNB4	11.443	5.703	0.961	0.00314	2.006 up
LOC105377527	37.925	12.953	0.643	0.00233	2.927 up	SEMA6B	34.721	14.059	0.938	0.00216	2.469 up	GNB4	11.707	5.836	0.953	0.00314	2.006 up	UGT2A3	3.414	7.35	0.961	0.00318	2.153 down
OR4A47	65.459	33.17	0.643	0.00246	1.973 up	CHRM5	50.497	31.926	0.938	0.00256	1.581 up	SPATA16	7.152	18.795	0.953	0.00366	2.627 down	LOC101927474	6.689	2.658	0.961	0.00384	2.516 up
LOC105376339	17.049	5.346	0.643	0.00262	3.188 up	LOC107987205	4.363	0.225	0.938	0.00274	19.352 up	LOC101927474	10.309	4.096	0.953	0.00384	2.516 up	SPATA16	4.374	12.132	0.961	0.00394	2.773 down
SLC5A2	23.883	7.922	0.643	0.00366	3.014 up	LOC105376339	15.424	6.461	0.938	0.00293	2.387 up	LOC105377527	20.364	7.538	0.953	0.00492	2.701 up	LOC107986385	0.116	5.348	0.961	0.00457	45.959 down
ABCA7	96.441	45.702	0.643	0.00395	2.110 up	RCAN1	44.786	23.062	0.938	0.00303	1.941 up	LOC107986385	0.047	2.626	0.953	0.00529	56.026 down	LOC105377527	10.589	3.919	0.961	0.00491	2.701 up
RNF43	76.138	39.694	0.643	0.00426	1.918 up	LLGL1	58.293	36.014	0.938	0.00318	1.618 up	SLC5A2	12.824	4.61	0.953	0.00661	2.781 up	LOC105373924	5.036	0.124	0.961	0.00646	40.539 up
CHRM5	53.446	29.717	0.643	0.00436	1.798 up	ABCA7	93.675	49.486	0.938	0.00356	1.892 up	LOC105373924	2.286	0.049	0.953	0.00712	47.140 up	SLC5A2	6.252	2.271	0.961	0.00717	2.752 up
BAG3	75.552	46.224	0.643	0.00448	1.634 up	LOC101927188	47.543	33.049	0.938	0.00377	1.438 up	ABCA7	51.784	26.597	0.953	0.00803	1.947 up	ABCA7	9.091	4.675	0.961	0.00732	1.944 up
LOC105373924	4.258	0.098	0.643	0.00465	43.448 up	BRAT1	39.227	19.6	0.938	0.00499	2.001 up	MROH5	30.966	17.36	0.953	0.00865	1.783 up	MROH5	8.019	4.488	0.961	0.00852	1.786 up
LOC107986385	0.098	4.513	0.643	0.00467	46.047 down	OR4A47	62.324	37.168	0.938	0.00504	1.676 up	LOC105376339	9.154	3.111	0.953	0.00874	2.942 up	LOC105376339	16.059	5.457	0.961	0.00874	2.942 up
LOC105373882	19.877	7.007	0.643	0.00499	2.836 up	LOC107984435	35.84	62.83	0.938	0.00519	1.753 down	LOC102467226	1.273	4.486	0.953	0.00907	3.523 down	LOC102467226	2.526	8.899	0.961	0.00908	3.522 down
BRAT1	41.794	16.972	0.643	0.00517	2.462 up	SLC5A2	21.972	9.627	0.938	0.00539	2.282 up	LOC105376196	8.606	2.787	0.953	0.00985	3.087 up	LOC101927188	11.792	7.722	0.961	0.00983	1.527 up
LOC107986124	53.723	100.555	0.643	0.00519	1.871 down	BAG3	72.11	49.169	0.938	0.00589	1.466 up	LOC101927188	27.065	17.725	0.953	0.00986	1.526 up	LOC105376196	12.711	4.116	0.961	0.00984	3.088 up
PTGIR	103.442	64.856	0.643	0.00522	1.594 up	LOC105373521	0.706	7.099	0.938	0.00589	10.060 down	UGT1A8	5.173	12.711	0.953	0.00994	2.457 down	ARHGAP19-SLIT1	8.984	0.263	0.961	0.0109	34.140 up
UGT2A3	18.834	38.973	0.643	0.00538	2.069 down	LOC107984025	8.44	3.113	0.938	0.00666	2.711 up	LOC105375171	16.863	10.103	0.953	0.0112	1.669 up	LOC105375171	8.805	5.274	0.961	0.0112	1.669 up
LOC102723321	33.088	69.667	0.65	0.00623	2.105 down	LOC101927474	17.451	8.102	0.938	0.00688	2.154 up	GRN	22.753	9.849	0.953	0.0125	2.310 up	UGT1A8	3.348	8.111	0.961	0.0124	2.422 down
SPATA16	13.32	32.295	0.675	0.00739	2.424 down	LOC105377527	35.508	14.065	0.938	0.00714	2.524 up	MIR6778	3.034	0.062	0.953	0.0132	48.604 up	GRN	12.767	5.525	0.961	0.0125	2.310 up
ARPC1B	31.487	14.928	0.712	0.00954	2.109 up	UGT1A8	9.349	23.774	0.938	0.00742	2.543 down	LOC107984435	20.506	34.733	0.953	0.0191	1.693 down	MIR6778	141.472	20.752	0.961	0.0196	260.021 up
MIR6778	5.65	0.129	0.712	0.00992	43.693 up	RNF43	72.892	44.093	0.938	0.0102	1.653 up	BRAT1	22.442	9.877	0.953	0.0196	2.272 up	LOC107984435	23.588	45.07	0.961	0.0204	1.910 down
LOC105376787	18.532	6.022	0.712	0.0108	3.077 up	LOC102467226	1.861	8.552	0.938	0.0117	4.595 down	ARHGAP19-SLIT1	17.214	0.256	0.953	0.0203	67.147 up	BRAT1	8.041	3.649	0.961	0.0213	2.203 up
LOC102467226	2.371	7.708	0.712	0.0162	3.250 down	LOC105373882	18.204	8.334	0.938	0.0122	2.184 up	OR4A47	35.356	18.891	0.953	0.0203	1.871 up	LOC105378663	13.77	0.455	0.961	0.0236	30.234 up
ARHGAP19-SLIT1	32.059	0.528	0.712	0.0175	60.748 up	TWIST1	6.448	0.883	0.938	0.0153	7.301 up	LOC105376787	9.951	3.505	0.953	0.0238	2.839 up	LOC105376787	13.649	4.806	0.961	0.0238	2.839 up
LOC107984435	38.19	59.681	0.712	0.0176	1.562 down	WFDC6	6.114	0.848	0.938	0.0154	7.210 up	LOC105373882	10.673	4.078	0.953	0.0244	2.617 up	LOC105373882	9.836	3.758	0.961	0.0244	2.617 up
NEU1	37.865	15.096	0.712	0.0178	2.508 up	LOC105374820	0.651	9.534	0.938	0.0155	14.638 down	LOC105378663	6.404	0.19	0.953	0.0258	33.753 up	OR4A47	37.777	20.752	0.961	0.0254	1.820 up
LOC283214	22.959	6.1	0.712	0.0216	3.763 up	LOC105376787	16.919	7.279	0.938	0.018	2.324 up	NEU1	20.332	8.786	0.953	0.0317	2.314 up	NEU1	16.29	7.038	0.961	0.0317	2.314 up
UGT1A8	9.633	22.207	0.712	0.0227	2.305 down	ARHGAP19-SLIT1	30.046	1.179	0.938	0.0183	25.486 up	LOC283214	12.328	3.55	0.953	0.0358	3.472 up	TWIST1	6.553	0.369	0.961	0.0347	17.759 up
LOC105378663	11.926	0.407	0.712	0.0247	29.306 up	MIR6778	4.881	0.51	0.938	0.0203	9.575 up	LOC105373521	0.333	3.693	0.953	0.0369	11.079 down	LOC283214	8.256	2.36	0.961	0.0348	3.498 up
WFDC6	6.999	0.318	0.712	0.0261	22.006 up	ARPC1B	29.042	16.258	0.938	0.0221	1.786 up	WFDC6	3.758	0.174	0.953	0.0399	21.569 up	LOC105373521	0.89	9.716	0.961	0.0357	10.915 down
TWIST1	7.432	0.33	0.712	0.0268	22.530 up	LOC105378663	10.608	0.803	0.938	0.0266	13.216 up	TWIST1	3.991	0.181	0.953	0.04	22.083 up	RNF43	17.345	9.786	0.961	0.0419	1.772 up
LOC105373521	0.66	6.346	0.712	0.0348	9.618 down	NEU1	35.06	17.307	0.938	0.0275	2.025 up	RNF43	40.882	23.101	0.953	0.0419	1.769 up	LOC105374820	0.544	12.069	0.961	0.0422	22.168 down
LOC105374820	0.42	8.6	0.712	0.0413	20.491 down	LOC283214	21.187	7.253	0.938	0.0359	2.921 up	LOC105374820	0.214	5.005	0.953	0.0448	23.424 down	WFDC6	14.399	0.591	0.961	0.0427	24.343 up

