**Supporting Information**

**Early Pregnancy Prediction of Spontaneous Preterm Birth Prior to**

**32 Completed Weeks of Pregnancy Using Plasma RNA:**

*Transcriptome Discovery and Initial Validation Study of an RNA Panel of Markers*

*Running Title:* Early Pregnancy Prediction of Spontaneous Preterm Birth

1,2 Carl P. Weiner, 2 Yafeng Dong, 1 Helen Zhou, 3 Howard Cuckle, 4 Risa Ramsey, 5 Robert Egerman, 6 Irina A. Buhimschi, 6 Catalin S. Buhimschi

1 Department of Obstetrics and Gynecology, University of Kansas School of Medicine, Kansas City, KS, USA

2 Rosetta Signaling Laboratory, Mission Hills, KS, USA

3 Faculty of Medicine, Tel Aviv University, Ramat Aviv, Israel

4 University of Tennessee Health Science Center in Memphis, Memphis, TH, USA

5 Department of Obstetrics and Gynecology, University of Florida, Gainesville FL, USA

6 Department of Obstetrics and Gynecology, University of Illinois Chicago, Chicago, IL, USA

Corresponding author:

Carl P Weiner MD MBA,

Department Obstetrics and Gynecology, KUMC,

3901 Rainbow Blvd, Kansas City, KS, 66160. USA

cweiner@kumc.edu and cpweiner@gmail.com

**METHODS**

**Plasma RNA Isolate QC**

**Agarose gel runs supplied the initial RNA QC assessment for the PCF RNA isolates.**

***QC for GeneChip Human Exon 1.0 ST array***

**QC of the biotin labeled single stranded cDNA (sscDNA) generated from the Ambion Sense Target Labeling Protocol using the Ambion Whole Transcript (WT) Expression Kit (Ambion 4411973) was performed using the Agilent Bioanalyzer 2100 running the Agilent RNA6000 Nano Kit series II assay. (Agilent 5067-1511). After the completion of the sscDNA fragmentation using the Affymetrix GeneChip WT Terminal Labeling kit (Affymetrix 900671), a 1ul aliquot of the of the fragmented sscDNA was retained prior to the biotin labeling procedure. After completing the terminal biotin labeling of the sscDNA, a 1ul aliquot was retained for the Agilent RNA6000 Nano Kit Series II assay. The 1ul aliquot of the fragmented, unlabeled sscDNA was loaded on the RNA6000 Nano chip in an adjacent well to the 1ul aliquot of the biotin labeled sscDNA, according to manufacturer’s instruction. The RNA6000 Nano chip was run on the Agilent 2100 Bioanalyzer using the Eukaryotic Total RNA Nano assay. A distinct shift due to the biotin labeling was seen compared to the unlabeled sscDNA and the biotin labeled sscDNA’s migration profiles indicating a successful labeling reaction. The remaining biotin labeled sscDNA was hybridized to the GeneChip Human Exon 1.0 ST array.**

***Microarray Data Analysis***

**Gene expression profiling was carried out using the Affymetrix GeneChip Human Exon 1.0 ST Array. This array consists of over 1.4 million probe sets, of which around 289,961 are core exon probe sets supported by putative full-length mRNA (RefSeq and full-length GenBank annotated alignments). These core probe sets map to approximately 28,869 genes of high confidence. These genes were the focus of the current study. The exon-arrays were background corrected, normalized and gene-level summarized using the Robust Multichip Average (RMA) procedure[[1]](#endnote-1). The QC plots indicate high quality data from all the arrays with significantly different expression profiles between conditions. The resulting log (base 2) transformed signal intensities (expression values) were used for ascertaining differentially expressed genes. The data was modelled using a linear regression model with linear contrast for obtaining fold change and p-value statistics for individual genes between the groups being compared. The raw p-values were adjusted for multiple hypothesis testing by the Benjamini and Hochberg procedure[[2]](#endnote-2), giving a corresponding false discovery rate (FDR). Genes with multiple transcripts in the microarray were resolved by taking the transcript with maximum variance across samples. All computations were performed in Matlab (R2018b, The MathWorks Inc, Natick, MA) and the Partek Genomic suite (v 6.5, Partek Inc., St. Louis, MO).**

