HHS Public Access

Author manuscript

J Proteome Res. Author manuscript; available in PMC 2023 November 04.

Published in final edited form as:

J Proteome Res. 2022 November 04; 21(11): 2687–2702. doi:10.1021/acs.jproteome.2c00391.

The human plasma proteome during normal pregnancy

Adi L. Tarca[1,2,3], **Roberto Romero**[1,4,5,6,7], **Gaurav Bhatti**[1,2], **Francesca Gotsch**[1,2], **Bogdan Done**[1,2], **Dereje W. Gudicha**[1,2], **Eunjung Jung**[1,2], **Roger Pique-Regi**[1,6], **Stanley M. Berry**[1,2], **Tinnakorn Chaiworapongsa**[1,2], **Nardhy Gomez-Lopez**[1,2,8]

[1]Perinatology Research Branch, Division of Obstetrics and Maternal-Fetal Medicine, Division of Intramural Research, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, US Department of Health and Human Services, Bethesda, MD, and Detroit, MI, USA

[2]Department of Obstetrics and Gynecology, Wayne State University School of Medicine, Detroit, MI, USA

[3]Department of Computer Science, Wayne State University College of Engineering, Detroit, MI, USA

[4]Department of Obstetrics and Gynecology, University of Michigan, Ann Arbor, MI, USA

[5]Department of Epidemiology and Biostatistics, Michigan State University, East Lansing, MI, USA

[6]Center for Molecular Medicine and Genetics, Wayne State University, Detroit, MI, USA

[7]Detroit Medical Center, Detroit, MI, USA

[8]Department of Biochemistry, Microbiology and Immunology, Wayne State University School of Medicine, Detroit, MI, USA

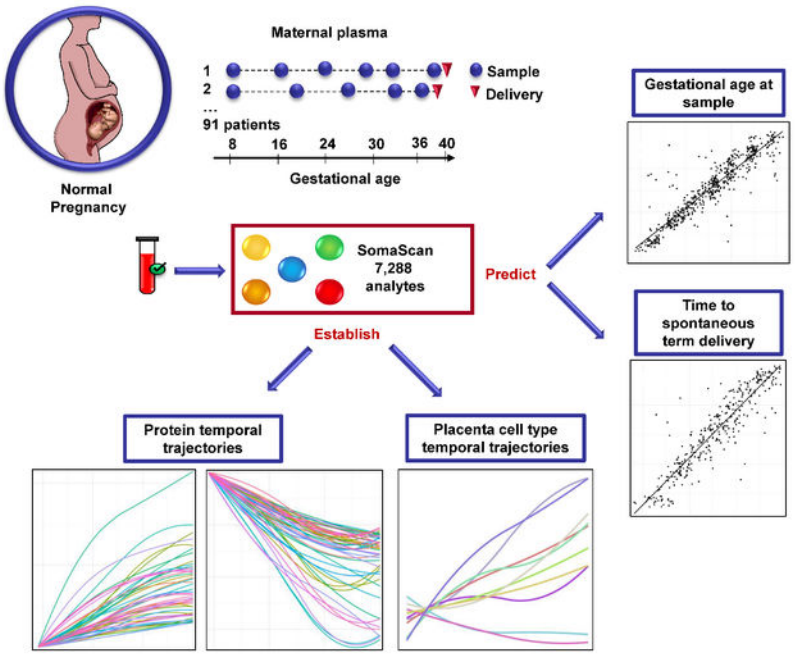
Abstract

The human plasma proteome is underexplored, despite its potential value for monitoring health and disease. Herein, using a recently developed aptamer-based platform, we profiled 7,288 proteins in 528 plasma samples from 91 normal pregnancies (Gene Expression Omnibus identifier GSE206454). The coefficient of variation was <20% for 93% of analytes (median 7%), and cross-platform correlation for selected key angiogenic and anti-angiogenic proteins was significant. Gestational age was associated with changes in 953 proteins, including highly modulated placenta- and decidua-specific proteins, and they were enriched in biological processes including regulation of growth, angiogenesis, immunity, and inflammation. The abundance of proteins corresponding to RNAs specific to populations of cells previously described by single-cell RNA-Seq analysis of the placenta was highly modulated throughout gestation. Furthermore, machine learning-based prediction of gestational age and of time from sampling to a term delivery compared favorably with transcriptomic models (mean absolute error of 2 weeks). These results suggested that the plasma proteome may provide a non-invasive readout of placental cellular dynamics and serve as blueprint for investigating obstetrical disease.

Correspondence atarca@med.wayne.edu (ALT); prbchiefstaff@med.wayne.edu (RR); and ngomezlo@med.wayne.edu (NGL).

The authors declare no conflicts of interest.

Graphical Abstract



Keywords

aptamer; biomarker; machine learning; proteomic standards; single-cell RNA signature

INTRODUCTION

Prediction, prevention, and treatment of obstetrical diseases such as preterm labor, preeclampsia, small for gestational age (SGA), and fetal death are currently sub-optimal due to the syndromic nature and multiple etiologies of these conditions. Therefore, a personalized medicine approach is required to avoid dependence on non-specific clinical symptoms and signs. The success of such an approach depends on the accuracy, practicality, and low cost of generating patient-specific molecular readouts from non-invasive samples, such as the maternal blood.

High-throughput molecular studies of the maternal blood were proposed based on the analysis of whole-blood (cellular) RNA[1–5] or cell-free RNA[6–9], plasma or serum proteome[10–12], and metabolome[13–14], among other techniques. Often, in such studies, gestational age at venipuncture was used as a physiologic endpoint to assess the reliability of the omics platforms and to gauge their suitability prior to attempting the prediction of pathology such as preterm birth and preeclampsia[5, 7–8, 10, 15–17]. While the optimal blood omics platform to use in pregnancy is still a subject of research and may depend on the condition of interest, our earlier work suggests that plasma proteomics may have an advantage over cellular RNA for predicting spontaneous preterm birth[5].

Blood proteins were shown to be comprehensive indicators of human health as they are purposefully secreted as effectors of biological processes or they leak into circulation upon cell damage or death [18–19]. In pregnancy, in particular, proteins can also enter into the circulation from gestational tissues and therefore may reflect maternal adaptations to the developing fetus. For instance, human chorionic gonadotrophin (hCG) in human blood has allowed the early detection of pregnancy[20–30], while maternal alpha fetoprotein[31–34] is being used for biochemical screening of congenital anomalies[35–42]. Proteins with high modulation in the maternal circulation during normal pregnancy, including pro-angiogenic placental growth factor (PlGF) and anti-angiogenic vascular endothelial growth factor receptor (VEGFR)-1, also known as soluble fms-like tyrosine kinase-1 (sFlt-1), were shown to be dysregulated in preeclampsia[43–54], fetal death[55–56], SGA[57], and maternal floor infarction[52, 58–59]. Increasing sensitivity and specificity of prediction of such pregnancy complications would however require identification of additional biomarkers.

To enable highly multiplexed profiling of the human proteome, an aptamer-based platform was developed[60–61] and utilized in obstetrics by our group and others for proteomic profiling of the maternal plasma[5, 10–12, 16] and amniotic fluid[62], among many other proteomic studies in pregnancy[63–64]. The recently expanded version of the SomaScan® platform v4.1, which allows simultaneous profiling of 7,288 proteins, i.e. over one third of the human proteome[65], has not been applied in obstetrics. Therefore, we sought to 1) evaluate this high-throughput proteomic platform in pregnancy and define the expected protein values for gestational age and maternal characteristics, and 2) to determine the accuracy of the proteomic profiles for prediction of gestational age and time from venipuncture to spontaneous term delivery. We believe that such contribution has the potential to enable the development and implementation of predictive models in obstetrics.

MATERIALS AND METHODS

Study Design

Based on a prospective longitudinal biomarker study [54, 66], we conducted a retrospective analysis of 528 plasma samples collected from 91 women who had a normal pregnancy. Only singleton pregnancies without major medical or surgical complications, who delivered an appropriate-for-gestational-age infant, with a birthweight between the 10th and 90th percentiles, without major congenital anomalies were included in the study. Patients were enrolled at the Center for Advanced Obstetrical Care and Research of the Perinatology Research Branch, NICHD, the Detroit Medical Center, and Wayne State University. For each of the 91 women, 3 to 7 plasma samples were obtained from the first trimester up to two days before the spontaneous onset of term labor [median number of samples=6, interquartile range (IQR)=5–6]. Blood samples were collected in tubes containing EDTA and plasma was separated by centrifugation (1300 × g, 10 min). Plasma samples were immediately stored at −80 °C until proteomic analysis. Maternal plasma protein abundance was determined by using the SomaScan® platform v4.1 and its reagents.