연세대학교
YONSEI UNIVERSITY

3. qRT-PCR based validation of differentially expressed genes in the extended samples

The expression levels of the DEGs were confirmed using qRT-PCR. The top 6 upregulated DEGs (LOC105374432, RDH14, ZNF572, VOPP1, SERPINA12, and TCF15) were chosen from the ranked gene lists. For qRT-PCR based validation of the candidate genes, 21 second-trimester AFS samples from the preterm birth group and 40 from the term birth group, with an RNA quantity of more than 10 ng/ul were chosen. Consistent with the sequencing results, all 5 DEGs (RDH14, ZNF572, VOPP1, SERPINA12, and TCF15) were found to be differentially expressed in the preterm birth samples with the same trends ($P < 0.05$; Fig. 3 and Table 4).

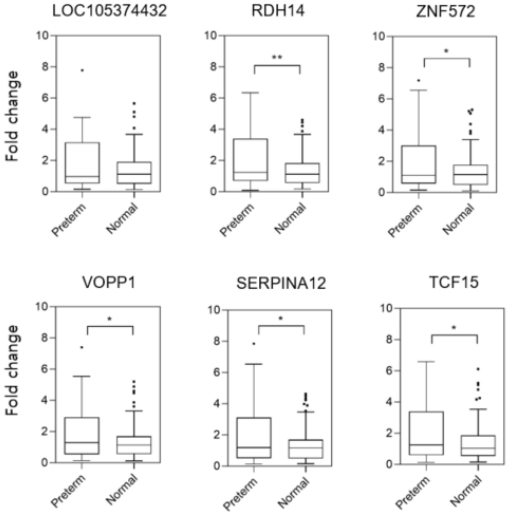

Figure 3. qRT-PCR validation of the expression of 5 chosen differentially expressed genes in the extended samples

Asterisks indicate levels of significance of differential expression (T-test: * $P \leq 0.05$, ** $P \leq 0.01$).