***QC for the GeneChip miRNA arrays (version 1.0, 2.0, 3.0)***

**A colorimetric ELOSA assay was performed to QC the labeled miRNA generated using the FlashTag Biotin HSR RNA labeling kit (Life Technologies 901910). The ELOSA assay was prepared by loading 75ul of a 1:50 dilution of ELOSA Spotting Oligos (contained in the FlashTag kit) in 1xPBS into each well of a Flat Bottom Immobilizer Amino – 8 well strip plate (Nunc 436013) and incubating 12 hours at 4◦C. The Spotting Oligos were expelled and the wells washed 2X’s with 1xPBS supplemented with 0.02% Tween-20. The wells were blocked with 150ul of 1xPBS supplemented with 5% BSA and incubated 1 hour. During the blocking procedure the labeled RNA sample, positive control and negative control were prepared. 2ul of the 3DNA dendrimer signal amplified and biotin labeled RNA was combined with 48ul of 5xSSC supplemented with 0.05% SDS and 0.005% BSA and 2.5% of 25% Dextran sulfate solution. An ELOSA Positive Control RNA (FlashTag kit) was prepared using 2ul Control RNA combined with 48ul of 5xSSC supplemented with 0.05% SDS and 0.005% BSA and 2.5% of 25% Dextran sulfate solution. A Negative control was prepared containing 50ul of 5xSSC supplemented with 0.05% SDS and 0.005% BSA and 2.5% of 25% Dextran sulfate solution. After purging the 5% Blocking reagent from the wells, 52.5ul of each hybridization mixture was added to the designated well which were sealed with an adhesive plate seal and incubated at room temperature for 1 hour. HRP-Conjugated Streptavidin (Thermo Scientific ENN100) is diluted 1:6000 in 1xPBS and supplemented with 5% BSA. The hybridization solution was expelled and the wells vigorously washed 4Xs with 1xPBS, 0.02% Tween-20. Wells are drained and 75ul of the diluted HRP-Conjugated Streptavidin added. Wells were sealed with an adhesive plate seal and incubated at RT 30 minutes. The HRP-Conjugated Streptavidin was expelled and the wells vigorously washed 4Xs with 1xPBS, 0.02% Tween-20. Wells were drained and bubbles removed with compressed air. 100ul of TMB Substrate Solution (Thermo Fisher N301) was added to each well and the wells sealed with an adhesive plate seal. The plate was incubated at room temperature for 30 minutes. The sample, positive control and negative control were assessed for a blue substrate color which indicates a positive biotin labeling with the 3DNA dendrimer. Positive control and RNA sample yielded a blue color and the negative control remained clear.**

***Micro RNA Data Analysis***

**Micro RNA expression profiling was carried out using the Affymetrix GeneChip miRNA 1.0 Arrays. This array interrogates all miRNA sequences listed in miRBase Release 11. In total, it interrogates 6,703 mature miRNAs from 71 organisms of which 847 are from human. The arrays were background corrected, normalized and gene-level summarized using the Robust Multichip Average (RMA) algorithm1. The resulting log (base2) transformed intensity values were used for quality control and further analysis. The QC plot indicate high quality data from all the arrays with good separation between conditions. Micro RNA fold changes were obtained by taking the exponentiation of the linear contrast between the least square means of the PTB and control samples. The significance of these fold changes was calculated using a one-way analysis of variance model. The raw significance p-values were adjusted for multiple testing using the Benjamini-Hochberg FDR correction2.**