All patients provided written informed consent, and the use of biological specimens and clinical and ultrasound data for research purposes was approved by the Institutional Review Boards of Wayne State University and NICHD.

Proteomics Techniques

Maternal plasma protein abundance was determined by using the SomaScan® platform v4.1, which is based on SOMAmer® (Slow Off-rate Modified Aptamer) reagents. This platform allowed multiplexed quantification of 7,288 analytes corresponding to 6,596 unique human protein targets in maternal plasma samples[60–61, 67]. Of these, 88.7% of proteins had one single assay, 10.4% had two, and less than 1% of proteins had 3 to 9 assays present on the platform. The experiments were run in batches of up to 85 samples per plate.

The plasma samples were diluted and then incubated with the respective SOMAmer® mixes pre-immobilized onto streptavidin-coated beads. The beads were washed to remove all non-specifically bound proteins and other matrix constituents. Proteins that remained specifically bound to their cognate SOMAmer® reagents were tagged by using an NHS-biotin reagent. After the labeling reaction, the beads were exposed to an anionic competitor solution that prevents non-specific interactions from reforming after dissociating.

Using this approach, pure cognate-SOMAmer® complexes and unbound (free) SOMAmer® reagents are released from the streptavidin beads using ultraviolet light that cleaves the photo-cleavable linker. The photo-cleavage eluate, which contains excess anionic competitor and all SOMAmer® reagents (some bound to a biotin-labeled protein and some free), was separated from the beads and then incubated with a second streptavidin-coated bead that binds the biotin-labeled proteins and the biotin-labeled protein-SOMAmer® complexes. The free SOMAmer® reagents were then removed during subsequent washing steps. In the final elution step, protein-bound SOMAmer reagents were released from their cognate proteins, using denaturing conditions. These SOMAmer® reagents were then hybridized to custom DNA microarrays. The Cyanine-3 signal from the SOMAmer® reagent was detected and measured on microarrays[60–61, 67]. Proteomics profiling was performed by SomaLogic, Inc. (Boulder, CO). In addition to the SomaScan® platform data generated herein, data for PlGF, sFlt-1 and soluble endoglin (sEng) were previously determined by immunoassays (R&D Systems, Minneapolis, MN, USA).[48] The inter- and intra-assay coefficients of variation of the assays were 1.4% and 3.9% for sFlt-1, 2.3% and 4.6% for sEng, and 6.02% and 4.8% for PlGF, respectively. The sensitivity of assays were 16.97 pg/ml for sFlt-1, 0.08 ng/ml for sEng, and 9.52 pg/ml for PlGF. Sample collection methods, biospecimen processing, and validation of the assays used were previously reported in greater detail [54, 68].

Statistical Analysis

Data reproducibility: The proteomic data preprocessing, including an adaptive normalization by maximum likelihood (ANML) step and a calibration step, were performed by SomaLogic, Inc. The goal of these steps was to make data comparable across samples by calculating plate-specific and analyte-specific scale factors. Based on such scale factors, a quality control flag was assigned to each sample and each analyte[67]. Using preprocessed data for samples and analytes that passed quality controls, the Spearman's correlation coefficient and coefficient of variation for each protein were determined based on 14 samples collected from 2009–2010 and profiled in duplicates in different batches. These samples spanned the full range of gestational ages considered (10.4–39.4 weeks), capturing gestational age-related variability in the proteome, and hence provided an opportunity to

observe correlations between duplicate values. The coefficient of variation from duplicates was determined by a method that accounted for the measurement error being potentially dependent on the mean protein abundance[69]. Proteomic data and sample annotation is available from the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE206454).

Determining sources of variability in the meta-proteome: The proteome data for all 519 samples and 6,277 analytes that passed quality checks were analyzed using principal components (PC) analysis to reduce the dimensionality of the proteins to a few PCs (meta-proteins) using the *pca* function in the *PCAtools* R package. Next, the *eigencorplot* function from the same package was used to determine Pearson's correlation coefficients between each PC and maternal characteristics and gestational age. A p-value <0.05 was considered a significant result.

Differential abundance analysis: Protein abundance expressed as relative fluorescence units (RFU) was log (base 2) transformed to improve normality. Linear mixed-effect models with quadratic splines (one knot) were used to model protein abundance as a function of gestational age. Briefly, such a model assumes that $\log_2$ protein data throughout one-half of the gestational age span can be adequately modeled using a quadratic function of gestational age. Maternal age, parity, body mass index (BMI), and smoking status were considered as covariates and retained if they improved the model fit for a given protein. Covariate selection was based on the significance ($p<0.05$) of likelihood ratio tests implemented in the *glmerselect* function available in the *StatisticalModels* package under the R statistical language and environment (www.r-project.org). The models included patient identifiers as random effects to account for the repeated and likely correlated measurements from the same patient. Protein abundance was considered to have changed significantly with gestational age if the fold change was >1.25 and false discovery rate (FDR)[70] adjusted p-value (q-value) was <0.1. Fold change was defined as the ratio of highest to lowest mean protein abundance across the 9 to 40 weeks gestational age span. Linear mixed-effects models were fit using the *lme4* package[71].

Clustering proteins by average profile: The expected protein abundances determined by linear mixed-effects models across the 9 to 40 weeks of gestational-age span were used to perform hierarchical clustering of protein trajectories. A correlation distance measure was used in the clustering so that proteins with a similar trend vs. gestation but possibly different magnitude of changes were clustered together. Clustering was performed with the *WGCNA* package[72].

Gene ontology enrichment analysis: Proteins were mapped to Entrez gene database[73] identifiers based on SomaLogic, Inc. protein annotation, and then to gene ontology[74]. Biological processes over-represented among a given list of proteins (e.g. those differentially expressed with gestational age) were identified by using Fisher's exact tests. Gene ontology terms with three or more hits and having an adjusted enrichment q-value <0.1 were considered significantly enriched. Enrichment analysis was performed with the *GOStats* package[75] in Bioconductor[76]. The reference list used in all enrichment analyses was the

list of genes corresponding to the proteins profiled in the study, which is in line with best practices in the field[77–78]. This ensures that the enrichment we attributed to differential abundance with advancing gestation is not confounded by any biases in the design of the SomaScan platform relative to the universe of all known human proteins.

Quantification of single cell RNA-seq signatures of the placenta: To quantify the expression of previously defined signatures of cell sub-types identified by single cell RNA studies of the placenta[79], we have first selected the RNAs specific (most highly abundant 20 RNAs) to a given cell type that had a corresponding protein in our dataset. Then, the $\log_2$ protein abundance was transformed into a Z-scores based on the mean and standard deviation observed in samples at 8–16 weeks for each protein. Average Z-scores for each signature were used as a summary and further tested for association with gestational age at venipuncture using the same types of linear mixed-effects models described above for analysis of data of individual proteins.

Prediction of gestational age and time from sample to delivery: Predictive models for gestational age at venipuncture were based on all samples profiled, while prediction of time from venipuncture to delivery was limited to samples from patients who had a spontaneous term delivery. Prediction models were fit and evaluated using a leave-one-out cross-validation procedure. With this approach, a random forest model[80] was fit using data from all but one patient including all corresponding longitudinal samples, and it was then applied to the data of the patient left out during model training. Lasso regression[81], a procedure designed for fitting a continuous response variable (e.g. gestational age) using more predictors (i.e. proteins) than available samples, was utilized for multi-variate protein selection, and data for the selected proteins were used as input in random forest models. Prediction performance metrics were the Pearson correlation coefficient between actual and predicted values. The root mean squared error, i.e., the standard deviation of prediction errors (error=actual - predicted), and mean absolute error were also determined to enable direct comparison with previous reports. Lasso regression and random forest models fitting were implemented in *glmnet* and *randomforest* packages, respectively under the R statistical language and environment.