연세대학교
YONSEI UNIVERSITY

Table 4. Validation data of selected differentially expressed genes by qRT-PCR

		LOC105374432		RDH14		ZNF572	
		Preterm	Normal	Preterm	Normal	Preterm	Normal
n		21	40	21	40	21	40
Mean		1.777	1.416	1.98	1.36	1.91	1.435
Std. Deviation		1.699	1.193	1.696	1.044	1.819	1.223
Std. Error		0.2622	0.1333	0.2617	0.1167	0.2807	0.1367
95% CI		1.247 to 2.306	1.150 - 1.681	1.451 to 2.508	1.128 to 1.592	1.343 to 2.477	1.163 to 1.707
unpaired t-test	two-tailed	0.175		0.014*		0.0889	
	one-tailed	0.0873		0.007**		0.0444*	

		VOPP1		SERPINA12		TCF15	
		Preterm	Normal	Preterm	Normal	Preterm	Normal
n		21	40	21	40	21	40
Mean		1.941	1.394	1.963	1.391	2.079	1.437
Std. Deviation		1.778	1.157	1.868	1.13	1.801	1.287
Std. Error		0.2743	0.1293	0.2883	0.1264	0.2778	0.1449
95% CI		1.387 to 2.495	1.137 to 1.652	1.380 to 2.545	1.139 to 1.642	1.518 to 2.640	1.148 to 1.725
unpaired t-test	two-tailed	0.0425*		0.0373*		0.0252*	
	one-tailed	0.0213*		0.0187*		0.0126*	

We used the limma package to calculate the discrepant expression of the genes between the preterm birth and term birth groups. Then, we obtained the $\log_2$ fold change value of each gene according to the data with a linear fit. To overview statistical significance along with differential gene expression, volcano plots were drawn, in which the magnitude of gene expression ratios and the significance of the difference in gene expression between pools were displayed on the x-axis and y-axis, respectively. Red dots indicate the 5 upregulated DEGs (Fig. 4).

연세대학교
YONSEI UNIVERSITY

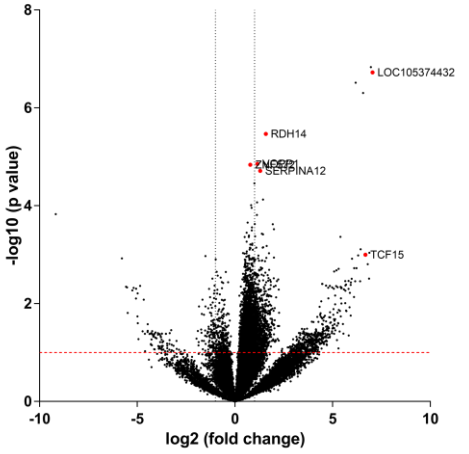

Figure 4. Volcano plot of sequencing data

Raw P-values in negative logarithmic scale on the y-axis as a function of the log2-fold change on the x-axis. Red dots represent the 5 upregulated candidate genes.

4. Functional GO analysis and KEGG analysis of differentially expressed genes

To gain insights into the biological implications of the DEG analysis and examine the patterns of biological pathway dysregulation leading to preterm birth, we analyzed our curated list using the gene annotation enrichment program DAVID. Application of DAVID to the transcriptome of a preterm birth group compared with that of the term birth group showed only a few gene sets with an FDR adjusted P-value <0.1. Based on the DAVID analysis, enriched GO terms were grouped into four categories: biological process, cellular component, molecular function, and KEGG pathway. In the biological process category, 4 terms were significantly enriched, including cellular glucuronidation (GO: 0052695, P=0.024), flavonoid glucuronidation (GO: 0052695, P=0.033), response to estrogen (GO: 0043627, P=0.1), and regulation of catalytic activity (GO: 0050790, P=0.1). The two enriched GO terms under the cellular component category were intracellular membrane-bounded organelles (GO: 0043231, P=0.044) and lysosomal membrane (GO: 005765, P=0.06). With respect to the

연세대학교
YONSEI UNIVERSITY

molecular function category, two terms were significantly enriched, including glucuronosyltransferase activity (GO: 00015020, P=0.04) and transferase activity, transferring hexosyl groups (GO: 0016758, P=0.042). When analyzing KEGG pathways, five terms were significantly enriched in ascorbate and aldarate metabolism (hsa00053, P =0.046), pentose and glucuronate interconversions (hsa00040, P=0.056), porphyrin and chlorophyll metabolism (hsa00860, P=0.071), drug metabolism-other enzymes (hsa00983, P=0.077), and steroid hormone biosynthesis (has 00140, P=0.1) (Fig. 5).