**Miscellaneous Laboratory Methods**

*Myometrial immunohistochemistry for APOA4 protein:*

 Tissues fixed in 10% neutral buffered formalin and embedded in paraffin blocks were obtained from the KUMC pregnancy biobank. Four 10-12 micrometer sections were mounted on slides, deparaffinized, and rehydrated through graded alcohols to water. Sections were then incubated with 0.025% Triton X-100 in Tris buffered saline (TBS) for 10 minutes. Endogenous peroxidase was quenched with a 10-minute incubation of 0.3% hydrogen peroxide in TBS. Samples were washed in TBS and blocked with 10% normal rabbit serum with 1% BSA in TBS for 2 hours at room temperature. The primary anti-human APOA4 antibody (Invitrogen) was diluted 1:500 in TBS with 1% BSA and incubated with tissue overnight in a humidified chamber overnight at 4°C. The following day, the sections were triple washed and a biotin-conjugated anti-rabbit secondary antibody (Vector Labs, Burlingame, California) was applied followed by 30-minute incubation followed by triple wash prior to applying Vectastain Elite ABC reagent (Vector Labs). The slides were triple washed and colorized with ImPACT diaminobenzidine tetrahydrochloride substrate kit (Vector Labs) for 3 minutes and rinsed with water. They were then counterstained for 1 minute with Vector Hematoxylin QS (Vector Labs), rinsed with tap water, dehydrated, cleared, and mounted with mounting media. Sections were visualized under a light microscope.

*APOA4 expression plasmid construction*

The APOA4 gene (NM\_000482; 1471 bp) was amplified by PCR. The purified PCR product ligated in frame into pcDNA™3.1D/V5-His-TOPO expression vector according to the manufacturer’s protocol (Life Technologies) and transfected into competent cells. The presence of the insert and its orientation was confirmed by double restriction enzyme digestion. The fidelity of cloned sequence was confirmed by DNA sequencing.

*Cell culture and transfection*

The immortalized human pregnant myometrial cells (PHM) were cultured in DMEM with 10% fetal bovine serum and 1% penicillin/streptomycin (10,000 U) under standard cell culture conditions (37 °C, 5% CO2). For transfection, cells were seeded at 0.5 × 106 cells per well in 6 well plates and cultured for 24 hours. The APOA4 plasmid (2.5 μg per well), the synthetic siRNA-IFNGR (100 nM per well) or empty plasmid controls were transfected into PHM cells using Lipofectamine 2000™ (Life Technologies) 10 μl/well. The cells were harvested for RNA and protein isolation after 24 h incubation.

*Western blot analysis:*

Protein was extracted from cells with ice-cold lysis buffer (150 mM NaCl, 1% IGEPAL Ca- 630, 0.5% sodium deoxycholate, 0.1% SDS, 1 mM EDTA, 50 mM Tris,pH 7.8) containing protease and phosphatase inhibitor cocktails (Sigma). Protein concentration was quantified using the Bradford™ reagent (Bio-Rad, Hercules CA). Cleared lysates containing equivalent amounts of total protein (40 μg) were loaded onto a 15% acrylamide gel and electrophoresed in Tris/Glycine/SDS running buffer (Bio- Rad). Proteins were then transferred to polyvinylidene difluoride membranes with Tris/Glycine transfer buffer (Bio- Rad). Membranes were blocked in 2% non-fat milk and incubated overnight at 4 °C with the anti-APOA4 (1:500, Invitrogen), anti- β-actin (1:500, Abcam). Primary antibodies were visualized using horseradish peroxidase-labeled species-appropriate IgG (1:5000 or 1:10,000; GE Healthcare Life Sciences) and enhanced chemiluminescence (GE Healthcare Life Sciences). Western blot results were captured by ChemiDOC Imaging System™ (Bio-Rad) and quantified by Image Lab Software (Bio-Rad).