RESULTS AND DISCUSSION

Proteomic signals in the maternal blood are known to be correlated with both physiologic and pathologic endpoints in pregnancy. Large studies of maternal blood proteins were based on targeted profiling of specific angiogenic (PlGF) and anti-angiogenic factors (sFlt-1 and sEng)[52, 61, 82], or targeted profiling of pro-inflammatory proteins, cytokines, and chemokines[83–85]. Given the sub-optimal performance of current biomarkers for early prediction of obstetrical complications, high-throughput discovery platforms have been proposed to identify novel candidate biomarkers. Using earlier versions of the SomaScan® platform, measuring up to 1,310 proteins (v2 and v3), physiologic changes with gestational age[12], as well as pathologic perturbations in preeclampsia[10–11, 16], placenta accreta spectrum[86], and spontaneous preterm birth[5], have been reported, hence spurring the interest in this omics platform. In the current study, we utilized the SomaScan® platform v4.1 to

generate data for 7,288 proteomic assays in blood samples collected longitudinally from pregnant women throughout gestation.

Characteristics of the Study Population

The study population included 91 pregnant women, 32% (29/91) of these were nulliparous, the median maternal age was 23 years (IQR: 20–26), and the median BMI was 25.8 kg/m^2 (IQR: 22.5–30.5). All patients delivered at term gestation [median gestational age of 39.6 (IQR: 38.8–40.7) weeks] appropriate-for-gestational-age[87] neonates with a median birth weight of 3,400 (IQR: 3,137.5–3,702.5) grams (see Supplementary Table 1). The generation of data in a majority (92%) African-American population is important given the higher rate of pregnancy complications in this group of women. However, a more diverse cohort would have been ideal.

Proteomic Data Reproducibility

Of the 528 samples profiled, data for 519 (98.3%) of the samples passed the quality control checks, while of the 7,288 human protein assays, 6,277 (86%) passed the calibration filter and were included in the analyses. Based on data collected from 14 duplicate samples, the median of the Spearman's correlation coefficients of protein abundance was 0.77 (IQR 0.64%-0.87%), and of the coefficients of variation was 7% (IQR 4.5%-11.4%). Of importance, the coefficients of variation were below 20% for 93% of the assays. The median coefficient of variation found herein was somewhat higher than the 5% previously reported for a lower level of multiplexing of 4,000 proteins[88]. This can be explained in part by the higher multiplexing, the longer storage duration of the blood samples prior to profiling, and perhaps the lower biological variability in the current study compared to that in the study of Tin et al. [88] which included both male and females of various ages. Longer storage time is expected to negatively affect the reproducibility, while higher biological variability is essential to put in perspective the magnitude of technical noise in the data relative to biological variability.

Moreover, the proteomic data from the SomaLogic SomaScan® Platform v4.1 was well correlated with enzyme-linked immunosorbent assay (ELISA)-based measurements for key biomarkers in pregnancy, including PlGF (ρ=0.87) and sFlt-1 (ρ=0.77), but only modestly for sEng (ρ=0.38) ($p<0.001$ for all Figure 1).

The high correlation of individual patient data between SomaScan and ELISA-derived data for PlGF and sFlt-1, currently used in screening to prevent preeclampsia[89–90], further supports the utility of this platform for research in obstetrics. The lower correlation between ELISA-based concentrations and SomaScan for sEng may be reflective of differences in epitopes being recognized between technologies and possible modifications such as misfolding, protein-protein interactions, and impact of genetic variants in protein structure. ELISA-based sEng was shown to add predictive value relative to PlGF and sFlt-1 when distinguishing between women with chronic hypertension from those who develop superimposed preeclampsia[54], and therefore, further studies on the value of SomaScan-derived sEng measurements in obstetrics is warranted.

Global Sources of Variation in the Maternal Plasma Proteome

Next, we aimed to assess the primary sources of variability in the plasma proteome by deriving PCs and correlating such meta-proteins with gestational age at venipuncture and maternal characteristics. We found a substantial modulation of proteomics data with advancing gestational age and, to a lesser extent, with maternal characteristics (Figure 2). Six of the top 10 PCs (ranked by % of variance explained), which explained 38% of the variance in the data (Figure 2A), were significantly correlated with gestational age at venipuncture ($p<0.05$ for all) (Figure 2B). As an example, the Pearson correlation of the PC4 and gestational age was $\rho=0.66$, and PC6 was $\rho=0.36$ (Figure 2 B, C, D). PCs derived from proteomics data were also correlated with maternal BMI (7/10 PCs), parity (5/10 PCs), and maternal age (2/10 PCs) ($p<0.05$ for all), although the magnitude of such correlations was lower than that observed for gestational age at venipuncture (e.g., Pearson $\rho=-0.24$ for PC3; $\rho=-0.22$ for PC4 for correlation with BMI) (Figure 2, B, E, F).

Protein Level Changes with Gestational Age

Subsequently, we analyzed the data from individual proteomic assays in relation to gestational age and maternal characteristics by fitting multivariate linear-mixed effects models. Of the 6,277 human protein targets that passed quality control filters, 953 (15.2%) changed in abundance as a function of gestational age while accounting for maternal characteristics (fold change >1.25 and q-value <0.1) (Figure 3a, Supplementary File 1 and Supplementary File 2).

The sizable fraction of the maternal plasma proteins modulated in abundance as a function of gestational age in normal pregnancy can be understood as protein abundance reflects both fetal development and maternal adaptations throughout gestation. This result is in line with the previously reported estimate (10% of proteins modulated during gestation), when considering that the later estimate was obtained using more conservative cut-offs ($q<0.1$, fold-change >1.5)[12]. The fraction of blood proteins modulated with advancing gestation that we report herein is about one order of magnitude higher than that in whole blood RNA (2.3%) determined in a similar population based on comparable sample size, modeling methods, and the same significance cut-offs[91]. This suggests that the plasma proteome could be a more abundant source of disease biomarkers relative to the cellular transcriptome, given the known association between modulation with advancing gestation in normal pregnancy and dysregulation in obstetrical disease, such as preeclampsia (odds ratio = 4.3)[11].

The $\log_2$ fold changes in protein abundance with gestational age were highly correlated with estimates derived from an independent set of patients profiled by using a lower throughput version (v3) of the SomaScan® platform[12]. This finding was based on all proteins with significant change with gestational age in the current study that were also measured in Romero et al. 2017[12] (N=255 proteins, $\rho=0.89$, $p<0.01$, see Figure 3B). Not only specific proteins were confirmed herein to be highly modulated with gestational age (>5-fold change), such as PlGF, Siglec-6, glypican-3, and Prolactin, but also the magnitude of changes were highly correlated between studies, hence providing an *in-silico* validation of the current results, despite a small fraction [9.4% (24/255)] of proteins with opposite direction of change between studies. However, the current study has identified almost 9-fold

as many proteins changing with gestational age than in the previous report[12]. Notably, several proteins assayed by the newly developed SomaScan® v4.1 platform, but not by the previous platform version, were highly modulated with gestational age and were placenta- and decidua-specific. These included: ABP1 (amiloride-sensitive amine oxidase, copper-containing), DLK1 (Protein delta homolog 1), EMBP (bone marrow proteoglycan), IGSF3 (immunoglobulin superfamily member 3), and SVEP1 (sushi, von Willebrand factor type A, EGF, and pentraxin domain-containing protein) (fold change >5 for all, see Supplementary File 1).

SVEP1, a protein highly expressed in the human placenta[92], has been characterized as an extracellular matrix protein important for cell adhesion[92] and plays a role in lymphangiogenesis[93–94], septic shock[95–96] and endotoxinemia[97], atherosclerosis, and coronary artery disease[98–100]. ABP1, also called diamine oxidase (DAO), is regulated by estrogens and is mainly localized in the decidua[101–102]. Maternal plasma ABP1 levels rise exponentially during the first 20 weeks of gestation and are potentially indicative of fetoplacental integrity[103–104]. IGSF3, also known as EWI-3, is widely expressed in the placenta, kidney, and lung. Although placental gene expression of IGSF3 has been described in a rat model of placental insufficiency[105], data in pregnant human subjects are lacking. Pregnancy-associated plasma protein A (PAPP-A) and the proform of eosinophil major basic protein (EMBP) are produced by the placenta.. EMBP has been implicated in placenta-mediated obstetrical syndromes such as preeclampsia[106], SGA[107], and preterm birth[108]. DLK1, also known as fetal antigen 1 and pre-adipocyte factor 1, is a transmembrane protein encoded by the DLK1 gene expressed in the placenta, yolk sac, fetal liver, adrenal cortex, and pancreas and in the beta cells of the islets of Langerhans in the adult pancreas[109]. In the placenta, DLK1 is specifically expressed by the stromal cells of the villi that are in close contact with the vasculature[109–110] and has been identified as a potential biomarker of fetal growth restriction[111–112].