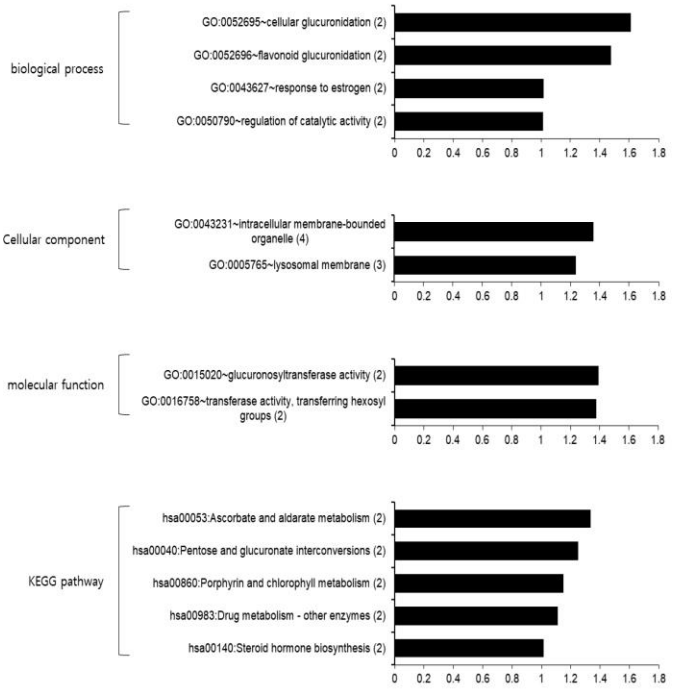

Figure 5. DAVID functional GO analysis and KEGG analysis of differentially expressed genes

연세대학교
YONSEI UNIVERSITY

5. Validation of candidate genes in H_2O_2-treated human trophoblast cell line HTR-8/SVneo and JEG-3 using qRT-PCR

A high level of systemic oxidative stress in the placenta is known to be associated with preterm birth. To explore the association between oxidative stress and preterm birth, we further compared the mRNA expression of candidate genes in H_2O_2-treated human trophoblast cell line HTR-8/SVneo and choriocarcinoma JEG-3. First, to compare the cytotoxic effects of H_2O_2 in each cell line, we performed a cell viability assay following treatment with different concentrations of H_2O_2 for 24 h and 48 h. As shown in Figure 6, H_2O_2 significantly inhibited the viability of HTR-8/SVneo cells at concentrations of 25, 50, 100, and 250 μM after 24 h ($P<0.05$), and the inhibitory effect of H_2O_2 on both cells was dose-dependent. HTR-8 cells were more sensitive than JEG-3 cell lines to H_2O_2 cytotoxicity. The IC50 values of HTR-8/SVneo and JEG-3 cells were 34.50 and 78.66 uM after 24 h and 32.73 and 71.54 uM after 48 h of treatment, respectively.

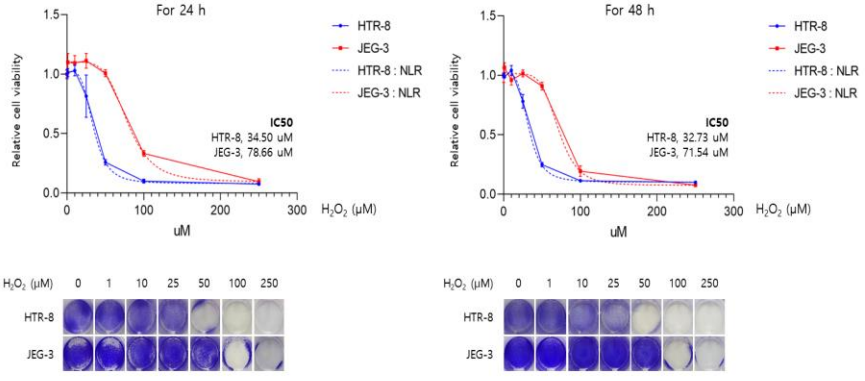

Figure 6. Time- and dose-dependent cytotoxic activity in the HTR-8/SVneo and JEG-3 cell lines

연세대학교
YONSEI UNIVERSITY

Next, we performed qRT-PCR to validate the candidate genes obtained through sequence analysis and to identify the preferentially expressed genes in two cell lines with oxidative stress. Based on the results that aberrantly expressed RNA transcripts do exist in the trophoblast cell line with oxidative stress, we further hypothesized that these DEGs could be detected in maternal AFS before the occurrence of preterm birth. To further characterize whether the 5 selected DEGs were related to oxidative mechanisms associated with preterm birth, we quantified their mRNA expression levels in HTR-8/SVneo and JEG-3 cell lines with H_2O_2 treatment. Expression analyses of RDH14, ZNF572, VOPP1, SERPINA12, and TCF15 were performed using RT-qPCR analysis in the RNA samples extracted from two cell lines with H_2O_2 treatment at different concentrations for 24 h and 48 h. Five candidate genes displayed a significant increase in mRNA expression in the HTR-8/SVneo cells after H_2O_2 treatment, but not in the JEG-cells resistant to H_2O_2 treatment (Fig. 7).