*Fura- Ca2+ measurements:*

Intracellular Ca2+ was determined using FLIPR calcium 6 assay kit (Molecular Devices, Sunnyvale, California), according to the manufacturer’s instructions. Briefly, cells were loaded with calcium 6 and read using FlexStation 3 (Molecular Devices, San Jose CA). After recording the baseline fluorescence for 20 seconds, oxytocin [10 nM] was added to the cells through a multichannel pipettor included as a part of the fluidics module of FlexStation 3 and fluorescence changes were monitored over time. Thapsigargin, a sarco-endoplasmic reticulum Ca2+ ATP-ase inhibitor, was used to verify Ca2+ mobilization from internal stores. Thapsigargin (5 mM) was added for >10 minutes prior to oxytocin to block SERCA and deplete the SR of Ca2+. The relative fluorescence changes for each well were expressed as DF ¼ (Fmax - Fmin), where Fmax and Fmin was the peak relative fluorescence unit achieved and the lowest fluorescence values at baseline, respectively.

*Myometrial cell Contraction Frequency:*

Collagen Gel Contraction Assay was performed using Cell Contraction Assay Kit (Cell Biolabs, Cat. # CBA-201). Myometrial cells were harvested and resuspend in serum-free medium. Collagen lattice was prepared by mixing 2 parts of cell suspension and 8 parts of cold collagen gel solution. One half milliliter of the cell-collagen mixture was added per well in a 24-well plate and incubated 1h at 37ºC. After collagen polymerization, 1.0 mL of serum-free culture medium was added atop each collagen gel lattice. Cultures were incubated for 2d, during which stress developed. Before release of the stressed matrix, cells were treated with agents to stimulate or inhibit contraction, such as oxytocin. To initiate contraction, the collagen gels were gently released from the sides of the culture dishes. The collagen gel size change was measured at various times with a quantified image analysis software (Image Pro Plus, Medi Cybernetics, Rockville MD).

**Statistical Analysis**

Laboratory Statistics

Statistical analysis used GraphPad Prism version 6.0. Statistical comparisons between two groups used a Student's t-test after determining samples were normally distributed. The difference among multiple groups was analyzed by one-way ANOVA followed by Student-Newman-Keuls multiple comparisons post hoc tests. P values below 0.05 were considered statistically significant. Data are presented as mean ± SD

**Discussion**

Biologic Plausibility

APOA4 and its potential downstream targets are of interest. *In silica* studies suggested APOA4 could interact either directly or indirectly with five (5) Preterm Initiator genes. IFNG (interferon gamma) appeared as a branch point to: CD3E, a component in the T-cell receptor-CD3 complex; HLA-DOA, which regulates HLA-DM-mediated peptide loading on MHC class II molecules; NDRG4, a member of the N-myc down regulated gene, which may be also be associated with mobilization of extracellular calcium[[3]](#endnote-3); and VPS33A, (vacuolar protein sorting 33 homolog A) is associated with calbindin[[4]](#endnote-4). Calbindin is a calcium-binding protein in the troponin C superfamily and expressed mainly in intestine, lung, and uterus in a tissue-specific manner.[[5]](#endnote-5)

Low dietary choline or a deficiency of methylenetetrahydrofolate reductase (MTHFR) can cause hyperhomocysteinemia which has been associated with adverse reproductive outcomes [[6]](#endnote-6). Disturbances in 1-carbon metabolism could also reduce maternal PPARα expression, perhaps through promoter hypermethylation, and increase IFNγ and decrease IL-10. The disturbance of the T helper (Th1) (IFNγ):Th2 (IL-10) ratio and an increase in inflammatory mediators may also contribute downstream to pregnancy complications.[[7]](#endnote-7)

The precise function of the APOA4 gene is unknown. It is increased in the follicular fluid of women with polycystic ovary disease[[8]](#endnote-8). At least one apolipoprotein (A-1) reportedly binds with IFN-α. The resulting fusion protein crosses the blood-brain barrier by a saturable transport mechanism[[9]](#endnote-9). One relatively unstudied role of APOA is its role as an HDL chaperone for plasma RNAs[[10]](#endnote-10). So in addition to changes in PCF RNA associated with extracellular vesicles, changes in plasma RNAs might be associated with an high density lipoproteins via an APOA-dependent mechanism.

Vps33a is a Sec1-related protein that plays a role in vesicular transport to the lysosomal compartment.