Distinct Types of Longitudinal Trajectories and Functional Profiling

Given the complex patterns of protein abundance modulation during normal pregnancy (see examples in Supplementary File 2), we sought to identify clusters of proteins based on the similarity of their longitudinal trajectories. Figure 4 depicts the trajectories of the top 50 most highly modulated proteins in each of the three protein clusters identified in this analysis.

Cluster 1 includes 249 proteins with a trajectory characterized by a steady increase in abundance throughout gestation. Member proteins were specifically associated with regulation of growth, angiogenesis, immune (e.g., T-helper 1 type immune response) and inflammatory processes (e.g., regulation of macrophage cytokine production) (Figure 5, Supplementary File 3). Cluster 2 includes 151 proteins with trajectories that had an early decrease or remained unchanged early in pregnancy followed by an increase later in pregnancy. Cluster 2 proteins were involved in the regulation of blood vessel remodeling, response to steroid hormone, and metabolic and catabolic processes (Figure 5, Supplementary File 3). Cluster 3 includes 522 proteins that demonstrate an overall decrease in abundance throughout gestation and were associated with various processes ranging from

immune and defense response to hemostasis. The complexity of immune and inflammatory processes regulation is highlighted by the involvement of proteins with both increasing (Cluster 1 and 2) and decreasing (Cluster 3) trajectories (Figure 5).

Placental Single-Cell RNA Signatures are Modulated with Advancing Gestation

We have next evaluated whether the meta-proteome corresponding to RNAs specific to previously described placental-derived cell types changes with advancing gestation and hence can provide a non-invasive molecular readout of their activity in the placenta. Figure 6 shows that the activity of placental cell types undergoes a complex modulation as captured by the corresponding plasma protein abundance. Among the top ten single-cell signatures most strongly associated with gestational age, the signatures of B cells, cytotrophoblasts, dendritic macrophages, extravillous trophoblast, and stromal cells were increased, while the signatures of hematopoietic stem cell and monocytes were decreased ($q<0.1$). The decidual cell signature displayed a complex pattern of modulation culminating to an increase at term gestation (Figure 6). Tracking placental single cell signatures throughout gestation was previously shown using amniotic fluid cell free RNA[113] and maternal blood cell free[6] and cellular RNA[114]. Here we show that maternal plasma proteome also captures placenta cell population activity during gestation.

Prediction of Gestational Age at Venipuncture and of Time-from-Sample-to-Delivery

Given the strong modulation of maternal plasma proteins during normal pregnancy, we sought to determine whether gestational age could be ascertained by using machine learning methods based solely on proteomic profiles of pregnant women. Indeed, random forest models, trained using proteins selected by lasso regression and evaluated via cross-validation, significantly predicted the gestational ages at venipuncture in patients not included during protein selection and model training (test sets). The Pearson correlation between actual and predicted gestational ages was 0.92 ($p<0.001$), the root mean squared error (RMSE) of predictions was 3.1 weeks, and the mean absolute error (MAE) was 1.99 weeks (Figure 7A). Of note, the samples with the largest (top 10%) gestational age prediction errors, among those shown in Figure 7A, did not tend to cluster by patient, nor by specific maternal characteristics, suggesting lack of systematic biases.

We next used the proteomics data to predict the time from venipuncture to delivery using the data form 334 samples collected from the subset of 61/91 women included in the study who had a spontaneous term delivery, hence excluding the cases for whom the pregnancy was truncated by a selective cesarean section. The accuracy of the random forest model for this analysis was similar to that of the model for prediction of gestational age (Pearson correlation 0.92, $p<0.001$; RMSE=3.0 weeks; MAE=2.15 weeks, Figure 7B). This suggests that eventual biases in the gold standard of gestational age, defined by last menstrual period (LMP) and ultrasound, were minimal. The fraction of term deliveries predicted within one week of the actual delivery was 26%, 33%, and 43% based on samples collected in the first, second, and third trimesters, respectively, while LMP and ultrasound-based accuracy was 57%, almost identical to the 55.1% reported in the literature[115].

Finally, when 42 of the best protein predictors of time from venipuncture to a term delivery identified herein, that were also profiled in Romero et al. 2017[12], were used to predict time from sample to delivery in the Romero et al. 2017 dataset, the cross-validated prediction was still high (Pearson correlation 0.95, RMSE=2.99 weeks; MAE=2.35 weeks), suggesting external validity of the protein signature predicting time from sample to delivery in the absence of obstetrical disease. The correlation between the plasma proteome and the time from venipuncture to delivery assessed via multi-variate models (Pearson correlation of predicted vs. actual, ρ=0.91) was stronger compared to estimates obtained using an unbiased analysis of whole blood transcriptome via microarrays (ρ=0.83)[5]. This is in line with previous evidence of higher accuracy for the plasma proteome compared to whole blood transcriptome for prediction of subsequent spontaneous preterm birth based on analyses of the same blood samples[5]. SomaScan-based prediction of time from sampling to term delivery, however, was similar to estimates reported based on more than 100 placenta-, immune-, and fetal liver-specific cell-free RNAs profiled by real-time polymerase chain reaction in maternal circulation (ρ=0.89)[7]. The stronger prediction reported in the studies that used cell-free RNA compared to those using cellular transcriptomics can be explained by the fact that cell-free transcripts are more likely to be derived from the placenta, fetus, and maternal reproductive tissues[116], and also by the targeting of analysis based on biological plausibility. Indeed, when only proteins found herein to be predictive of time from sampling to delivery were used to narrow the pool of candidate predictor proteins in the Romero et al. 2017 dataset[12], the accuracy was slightly better (Pearson correlation 0.95) compared to the estimate (Pearson correlation 0.92) obtained using an unbiased analysis. Other approaches to narrow the search of biomarkers based on omics profiles leverage single cell-derived signatures of the placenta[6, 79, 113–114, 117–118].

The Effect of Maternal Characteristics on Protein Abundance

Among all the maternal characteristics and obstetrical history covariates considered, the maternal BMI had the strongest effect on the plasma protein abundance. Indeed, of the 6,277 human protein targets that passed quality control filters, 211 (3.4%) were significantly associated with BMI. These included fatty acid-binding protein (FABP), leptin, insulin-like growth factor-binding protein 1 (IGFBP-1), PlGF, and matrix metalloproteinase (MMP)-7 ($q<0.1$, Supplementary File 4). The top two proteins in this analysis, FABP3 and LEP (Supplementary File 4), were also among the top three most important predictors the percentage body fat out of 4000 proteins profiled with the SomaScan® platform in an independent study of 6,000 individuals[18]. Leptin, an appetite and metabolism regulator[18], was described also as a marker of placental function[119], and was implicated in several pathologies in pregnancy such as gestational age diabetes[120] and preterm delivery[121]. FABP3, expressed in adipocytes, is strongly linked to metabolic and inflammatory pathways[18]. Biological processes associated with BMI-modulated proteins included eating behavior, response to nutrient levels, regulation of growth, and inflammatory-related processes. In multivariate models, nulliparity was associated with an increase in Trefoil factor 3 (TFF3) and Serine/threonine-protein kinase (DCAK1) and with a decrease in Mesoderm development candidate 1 (MESD1). Moreover, maternal age was associated with a decrease in Collagen alpha-2(XI) chain (COL11A2) and with an increase in Spectrin alpha chain, non-erythrocytic 1 (SPTA2). These results suggest that among maternal

characteristics, maternal BMI, in particular, needs to be accounted for in analyses involving pathology to avoid confounding of effects.

Proteomic Standards to Enable the Discovery of Disease-Related Proteomic Perturbations

Last, given the current and potential use of maternal plasma protein dysregulation in predicting obstetrical complications, we sought to generate proteomic standards that would allow the comparison of data across studies and enable the discovery of new disease-related protein perturbations. To this end, we fitted the protein relative fluorescence data ($\log_2$ thereof) as a function of gestational age and maternal characteristics so that the expected levels could be determined and used to derive multiple of the mean values. The approach to convert raw protein concentration data into multiples of the mean or median for gestational age and maternal characteristics and obstetrical history is currently being used in obstetrics as a way to overcome assay- and population-specific biases[122–125]. For instance, for preeclampsia screening, the *a priori* risk of disease can be derived from maternal characteristics and obstetrical history (prior risk factors). This risk estimate, is then combined with biomarker-based risk scores that need to be independent of prior risk factors so that the two probabilities can be combined into a posterior risk of developing the disease[54, 82, 126].