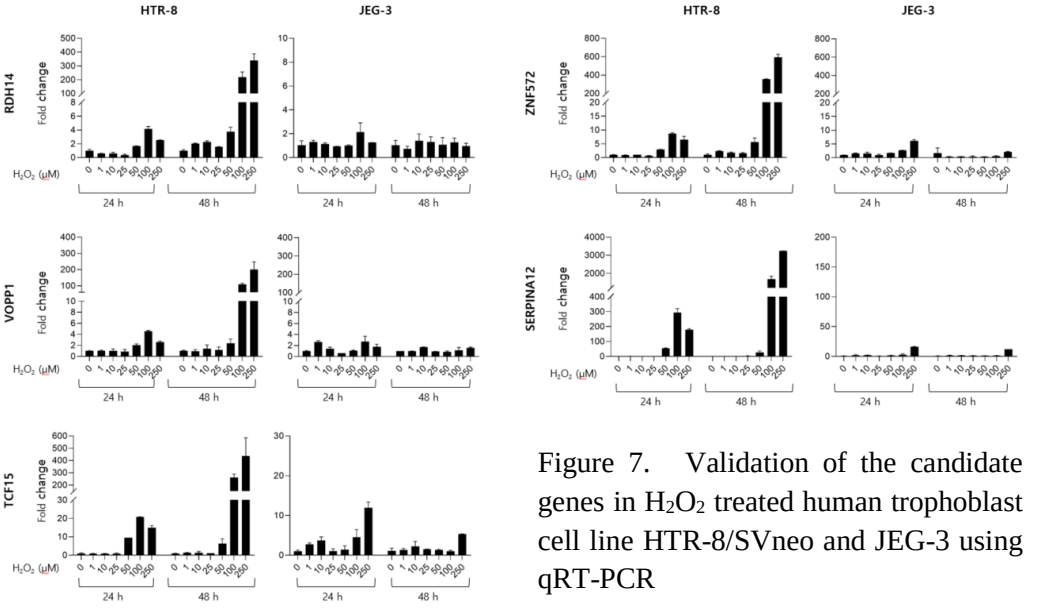

Figure 7. Validation of the candidate genes in H_2O_2 treated human trophoblast cell line HTR-8/SVneo and JEG-3 using qRT-PCR

연세대학교
YONSEI UNIVERSITY

6. Effect of liposaccharide treatment and glucose deprivation on the expression of candidate genes

It has been shown that LPS-induced oxidative stress is associated with preterm labor and that glucose deprivation results in oxidative stress. Thus, to examine the effect of LPS treatment and glucose deprivation on the expression of genes, we analyzed the expression of five candidate genes in glucose-deprived, LPS-induced HTR-8/SVneo cells using qRT-PCR analysis. The mRNA expression of all five candidate genes, RDH14, VOPP1, TCF15, ZNF572, and SERPINA12 were significantly increased in HTR-8/SVneo cells pretreated with LPS at concentration of 10 ng/ml and 100 ng/ml. The expression of five candidate genes was increased more than 2-fold by pretreatment with LPS in HTR-8/SVneo cells. It was observed that the expression levels of TCF15 and SERPINA12 were upregulated in HTR-8/SVneo cells under glucose deprivation conditions in comparison with the control cells growing with glucose. Thus, LPS treatment and glucose deprivation led to the upregulation of candidate genes (Fig. 8).

연세대학교
YONSEI UNIVERSITY

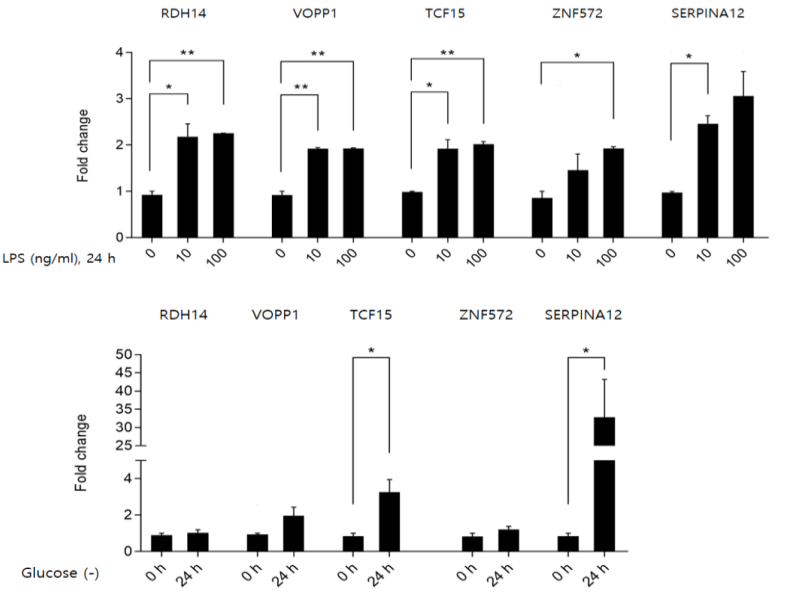

Figure 8. Effect of liposaccharide and glucose depriviation on the expression of the candidate genes in HTR-8/SVneo cells

Asterisks indicate levels of significance of differential expression (T-test: * $P \leq 0.05$, ** $P \leq 0.01$).

IV. DISCUSSION

To our knowledge, this is the first study comparing gene expression between preterm and term births using next-generation sequencing (NGS). This study investigated differences in gene expression levels using cf NAs isolated from AFS in the second-trimester of gestation for the prediction of preterm birth. This information regarding the variations occurring at the genome or transcriptome level will be help in improved antenatal recognition, thereby reducing the morbidity and mortality associated with preterm birth.

연세대학교
YONSEI UNIVERSITY

It is well known that AF is a complex and dynamic solution that can provide valuable information about fetal development and health status, and many studies have therefore reported the presence of cf NAs in AF. Although numerous studies have reported the use of AF cf NAs for discovering novel biomarkers for complications of pregnancy, there have been no studies using AF cf NAs to explore the pathogenesis of preterm birth. In this study, we analyzed the differences in global gene expression in the AF cf NAs before the onset of preterm birth to identify the pathogenesis of preterm birth. While other researchers have reported gene expression via microarray in AFS, few studies have performed whole genome and transcriptome analyses using sequencing analysis from AFS.