Calbindin-D9k is encoded by the S100G gene in humans (also known as CALB3). It has two EF-hands sequences which bind Ca2+ tightly. Calbindin-D9k mediates the transport of calcium across enterocytes from the apical side, where entry is regulated by the calcium channel TRPV6, to the basolateral side, where calcium pumps such as PMCA1 utilize intracellular ATP to pump calcium into the blood [[11]](#endnote-11). The transport of calcium across the enterocyte cytoplasm appears the rate-limiting step for calcium absorption in the intestine; the presence of calbindin increases the amount of calcium crossing the cell without raising the free concentration [[12]](#endnote-12). Calbindin-D9k may also stimulate the basolateral calcium-pumping ATPases. Expression of calbindin-D9k, like that of calbindin-D28k, is stimulated by the active vitamin D metabolite, calcitriol although the mechanisms remain controversial [[13]](#endnote-13).

Tables:

1. Combined list of differentially expressed RNAs in women 22-24w with sPTB≦32w according to Affymetrix Human Exon and GeneChip miRNA arrays. The samples were derived from 5 women with no comorbidities and sPTB≦32w (2 preceded by PPROM) and 5 of similar age and parity with no comorbidities delivered spontaneously at term. Half the women were Black and half White. All were selected for clarity of diagnosis, the availability of longitudinal samples and a complete medical record.

**Legends**

**Figure s1:** Four (4) plasma mRNAs were selected during Discovery as potential genes for normalizing expression among patients. Expression was measured by qRT PCR across gestational age in the same 10 subjects in the Discovery arm plus an additional 10 women who experienced PPROM. Only PPIA mRNA (D) fulfilled the desired stability criteria. Note that 2 RNAs commonly used for normalization, 18s and GAPDH were in plasma samples variably impacted by gestation and or PTB.

**Figure s2:** Four (4) plasma mRNAs were selected during Discovery as potential genes for normalizing expression among patients. Expression was measured by qRT PCR across gestational age in the same 10 subjects in the Discovery arm plus an additional 10 women who experienced PPROM. Only U6 snRNA (D) fulfilled the desired stability criteria.

**Figure s3:** Genomic assessment of the 86 mRNA and 13 miRs selected through Discovery: The top three Disease associations were Hypertension, Neurologic Manifestations and Neurobehavioral Manifestations. The top three Maps included DNA damage, cell cycle and Development of PIP3 signaling in cardiac myocytes. The top three Networks were the [O-hexadecanoyl-(L)-carnitine, 2-arachidonoyl-glycerol 3 phosphocholine, and Phosphatidylinositol-3,4-diphosphate pathways.

**Figure s4:**

A: qRT PCR Confirmation studies of the 13 miRs identified as being differentially expressed in women destined for sPTB ≦32 weeks measured in the same cDNA used for Discovery. Each of the 3 miR increased in Discovery were confirmed, as well as 7 of the 10 miRs that declined in association with sPTB ≦32 weeks.

B: Expression levels of miR-548L and miR-99a were measured longitudinally across gestational age. While miR-99a was increased by 16 weeks’ in women destined for sPTB ≦32 weeks, the rise in miR-548L did not reach statistical significance until 24 weeks’ gestation.

**Figure s5:** Initial Validation Study of the PCF RNA Panel for sPTB≦32 weeks – Individual patient expression levels of each confirmed PCF RNA. Red- sPTB≦32; Blue control.

**Figure s6:** Immunohistochemistry for the APOA4 protein was seen only in placenta from women with sPTB ≦32 weeks.

**Figure s7:** APOA4 mRNA was cloned into the pcDNA 3.1 vector and successful transfected into immortalized pregnant human myometrial cells (IPHM) to eliminate the vagaries of cell RNA uptake (A and B). Transfection more than doubled the intracellular APOA4 protein levels (C). APOA4 overexpression also lead to a 5-fold increase in intracellular calcium as measured with fura-Ca2+ (D). Transfection with empty phage had no effect (not shown).