The proteomics standards proposed herein were implemented and made available as an *R* software package called *SomaPreg*, which is available from the author's website at http://bioinformaticsprb.med.wayne.edu/software/ and also as Supplementary File 5. The functionality of the software package is illustrated in Figure 8 and Supplementary File 6. Blood sample annotation data, paired with normalized proteomic data (RFU) are used as inputs (Figure 8A). The expected proteomic abundance is determined from prediction models stored in the software package. Such models also allow exploring the effect of maternal characteristics (Figure 8B) and calculating multiple of the mean values for gestational age and maternal characteristics (Figure 8C). This functionality encapsulated in the *SomaPreg* package is expected to facilitate the discovery of proteomic disease signatures and the implementation of risk prediction models by removing physiologic variability from proteomic signals.

Previous work suggested that the normalization of RFUs via internal control samples present on the SomaScan® platform already allows for a significant prediction of gestational age at venipuncture across batches and cohorts[16]. However, prediction of preeclampsia across cohorts was not feasible, likely due to the heterogeneity of the disease and much weaker within-cohort proteomic dysregulation with preeclampsia than that with gestational age. Improving cross-cohort prediction of disease based on biomarkers requires transforming the data into multiple of the mean (MoM) values to account not only for gestational age but also maternal characteristics that affect such measurements in control pregnancies. Risk models in pregnancy based on MoM-transformed biophysical and biochemical data were previously described[54, 82, 90, 127].

CONCLUSIONS

Herein, we have presented the most comprehensive characterization of the maternal plasma proteome in normal pregnancy. The proteome undergoes dramatic modulation with advancing gestation and is substantially affected by maternal body mass index. The models we have proposed and implemented in freely available software may enable the discovery of disease-related perturbations and the implementation of disease-prediction models in obstetrics.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The authors are grateful to Clare Paterson, Erin Hales, Michael Hinterberg and Sama Shrestha, and Will Schwarzmann from SomaLogic, Inc. for feedback on the manuscript, and to Marcia Arenas-Hernandez (Wayne State University) for support with sample processing. Maureen McGerty (Wayne State University) is acknowledged for support with manuscript proofreading.

This research was supported, in part, by the Perinatology Research Branch, Division of Obstetrics and Maternal-Fetal Medicine, Division of Intramural Research, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, U.S. Department of Health and Human Services (NICHD/NIH/DHHS); and, in part, with Federal funds from NICHD/NIH/DHHS under Contract No. HHSN275201300006C. ALT and NGL were also supported by the Wayne State School of Medicine Perinatal Initiative. RR has contributed to this work as part of his official duties as an employee of the United States Federal Government.

ABBREVIATIONS

ANML	adaptive normalization by maximum likelihood
APB1	amiloride-binding protein 1
BMI	body mass index
DAO	diamine oxidase
DLK1	delta-like homolog 1
ELISA	enzyme-linked immunosorbent assay
EMBP	bone marrow proteoglycan
FABP	fatty acid-binding protein
FDR	false discovery rate
hCG	human chorionic gonadotrophin
IGSF3	immunoglobulin superfamily member 3
IQR	interquartile range
LMP	last menstrual period

MAE	mean absolute error
MMP	matrix metalloproteinase
MoM	multiple of the mean
PAPP-A	pregnancy-associated plasma protein A
PC	principal component
PlGF	placental growth factor
RFU	relative fluorescence unit
RMSE	root mean squared error
sFlt-1	soluble fms-like tyrosine kinase-1
sEng	soluble endoglin
SGA	small for gestational age
SOMAmer®	slow off-rate modified aptamer
SVEP1	sushi, von Willebrand factor type A, EGF, and pentraxin domain-containing protein 1
VEGFR	vascular endothelial growth factor receptor

References

1. Gomez-Lopez N; Romero R; Galaz J; Bhatti G; Done B; Miller D; Ghita C; Motomura K; Farias-Jofre M; Jung E; Pique-Regi R; Hassan SS; Chaiworapongsa T; Tarca AL, Transcriptome changes in maternal peripheral blood during term parturition mimic perturbations preceding spontaneous preterm birthdagger. Biol Reprod 2022, 106 (1), 185–199. [PubMed: 34686873]
2. Tarca AL; Romero R; Erez O; Gudicha DW; Than NG; Benshalom-Tirosh N; Pacora P; Hsu CD; Chaiworapongsa T; Hassan SS; Gomez-Lopez N, Maternal whole blood mRNA signatures identify women at risk of early preeclampsia: a longitudinal study. J Matern Fetal Neonatal Med 2020, 1–12. [PubMed: 29804488]
3. Tarca AL; Romero R; Xu Z; Gomez-Lopez N; Erez O; Hsu CD; Hassan SS; Carey VJ, Targeted expression profiling by RNA-Seq improves detection of cellular dynamics during pregnancy and identifies a role for T cells in term parturition. Sci Rep 2019, 9 (1), 848. [PubMed: 30696862]
4. Heng YJ; Pennell CE; McDonald SW; Vinturache AE; Xu J; Lee MW; Briollais L; Lyon AW; Slater DM; Bocking AD; de Koning L; Olson DM; Dolan SM; Tough SC; Lye SJ, Maternal Whole Blood Gene Expression at 18 and 28 Weeks of Gestation Associated with Spontaneous Preterm Birth in Asymptomatic Women. PLoS One 2016, 11 (6), e0155191. [PubMed: 27333071]
5. Tarca AL; Pataki BA; Romero R; Sirota M; Guan Y; Kutum R; Gomez-Lopez N; Done B; Bhatti G; Yu T; Andreoletti G; Chaiworapongsa T; Consortium DPBPC; Hassan SS; Hsu CD; Aghaeepour N; Stolovitzky G; Csabai I; Costello JC, Crowdsourcing assessment of maternal blood multi-omics for predicting gestational age and preterm birth. Cell Rep Med 2021, 2 (6), 100323. [PubMed: 34195686]
6. Tsang JCH; Vong JSL; Ji L; Poon LCY; Jiang P; Lui KO; Ni YB; To KF; Cheng YKY; Chiu RWK; Lo YMD, Integrative single-cell and cell-free plasma RNA transcriptomics elucidates placental cellular dynamics. Proc Natl Acad Sci U S A 2017, 114 (37), E7786–E7795. [PubMed: 28830992]