In this study, 54 DEGs were identified in the preterm birth cases compared with the term birth using nucleic acid sequencing. In particular, the expression patterns of 5 DEGs were verified by qRT-PCR in the second trimester AFS of preterm birth group and were highly associated with oxidative stress in the pathophysiology of preterm birth. Oxidative stress, characterized by imbalances in the redox system in the maternal-fetal intrauterine compartments, has been reported to play a critical role in the pathogenesis of spontaneous preterm birth. The present study showed that even before the onset of symptoms of preterm birth, there is a possibility of a redox imbalance in the maternal-fetal intrauterine compartments. Vora et al. examined the differences in gene expression in the cf RNA from AFS in women who delivered preterm and those who delivered at term using a customized nanostring panel containing genes related to oxidative stress.[34] They reported that changes in fetal gene expression related to oxidative stress and inflammation occur before the onset of preterm birth, which may facilitate early

연세대학교
YONSEI UNIVERSITY

detection of pregnancies at higher risk of preterm birth.

The role of inflammation related to the pathophysiology of preterm birth has been well studied, and several investigators have reported various inflammatory biomarkers associated with preterm birth. Oxidative stress, characterized by the generation of reactive oxygen species (ROS), is an integral factor in the inflammatory process.[35] Recent studies have demonstrated that oxidative stress plays an important role in infection-induced apoptosis and in the pathogenesis of preterm labor. LPS triggers inflammatory responses in human gestational tissues by increasing the production of pro-inflammatory cytokines and phospholipid metabolites, which is mediated by activation of the NF-κB pathway. In addition, glucose deprivation results in oxidative stress, and the alteration of the redox status of cells triggers stress-activated or other signal transduction pathways leading to cell death. In the present study, the expression of candidate genes was increased under conditions of oxidative stress and ROS production, which was induced by H_2O_2, LPS treatment, and glucose deprivation.

The production of ROS is necessary for the regulation of aerobic energy production. Redox balance is maintained through an intricate equilibrium between ROS production and subsequent elimination, and an interactive network of enzymatic and non-enzymatic antioxidant systems. Cells can protect themselves against oxidative stress by increasing the levels of endogenous and exogenous antioxidants or decreasing the production of ROS.[36-38] The overproduction of ROS in inflammation and cell damage can generate pathways leading to preterm birth. The balance between ROS and antioxidants can maintain good maternal health and safe childbirth.[39,40] Redox

연세대학교
YONSEI UNIVERSITY

imbalance is closely related to the fundamental pathogenesis of many pregnancy complications. The risk factors for preterm birth can lead to redox imbalance caused by high levels of ROS, such as hydroxyl radicals, hydrogen peroxide, superoxide anion, and nitric oxide. As a result, ROS-mediated oxidative stress contributes to collagen degradation and consume antioxidant defenses. Collagen damage in human gestational tissue can result in uterine contractions and tears in the chorioamniotic sac, leading to preterm birth.

Our aim was to identify genetic factors that contribute to preterm birth prior to the onset of symptoms of preterm birth. Based on these results, we identified a genetic factor associated with redox imbalance as a contributing factor to preterm birth. This led to further investigations to define high-risk populations and characterize gene expression in the AFS. However, there are several limitations in this study. First, we performed a comparative study with sequencing data obtained from five samples each of preterm and term births. This relatively small sample size makes it difficult to draw definitive conclusions. While it is widely acknowledged that increasing the number of replicates in sequencing experiments usually results in more reliable results, the exact relationship between the number of replicates and the possibility of identifying DEGs has not been thoroughly investigated. However, Conesa et al. recommended that sequencing experiments have at least three biological replicates if sample availability is not limited to allow all sequencing analysis to utilize reproducibility between replicates.[41] Second, this study may have annotation biases because of inequality in gene annotations of bioinformatics analysis. Third, it lacks sufficient power to demonstrate the mechanisms leading to early delivery due to the lack of early preterm

연세대학교
YONSEI UNIVERSITY

births in our study. The preterm birth group in our study included only late preterm births (34–36 weeks), and there was only a difference of about 5 weeks compared to the mean gestational age at delivery of the term birth group. The absence of an early primordial group in sample collection resulted in a limitation, but it is still another significance of this study that gene expression differences occurred in second-trimester AF prior to 20–25 weeks before delivery.

Despite these limitations, this study has many strengths. Our study is the first to investigate gene signatures using AF cf NAs in preterm births. In addition, we analyzed AF samples that were obtained during the early second trimester before the subjects developed signs of preterm birth. The present study showed comparable differences in cell-free gene signatures in early second trimester AF between preterm birth and term birth groups.

V. CONCLUSION

In conclusion, specific expression patterns of genes associated with high oxidative stress in pregnant women may indicate placental tissue damage predisposing to preterm birth. In addition, they might suggest a high oxidative stress in preterm newborns due to direct blood exchange in the placenta. Expression patterns of candidate genes shown in our study may be a criterion for assigning therapeutic targets so as to regulate their function at the earliest, thereby modulating the expression of proteins and their pathogenic pathways in preterm birth. These differentially expressed genes may elucidate the underlying mechanisms of preterm birth and be used to predict preterm birth in early pregnancy.

연세대학교
YONSEI UNIVERSITY

REFERENCES

1. Romero R, Dey SK, Fisher SJ. Preterm labor: one syndrome, many causes. Science 2014;345:760-5.

2. Boyd HA, Poulsen G, Wohlfahrt J, Murray JC, Feenstra B, Melbye M. Maternal contributions to preterm delivery. Am J Epidemiol 2009;170:1358-64.

3. Plunkett J, Feitosa MF, Trusgnich M, et al. Mother's genome or maternally-inherited genes acting in the fetus influence gestational age in familial preterm birth. Hum Hered 2009;68:209-19.