References

1. Irizarry, R.A., et al. Exploration, normalization, and summaries of high-density oligonucleotide array probe level data. Biostatistics 4, 249-264 (2003). [↑](#endnote-ref-1)
2. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological), 289-300 (1995). [↑](#endnote-ref-2)
3. Brailoiu GC, Dun SL, Mizuo K, Brailoiu E, Yang J, Chang JK, et al. Smooth muscle-associated protein 8: distribution and biological activity in the rat brain. J Neurosci Res. 2007;85:1789-96. [↑](#endnote-ref-3)
4. Chintala S, Novak EK, Spernyak JA, Mazurchuk R, Torres G, Patel S, et al. The Vps33a gene regulates behavior and cerebellar Purkinje cell number. Brain Res. 2009 Apr 17;1266:18-28. doi: 10.1016/j.brainres.2009.02.035. Epub 2009 Feb 27. [↑](#endnote-ref-4)
5. Romagnolo B, Cluzeaud F, Lambert M, Colnot S, Porteu A, Molina T, et al. Tissue-specific and hormonal regulation of calbindin-D9K fusion genes in transgenic mice. J Biol Chem. 1996 Jul 12;271(28):16820-6. [↑](#endnote-ref-5)
6. Aubard Y, Darodes N, Cantaloube M. Hyperhomocysteinemia and pregnancy--review of our present understanding and therapeutic implications. Eur J Obstet Gynecol Reprod Biol. 2000; 93:157-65. [↑](#endnote-ref-6)
7. Mikael LG, Pancer J, Wu Q, Rozen R. Disturbed one-carbon metabolism causing adverse reproductive outcomes in mice is associated with altered expression of apolipoprotein AI and inflammatory mediators PPARα, interferon-γ, and interleukin-10. J Nutr. 2012 Mar;142(3):411-8. doi: 10.3945/jn.111.151753. Epub 2012 Jan 18. [↑](#endnote-ref-7)
8. Kim YS, Gu BH, Choi BC, Kim MS, Song S, Yun JH, et al. Apolipoprotein A-IV as a novel gene associated with polycystic ovary syndrome. Int J Mol Med. 2013 Mar;31(3):707-16. doi: 10.3892/ijmm.2013.1250. Epub 2013 Jan 17. [↑](#endnote-ref-8)
9. Fioravanti J, Medina-Echeverz J, Ardaiz N, Gomar C, Parra-Guillén ZP, Prieto J, et al. The fusion protein of IFN-α and apolipoprotein A-I crosses the blood-brain barrier by a saturable transport mechanism. J Immunol. 2012 Apr 15;188(8):3988-92. doi: 10.4049/jimmunol.1101598. Epub 2012 Mar 14. [↑](#endnote-ref-9)
10. Nik Mohamed Kamal NNSB, Shahidan WNS. Non-Exosomal and Exosomal Circulatory MicroRNAs: Which Are More Valid as Biomarkers? Front Pharmacol. 2020;10:1500. [↑](#endnote-ref-10)
11. Wasserman, RH; Chandler, JS, Meyer, SA, Smith, CA, Brindak, ME, Fullmer, CS, et al. "Intestinal calcium transport and calcium extrusion processes at the basolateral membrane". The Journal of Nutrition 1992; 122 (3 Suppl): 662–71. [↑](#endnote-ref-11)
12. Feher, JJ; Fullmer, CS, Wasserman, RH. "Role of facilitated diffusion of calcium by calbindin in intestinal calcium absorption". Am J Physio 1992; 262 (2 Pt 1): C517–26. [↑](#endnote-ref-12)
13. Barley, NF; Prathalingam, SR, Zhi, P, Legon, S, Howard, A, Walters, JR. "Factors involved in the duodenal expression of the human calbindin-D9k gene". The Biochemical Journal 1999; 341 (3): 491–500. doi:10.1042/0264-6021:3410491. [↑](#endnote-ref-13)