7. Ngo TTM; Moufarrej MN; Rasmussen MH; Camunas-Soler J; Pan W; Okamoto J; Neff NF; Liu K; Wong RJ; Downes K; Tibshirani R; Shaw GM; Skotte L; Stevenson DK; Biggio JR; Elovitz MA; Melbye M; Quake SR, Noninvasive blood tests for fetal development predict gestational age and preterm delivery. Science 2018, 360 (6393), 1133–1136. [PubMed: 29880692]
8. Moufarrej MN; Vorperian SK; Wong RJ; Campos AA; Quaintance CC; Sit RV; Tan M; Detweiler AM; Mekonen H; Neff NF; Baruch-Gravett C; Litch JA; Druzin ML; Winn VD; Shaw GM; Stevenson DK; Quake SR, Early prediction of preeclampsia in pregnancy with cell-free RNA. Nature 2022, 602 (7898), 689–694. [PubMed: 35140405]
9. Rasmussen M; Reddy M; Nolan R; Camunas-Soler J; Khodursky A; Scheller NM; Cantonwine DE; Engelbrechtsen L; Mi JD; Dutta A; Brundage T; Siddiqui F; Thao M; Gee EPS; La J; Baruch-Gravett C; Santillan MK; Deb S; Ame SM; Ali SM; Adkins M; DePristo MA; Lee M; Namsaraev E; Gybel-Brask DJ; Skibsted L; Litch JA; Santillan DA; Sazawal S; Tribe RM; Roberts JM; Jain M; Hogdall E; Holzman C; Quake SR; Elovitz MA; McElrath TF, RNA profiles reveal signatures of future health and disease in pregnancy. Nature 2022, 601 (7893), 422–427. [PubMed: 34987224]
10. Tarca AL; Romero R; Benshalom-Tirosh N; Than NG; Gudicha DW; Done B; Pacora P; Chaiworapongsa T; Panaitescu B; Tirosh D; Gomez-Lopez N; Draghici S; Hassan SS; Erez O, The prediction of early preeclampsia: Results from a longitudinal proteomics study. PLoS One 2019, 14 (6), e0217273. [PubMed: 31163045]
11. Erez O; Romero R; Maymon E; Chaemsaithong P; Done B; Pacora P; Panaitescu B; Chaiworapongsa T; Hassan SS; Tarca AL, The prediction of late-onset preeclampsia: Results from a longitudinal proteomics study. PLoS One 2017, 12 (7), e0181468. [PubMed: 28738067]
12. Romero R; Erez O; Maymon E; Chaemsaithong P; Xu Z; Pacora P; Chaiworapongsa T; Done B; Hassan SS; Tarca AL, The maternal plasma proteome changes as a function of gestational age in normal pregnancy: a longitudinal study. Am J Obstet Gynecol 2017, 217 (1), 67 e1–67 e21.
13. Handelman SK; Romero R; Tarca AL; Pacora P; Ingram B; Maymon E; Chaiworapongsa T; Hassan SS; Erez O, The plasma metabolome of women in early pregnancy differs from that of non-pregnant women. PLoS One 2019, 14 (11), e0224682. [PubMed: 31726468]
14. Luan H; Meng N; Liu P; Fu J; Chen X; Rao W; Jiang H; Xu X; Cai Z; Wang J, Non-targeted metabolomics and lipidomics LC-MS data from maternal plasma of 180 healthy pregnant women. Gigascience 2015, 4, 16. [PubMed: 25859341]
15. Jehan F; Sazawal S; Baqui AH; Nisar MI; Dhingra U; Khanam R; Ilyas M; Dutta A; Mitra DK; Mehmood U; Deb S; Mahmud A; Hotwani A; Ali SM; Rahman S; Nizar A; Ame SM; Moin MI; Muhammad S; Chauhan A; Begum N; Khan W; Das S; Ahmed S; Hasan T; Khalid J; Rizvi SJR; Juma MH; Chowdhury NH; Kabir F; Aftab F; Quaiyum A; Manu A; Yoshida S; Bahl R; Rahman A; Pervin J; Winston J; Musonda P; Stringer JSA; Litch JA; Ghaemi MS; Moufarrej MN; Contrepois K; Chen S; Stelzer IA; Stanley N; Chang AL; Hammad GB; Wong RJ; Liu C; Quaintance CC; Culos A; Espinosa C; Xenochristou M; Becker M; Fallahzadeh R; Ganio E; Tsai AS; Gaudilliere D; Tsai ES; Han X; Ando K; Tingle M; Maric I; Wise PH; Winn VD; Druzin ML; Gibbs RS; Darmstadt GL; Murray JC; Shaw GM; Stevenson DK; Snyder MP; Quake SR; Angst MS; Gaudilliere B; Aghaeepour N; Alliance for M; Newborn Health Improvement, t. G. A. t. P. P.; Stillbirth; the Prematurity Research Center at Stanford, U., Multiomics Characterization of Preterm Birth in Low- and Middle-Income Countries. JAMA Netw Open 2020, 3 (12), e2029655. [PubMed: 33337494]
16. Ghaemi MS; Tarca AL; Romero R; Stanley N; Fallahzadeh R; Tanada A; Culos A; Ando K; Han X; Blumenfeld YJ; Druzin ML; El-Sayed YY; Gibbs RS; Winn VD; Contrepois K; Ling XB; Wong RJ; Shaw GM; Stevenson DK; Gaudilliere B; Aghaeepour N; Angst MS, Proteomic signatures predict preeclampsia in individual cohorts but not across cohorts - implications for clinical biomarker studies. J Matern Fetal Neonatal Med 2021, 1–8.
17. Chen Y; He B; Liu Y; Aung MT; Rosario-Pabon Z; Velez-Vega CM; Alshawabkeh A; Cordero JF; Meeker JD; Garmire LX, Maternal plasma lipids are involved in the pathogenesis of preterm birth. Gigascience 2022, 11.
18. Williams SA; Kivimaki M; Langenberg C; Hingorani AD; Casas JP; Bouchard C; Jonasson C; Sarzynski MA; Shipley MJ; Alexander L; Ash J; Bauer T; Chadwick J; Datta G; DeLisle RK; Hagar Y; Hinterberg M; Ostroff R; Weiss S; Ganz P; Wareham NJ, Plasma protein patterns as comprehensive indicators of health. Nat Med 2019, 25 (12), 1851–1857. [PubMed: 31792462]

19. Ganz P; Heidecker B; Hveem K; Jonasson C; Kato S; Segal MR; Sterling DG; Williams SA, Development and Validation of a Protein-Based Risk Score for Cardiovascular Outcomes Among Patients With Stable Coronary Heart Disease. JAMA 2016, 315 (23), 2532–41. [PubMed: 27327800]
20. Mishell DR Jr.; Davajan V, Quantitative immunologic assay of human chorionic gonadotropin in normal and abnormal pregnancies. Am J Obstet Gynecol 1966, 96 (2), 231–9. [PubMed: 5921063]
21. Vaitukaitis JL; Braunstein GD; Ross GT, A radioimmunoassay which specifically measures human chorionic gonadotropin in the presence of human luteinizing hormone. Am J Obstet Gynecol 1972, 113 (6), 751–8. [PubMed: 4673805]
22. Landesman R; Saxena BB, Results of the first 1000 radioreceptorassays for the determination of human chorionic gonadotropin: a new, rapid, reliable, and sensitive pregnancy test. Fertil Steril 1976, 27 (4), 357–68. [PubMed: 944652]
23. Feldkamp CS; Pfeffer WH, The measurement of human chorionic gonadotropin for pregnancy testing. Henry Ford Hosp Med J 1982, 30 (4), 207–13. [PubMed: 7169378]
24. Nyberg DA; Filly RA; Mahony BS; Monroe S; Laing FC; Jeffrey RB Jr., Early gestation: correlation of HCG levels and sonographic identification. AJR Am J Roentgenol 1985, 144 (5), 951–4. [PubMed: 3885694]
25. Pittaway DE; Wentz AC, Evaluation of early pregnancy by serial chorionic gonadotropin determinations: a comparison of methods by receiver operating characteristic curve analysis. Fertil Steril 1985, 43 (4), 529–33. [PubMed: 3987923]
26. Birken S; Armstrong EG; Kolks MA; Cole LA; Agosto GM; Krichevsky A; Vaitukaitis JL; Canfield RE, Structure of the human chorionic gonadotropin beta-subunit fragment from pregnancy urine. Endocrinology 1988, 123 (1), 572–83. [PubMed: 2454811]
27. Klieber R; Panmoung W; Berger P; Wick G, [Monoclonal antibodies to non-assembled alpha and beta chain of human chorionic gonadotropin (hCG)]. Wien Klin Wochenschr 1990, 102 (10), 283–9. [PubMed: 1694052]
28. Chard T, Pregnancy tests: a review. Hum Reprod 1992, 7 (5), 701–10. [PubMed: 1639991]
29. Hinney B; Bertagnoli C; Tobler-Sommer M; Osmers R; Wuttke W; Kuhn W, Diagnosis of early ectopic pregnancy by measurement of the maternal serum to cul-de-sac fluid beta-hCG ratio. Ultrasound Obstet Gynecol 1995, 5 (4), 260–6. [PubMed: 7541302]
30. Vaitukaitis JL, Development of the home pregnancy test. Ann N Y Acad Sci 2004, 1038, 220–2. [PubMed: 15838116]
31. Wald N; Barker S; Peto R; Brock DJ; Bonnar J, Maternal serum alpha-fetoprotein and previous neural tube defects. Br J Obstet Gynaecol 1976, 83 (3), 213–6. [PubMed: 56196]
32. Weiss RR; Macri JN; Elligers K; Princler GL; McIntire R; Waldman TA, Amniotic fluid alpha-fetoprotein as a marker in prenatal diagnosis of neural tube defects. Obstet Gynecol 1976, 47 (2), 148–51. [PubMed: 55990]
33. Bond EB; Thompson W; Elwood JH; Cran GW, Evaluation of measurement of maternal plasma alpha-fetoprotein levels as a screening test for fetal neural tube defects. Br J Obstet Gynaecol 1977, 84 (8), 574–7. [PubMed: 70214]
34. Amniotic-fluid alpha-fetoprotein measurement in antenatal diagnosis of anencephaly and open spina bifida in early pregnancy. Second report of the U.K. Collaborative Study on Alpha-fetoprotein in Relation to Neural-tube Defects. Lancet 1979, 2 (8144), 651–62. [PubMed: 90757]
35. Cuckle HS; Wald NJ; Lindenbaum RH, Maternal serum alpha-fetoprotein measurement: a screening test for Down syndrome. Lancet 1984, 1 (8383), 926–9. [PubMed: 6201687]
36. Merkatz IR; Nitowsky HM; Macri JN; Johnson WE, An association between low maternal serum alpha-fetoprotein and fetal chromosomal abnormalities. Am J Obstet Gynecol 1984, 148 (7), 886–94. [PubMed: 6201071]
37. Heyl PS; Miller W; Canick JA, Maternal serum screening for aneuploid pregnancy by alpha-fetoprotein, hCG, and unconjugated estriol. Obstet Gynecol 1990, 76 (6), 1025–31. [PubMed: 1700348]
38. Crossley JA; Aitken DA; Connor JM, Prenatal screening for chromosome abnormalities using maternal serum chorionic gonadotrophin, alpha-fetoprotein, and age. Prenat Diagn 1991, 11 (2), 83–101. [PubMed: 1712100]