4. Smith R. Parturition. N Engl J Med 2007;356:271-83.

5. York TP, Eaves LJ, Neale MC, Strauss JF, 3rd. The contribution of genetic and environmental factors to the duration of pregnancy. Am J Obstet Gynecol 2014;210:398-405.

6. Underwood MA, Gilbert WM, Sherman MP. Amniotic fluid: not just fetal urine anymore. J Perinatol 2005;25:341-8.

7. Wilson RD. Amniocentesis and chorionic villus sampling. Curr Opin Obstet Gynecol 2000;12:81-6.

8. Larrabee PB, Johnson KL, Peter I, Bianchi DW. Presence of filterable and nonfilterable cell-free mRNA in amniotic fluid. Clin Chem 2005;51:1024-6.

9. Gahan PB, Stroun M. The virtosome-a novel cytosolic informative entity and intercellular messenger. Cell Biochem Funct 2010;28:529-38.

10. Bronkhorst AJ, Wentzel JF, Aucamp J, van Dyk E, du Plessis L, Pretorius PJ.

연세대학교
YONSEI UNIVERSITY

Characterization of the cell-free DNA released by cultured cancer cells. Biochim Biophys Acta 2016;1863:157-65.

11. Fernando MR, Jiang C, Krzyzanowski GD, Ryan WL. New evidence that a large proportion of human blood plasma cell-free DNA is localized in exosomes. PLoS One 2017;12:e0183915.

12. Thakur BK, Zhang H, Becker A, et al. Double-stranded DNA in exosomes: a novel biomarker in cancer detection. Cell Res 2014;24:766-9.

13. Wang W, Kong P, Ma G, et al. Characterization of the release and biological significance of cell-free DNA from breast cancer cell lines. Oncotarget 2017;8:43180-91.

14. Zhu J, Zhang F, Du M, Zhang P, Fu S, Wang L. Molecular characterization of cell-free eccDNAs in human plasma. Sci Rep 2017;7:10968.

15. Lo YM, Corbetta N, Chamberlain PF, et al. Presence of fetal DNA in maternal plasma and serum. Lancet 1997;350:485-7.

16. Bianchi DW, LeShane ES, Cowan JM. Large amounts of cell-free fetal DNA are present in amniotic fluid. Clin Chem 2001;47:1867-9.

17. Lun FM, Chiu RW, Leung TY, Leung TN, Lau TK, Lo YM. Epigenetic analysis of RASSF1A gene in cell-free DNA in amniotic fluid. Clin Chem 2007;53:796-8.

18. Maron JL, Bianchi DW. Prenatal diagnosis using cell-free nucleic acids in maternal body fluids: a decade of progress. Am J Med Genet C Semin Med Genet

연세대학교
YONSEI UNIVERSITY

2007;145c:5-17.

19. Sekizawa A, Yokokawa K, Sugito Y, et al. Evaluation of bidirectional transfer of plasma DNA through placenta. Hum Genet 2003;113:307-10.

20. Jung YW, Shim SS, Park JE, et al. Analysis of the cell-free amniotic fluid transcriptome expressed during the euploid mid-trimester of pregnancy. Eur J Obstet Gynecol Reprod Biol 2016;203:94-8.

21. Mersy E, Smits LJ, van Winden LA, et al. Noninvasive detection of fetal trisomy 21: systematic review and report of quality and outcomes of diagnostic accuracy studies performed between 1997 and 2012. Hum Reprod Update 2013;19:318-29.

22. Koide K, Slonim DK, Johnson KL, Tantravahi U, Cowan JM, Bianchi DW. Transcriptomic analysis of cell-free fetal RNA suggests a specific molecular phenotype in trisomy 18. Hum Genet 2011;129:295-305.

23. Hui L, Slonim DK, Wick HC, Johnson KL, Koide K, Bianchi DW. Novel neurodevelopmental information revealed in amniotic fluid supernatant transcripts from fetuses with trisomies 18 and 21. Hum Genet 2012;131:1751-9.

24. Massingham LJ, Johnson KL, Scholl TM, Slonim DK, Wick HC, Bianchi DW. Amniotic fluid RNA gene expression profiling provides insights into the phenotype of Turner syndrome. Hum Genet 2014;133:1075-82.

25. Hui L, Wick HC, Moise KJ, Jr., et al. Global gene expression analysis of amniotic fluid cell-free RNA from recipient twins with twin-twin transfusion syndrome.

연세대학교
YONSEI UNIVERSITY

Prenat Diagn 2013;33:873-83.

26. Reuter JA, Spacek DV, Pai RK, Snyder MP. Simul-seq: combined DNA and RNA sequencing for whole-genome and transcriptome profiling. Nat Methods 2016;13:953-8.

27. Bradford JR, Hey Y, Yates T, Li Y, Pepper SD, Miller CJ. A comparison of massively parallel nucleotide sequencing with oligonucleotide microarrays for global transcription profiling. BMC Genomics 2010;11:282.

28. Perkins JR, Antunes-Martins A, Calvo M, et al. A comparison of RNA-seq and exon arrays for whole genome transcription profiling of the L5 spinal nerve transection model of neuropathic pain in the rat. Mol Pain 2014;10:7.

29. Zhao S, Fung-Leung WP, Bittner A, Ngo K, Liu X. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. PLoS One 2014;9:e78644.

30. Ackerman WEt, Buhimschi IA, Eidem HR, et al. Comprehensive RNA profiling of villous trophoblast and decidua basalis in pregnancies complicated by preterm birth following intra-amniotic infection. Placenta 2016;44:23-33.

31. Chan YW, van den Berg HA, Moore JD, Quenby S, Blanks AM. Assessment of myometrial transcriptome changes associated with spontaneous human labour by high-throughput RNA-seq. Exp Physiol 2014;99:510-24.

32. Kaartokallio T, Cervera A, Kyllonen A, Laivuori K, Kere J, Laivuori H. Gene expression profiling of pre-eclamptic placentae by RNA sequencing. Sci Rep

연세대학교
YONSEI UNIVERSITY

2015;5:14107.

33. Ostling H, Kruse R, Helenius G, Lodefalk M. Placental expression of microRNAs in infants born small for gestational age. Placenta 2019;81:46-53.

34. Vora NL, Smeester L, Boggess K, Fry RC. Investigating the Role of Fetal Gene Expression in Preterm Birth. Reprod Sci 2017;24:824-8.

35. Polettini J, Dutta EH, Behnia F, Saade GR, Torloni MR, Menon R. Aging of intrauterine tissues in spontaneous preterm birth and preterm premature rupture of the membranes: A systematic review of the literature. Placenta 2015;36:969-73.

36. Menon R. Oxidative stress damage as a detrimental factor in preterm birth pathology. Front Immunol 2014;5:567.

37. Menon R, Boldogh I, Hawkins HK, et al. Histological evidence of oxidative stress and premature senescence in preterm premature rupture of the human fetal membranes recapitulated in vitro. Am J Pathol 2014;184:1740-51.

38. Romero R, Kusanovic JP, Chaiworapongsa T, Hassan SS. Placental bed disorders in preterm labor, preterm PROM, spontaneous abortion and abruptio placentae. Best Pract Res Clin Obstet Gynaecol 2011;25:313-27.

39. Agarwal A, Aponte-Mellado A, Premkumar BJ, Shaman A, Gupta S. The effects of oxidative stress on female reproduction: a review. Reprod Biol Endocrinol 2012;10:49.

40. Sultana Z, Maiti K, Aitken J, Morris J, Dedman L, Smith R. Oxidative stress, placental ageing-related pathologies and adverse pregnancy outcomes. Am J Reprod

연세대학교
YONSEI UNIVERSITY

Immunol 2017;77.

41. Conesa A, Madrigal P, Tarazona S, et al. A survey of best practices for RNA-seq data analysis. Genome Biol 2016;17:13.

연세대학교
YONSEI UNIVERSITY

ABSTRACT(IN KOREAN)

임신중기 무증상 산모의 양수를 이용한 새로운 조기분만예측 유전자 표지자 분석

<지도교수 김영한>

연세대학교 대학원 의학과

김민아

조기분만은 전체 임신의 5-7%를 차지하고 신생아 사망률과 이환율을 높이는 주된 원인이다. 양수상층액은 임신합병증과 관련된 유전정보를 제공할 수 있는 풍부한 세포유리핵산을 함유하고 있다. 본 연구에서는 조기진통이 발생하기 이전인 임신 16-19주 산모의 양수상층액의 세포유리핵산에서 sequencing 기술을 이용하여 조기분만과 만삭분만을 예측할 수 있는 후보유전자를 발굴하고자 하였다. 조기분만한 산모 5명과 만삭분만한 산모 5명의 16-19주 시기에 염색체검사를 하기 위하여 채취한 양수의 상층액을 이용하여 Illumina Nextseq 500 platform으로 sequencing 분석을 시행하였다. 후보유전자의 검증실험을 위하여 조기분만한 산모 21명과 만삭분만한 산모 40명의 양수상층액과 hydrogen peroxide, lipopolysaccharide 및 glucose-deprivation 처리한 융모막세포에서 qRT-PCR을 이용하여 후보유전자의 발현양상을 비교분석하였다. Sequencing 결과, 만삭분만한 산모에 비하여 조기분만한 산모의

연세대학교
YONSEI UNIVERSITY

양수상층액에서 43개의 유전자가 유의하게 높게 발현되었고 11개의 유전자가 유의하게 낮게 발현되었다. 후보유전자의 검증실험을 위하여 조기분만 21명, 만삭분만 40명의 양수상층액에서 qRT-PCR을 시행한 결과, 조기분만군에서 유전자 RDH14, ZNF572, VOPP1, SERPINA12, TCF15가 높게 발현됨을 볼 수 있었다. 또한 조기분만과 관련된 산화스트레스와의 관련성을 살펴보기 위하여 과산화수소 처리를 한 두 가지의 융모막세포주, 즉 HTR-8/SVneo 와 JEG-3 세포에서 5개의 후보유전자에 대한 qRT-PCR실험을 시행한 결과, 과산화수소 독성에 민감한 HTR-8/SVneo 세포에서 5개의 유전자가 높게 발현되었으며 lipospolysaccharide 처리한 HTR-8/SVneo 세포에서 5개의 유전자가 높게 발현되는 것을 관찰하였다. 본 연구의 결과를 통하여 조기분만이 발생하기 이전에도 무증상 임신중기 양수에서 조기분만과 관련된 유전자 발현의 차이를 보이며 이들 후보유전자는 산화스트레스와 관련되어 조기분만의 병태생리에 중요한 역할을 함으로써 잠재적 발병예측인자로서의 가능성을 제시할 수 있을 것으로 생각된다.

핵심되는 말: 조기분만, 세포유리핵산, 유전자분석, 산화스트레스