39. Wald NJ; Rodeck C; Hackshaw AK; Walters J; Chitty L; Mackinson AM, First and second trimester antenatal screening for Down's syndrome: the results of the Serum, Urine and Ultrasound Screening Study (SURUSS). Health Technol Assess 2003, 7 (11), 1–77.
40. Malone FD; Canick JA; Ball RH; Nyberg DA; Comstock CH; Bukowski R; Berkowitz RL; Gross SJ; Dugoff L; Craigo SD; Timor-Tritsch IE; Carr SR; Wolfe HM; Dukes K; Bianchi DW; Rudnicka AR; Hackshaw AK; Lambert-Messerlian G; Wald NJ; D'Alton ME, First-trimester or second-trimester screening, or both, for Down's syndrome. N Engl J Med 2005, 353 (19), 2001–11. [PubMed: 16282175]
41. Bredaki FE; Wright D; Matos P; Syngelaki A; Nicolaides KH, First-trimester screening for trisomy 21 using alpha-fetoprotein. Fetal Diagn Ther 2011, 30 (3), 215–8. [PubMed: 21952001]
42. Wright D; Syngelaki A; Bradbury I; Akolekar R; Nicolaides KH, First-trimester screening for trisomies 21, 18 and 13 by ultrasound and biochemical testing. Fetal Diagn Ther 2014, 35 (2), 118–26. [PubMed: 24356462]
43. Levine RJ; Maynard SE; Qian C; Lim KH; England LJ; Yu KF; Schisterman EF; Thadhani R; Sachs BP; Epstein FH; Sibai BM; Sukhatme VP; Karumanchi SA, Circulating angiogenic factors and the risk of preeclampsia. N Engl J Med 2004, 350 (7), 672–83. [PubMed: 14764923]
44. Chaiworapongsa T; Romero R; Kim YM; Kim GJ; Kim MR; Espinoza J; Bujold E; Goncalves L; Gomez R; Edwin S; Mazor M, Plasma soluble vascular endothelial growth factor receptor-1 concentration is elevated prior to the clinical diagnosis of pre-eclampsia. J Matern Fetal Neonatal Med 2005, 17 (1), 3–18. [PubMed: 15804781]
45. Levine RJ; Lam C; Qian C; Yu KF; Maynard SE; Sachs BP; Sibai BM; Epstein FH; Romero R; Thadhani R; Karumanchi SA, Soluble endoglin and other circulating antiangiogenic factors in preeclampsia. N Engl J Med 2006, 355 (10), 992–1005. [PubMed: 16957146]
46. Erez O; Romero R; Espinoza J; Fu W; Todem D; Kusanovic JP; Gotsch F; Edwin S; Nien JK; Chaiworapongsa T; Mittal P; Mazaki-Tovi S; Than NG; Gomez R; Hassan SS, The change in concentrations of angiogenic and anti-angiogenic factors in maternal plasma between the first and second trimesters in risk assessment for the subsequent development of preeclampsia and small-for-gestational age. J Matern Fetal Neonatal Med 2008, 21 (5), 279–87. [PubMed: 18446652]
47. Romero R; Nien JK; Espinoza J; Todem D; Fu W; Chung H; Kusanovic JP; Gotsch F; Erez O; Mazaki-Tovi S; Gomez R; Edwin S; Chaiworapongsa T; Levine RJ; Karumanchi SA, A longitudinal study of angiogenic (placental growth factor) and anti-angiogenic (soluble endoglin and soluble vascular endothelial growth factor receptor-1) factors in normal pregnancy and patients destined to develop preeclampsia and deliver a small for gestational age neonate. J Matern Fetal Neonatal Med 2008, 21 (1), 9–23. [PubMed: 18175241]
48. Kusanovic JP; Romero R; Chaiworapongsa T; Erez O; Mittal P; Vaisbuch E; Mazaki-Tovi S; Gotsch F; Edwin SS; Gomez R; Yeo L; Conde-Agudelo A; Hassan SS, A prospective cohort study of the value of maternal plasma concentrations of angiogenic and anti-angiogenic factors in early pregnancy and midtrimester in the identification of patients destined to develop preeclampsia. J Matern Fetal Neonatal Med 2009, 22 (11), 1021–38. [PubMed: 19900040]
49. Chaiworapongsa T; Romero R; Savasan ZA; Kusanovic JP; Ogge G; Soto E; Dong Z; Tarca A; Gaurav B; Hassan SS, Maternal plasma concentrations of angiogenic/anti-angiogenic factors are of prognostic value in patients presenting to the obstetrical triage area with the suspicion of preeclampsia. J Matern Fetal Neonatal Med 2011, 24 (10), 1187–207. [PubMed: 21827221]
50. Chaiworapongsa T; Romero R; Korzeniewski SJ; Kusanovic JP; Soto E; Lam J; Dong Z; Than NG; Yeo L; Hernandez-Andrade E; Conde-Agudelo A; Hassan SS, Maternal plasma concentrations of angiogenic/antiangiogenic factors in the third trimester of pregnancy to identify the patient at risk for stillbirth at or near term and severe late preeclampsia. Am J Obstet Gynecol 2013, 208 (4), 287 e1–287 e15.
51. Chaiworapongsa T; Chaemsaithong P; Korzeniewski SJ; Yeo L; Romero R, Pre-eclampsia part 2: prediction, prevention and management. Nat Rev Nephrol 2014, 10 (9), 531–40. [PubMed: 25003612]
52. Chaiworapongsa T; Chaemsaithong P; Yeo L; Romero R, Pre-eclampsia part 1: current understanding of its pathophysiology. Nat Rev Nephrol 2014, 10 (8), 466–80. [PubMed: 25003615]

53. Chaiworapongsa T; Romero R; Korzeniewski SJ; Cortez JM; Pappas A; Tarca AL; Chaemsaithong P; Dong Z; Yeo L; Hassan SS, Plasma concentrations of angiogenic/anti-angiogenic factors have prognostic value in women presenting with suspected preeclampsia to the obstetrical triage area: a prospective study. J Matern Fetal Neonatal Med 2014, 27 (2), 132–44. [PubMed: 23687930]
54. Tarca AL; Taran A; Romero R; Jung E; Paredes C; Bhatti G; Ghita C; Chaiworapongsa T; Than NG; Hsu CD, Prediction of preeclampsia throughout gestation with maternal characteristics and biophysical and biochemical markers: a longitudinal study. Am J Obstet Gynecol 2022, 226 (1), 126 e1–126 e22.
55. Espinoza J; Chaiworapongsa T; Romero R; Kim YM; Kim GJ; Nien JK; Kusanovic JP; Erez O; Bujold E; Goncalves LF; Gomez R; Edwin S, Unexplained fetal death: another anti-angiogenic state. J Matern Fetal Neonatal Med 2007, 20 (7), 495–507. [PubMed: 17674262]
56. Romero R; Chaiworapongsa T; Erez O; Tarca AL; Gervasi MT; Kusanovic JP; Mittal P; Ogge G; Vaisbuch E; Mazaki-Tovi S; Dong Z; Kim SK; Yeo L; Hassan SS, An imbalance between angiogenic and anti-angiogenic factors precedes fetal death in a subset of patients: results of a longitudinal study. J Matern Fetal Neonatal Med 2010, 23 (12), 1384–99. [PubMed: 20459337]
57. Chaiworapongsa T; Romero R; Whitten AE; Korzeniewski SJ; Chaemsaithong P; Hernandez-Andrade E; Yeo L; Hassan SS, The use of angiogenic biomarkers in maternal blood to identify which SGA fetuses will require a preterm delivery and mothers who will develop pre-eclampsia. J Matern Fetal Neonatal Med 2016, 29 (8), 1214–28. [PubMed: 26303962]
58. Whitten AE; Romero R; Korzeniewski SJ; Tarca AL; Schwartz AG; Yeo L; Dong Z; Hassan SS; Chaiworapongsa T, Evidence of an imbalance of angiogenic/antiangiogenic factors in massive perivillous fibrin deposition (maternal floor infarction): a placental lesion associated with recurrent miscarriage and fetal death. Am J Obstet Gynecol 2013, 208 (4), 310 e1–310 e11.
59. Chaiworapongsa T; Romero R; Korzeniewski SJ; Chaemsaithong P; Hernandez-Andrade E; Segars JH; DeCherney AH; McCoy MC; Kim CJ; Yeo L; Hassan SS, Pravastatin to prevent recurrent fetal death in massive perivillous fibrin deposition of the placenta (MPFD). J Matern Fetal Neonatal Med 2016, 29 (6), 855–62. [PubMed: 25893545]
60. Davies DR; Gelinas AD; Zhang C; Rohloff JC; Carter JD; O'Connell D; Waugh SM; Wolk SK; Mayfield WS; Burgin AB; Edwards TE; Stewart LJ; Gold L; Janjic N; Jarvis TC, Unique motifs and hydrophobic interactions shape the binding of modified DNA ligands to protein targets. Proc Natl Acad Sci U S A 2012, 109 (49), 19971–6. [PubMed: 23139410]
61. Gold L; Ayers D; Bertino J; Bock C; Bock A; Brody EN; Carter J; Dalby AB; Eaton BE; Fitzwater T; Flather D; Forbes A; Foreman T; Fowler C; Gawande B; Goss M; Gunn M; Gupta S; Halladay D; Heil J; Heilig J; Hicke B; Husar G; Janjic N; Jarvis T; Jennings S; Katilius E; Keeney TR; Kim N; Koch TH; Kraemer S; Kroiss L; Le N; Levine D; Lindsey W; Lollo B; Mayfield W; Mehan M; Mehler R; Nelson SK; Nelson M; Nieuwlandt D; Nikrad M; Ochsner U; Ostroff RM; Otis M; Parker T; Pietrasiewicz S; Resnicow DI; Rohloff J; Sanders G; Sattin S; Schneider D; Singer B; Stanton M; Sterkel A; Stewart A; Stratford S; Vaught JD; Vrkljan M; Walker JJ; Watrobka M; Waugh S; Weiss A; Wilcox SK; Wolfson A; Wolk SK; Zhang C; Zichi D, Aptamer-based multiplexed proteomic technology for biomarker discovery. PLoS One 2010, 5 (12), e15004. [PubMed: 21165148]
62. Bhatti G; Romero R; Gomez-Lopez N; Chaiworapongsa T; Jung E; Gotsch F; Pique-Regi R; Pacora P; Hsu CD; Kavdia M; Tarca AL, The amniotic fluid proteome changes with gestational age in normal pregnancy: a cross-sectional study. Sci Rep 2022, 12 (1), 601. [PubMed: 35022423]
63. Michaels JE; Dasari S; Pereira L; Reddy AP; Lapidus JA; Lu X; Jacob T; Thomas A; Rodland M; Roberts CT Jr.; Gravett MG; Nagalla SR, Comprehensive proteomic analysis of the human amniotic fluid proteome: gestational age-dependent changes. J Proteome Res 2007, 6 (4), 1277–85. [PubMed: 17373841]
64. Pereira L; Reddy AP; Jacob T; Thomas A; Schneider KA; Dasari S; Lapidus JA; Lu X; Rodland M; Roberts CT Jr.; Gravett MG; Nagalla SR, Identification of novel protein biomarkers of preterm birth in human cervical-vaginal fluid. J Proteome Res 2007, 6 (4), 1269–76. [PubMed: 17373840]
65. Omenn GS; Lane L; Overall CM; Paik YK; Cristea IM; Corrales FJ; Lindskog C; Weintraub S; Roehrl MHA; Liu S; Bandeira N; Srivastava S; Chen YJ; Aebersold R; Moritz RL; Deutsch EW, Progress Identifying and Analyzing the Human Proteome: 2021 Metrics from the HUPO Human Proteome Project. J Proteome Res 2021, 20 (12), 5227–5240. [PubMed: 34670092]

66. Korzeniewski SJ; Romero R; Chaiworapongsa T; Chaemsaithong P; Kim CJ; Kim YM; Kim JS; Yoon BH; Hassan SS; Yeo L, Maternal plasma angiogenic index-1 (placental growth factor/soluble vascular endothelial growth factor receptor-1) is a biomarker for the burden of placental lesions consistent with uteroplacental underperfusion: a longitudinal case-cohort study. Am J Obstet Gynecol 2016, 214 (5), 629 e1–629 e17.

67. SomaScan Assay v4.1 Technical Note https://somalogic.com/. https://somalogic.com/.

68. Ogge G; Romero R; Kusanovic JP; Chaiworapongsa T; Dong Z; Mittal P; Vaisbuch E; Mazaki-Tovi S; Gonzalez JM; Yeo L; Hassan SS, Serum and plasma determination of angiogenic and anti-angiogenic factors yield different results: the need for standardization in clinical practice. J Matern Fetal Neonatal Med 2010, 23 (8), 820–7. [PubMed: 20158394]

69. Bland JM; Altman DG, Measurement error proportional to the mean. BMJ 1996, 313 (7049), 106. [PubMed: 8688716]

70. Benjamini Y; Hochberg Y, Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Royal Stat Soc B 1995, 57 (1), 289–300.

71. Bates D; Mächler M; Bolker B; Walker S, Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 2015, 67 (1), 1–48.

72. Langfelder P; Horvath S, WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008, 9, 559. [PubMed: 19114008]

73. Maglott D; Ostell J; Pruitt KD; Tatusova T, Entrez Gene: gene-centered information at NCBI. Nucleic Acids Res 2005, 33 (Database issue), D54–8. [PubMed: 15608257]

74. Ashburner M; Ball CA; Blake JA; Botstein D; Butler H; Cherry JM; Davis AP; Dolinski K; Dwight SS; Eppig JT; Harris MA; Hill DP; Issel-Tarver L; Kasarskis A; Lewis S; Matese JC; Richardson JE; Ringwald M; Rubin GM; Sherlock G, Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000, 25 (1), 25–9. [PubMed: 10802651]

75. Falcon S; Gentleman R, Using GOstats to test gene lists for GO term association. Bioinformatics 2007, 23 (2), 257–8. [PubMed: 17098774]

76. Gentleman RC; Carey VJ; Bates DM; Bolstad B; Dettling M; Dudoit S; Ellis B; Gautier L; Ge Y; Gentry J; Hornik K; Hothorn T; Huber W; Iacus S; Irizarry R; Leisch F; Li C; Maechler M; Rossini AJ; Sawitzki G; Smith C; Smyth G; Tierney L; Yang JY; Zhang J, Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 2004, 5 (10), R80. [PubMed: 15461798]

77. Reimand J; Isserlin R; Voisin V; Kucera M; Tannus-Lopes C; Rostamianfar A; Wadi L; Meyer M; Wong J; Xu C; Merico D; Bader GD, Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. Nat Protoc 2019, 14 (2), 482–517. [PubMed: 30664679]

78. Wieder C; Frainay C; Poupin N; Rodriguez-Mier P; Vinson F; Cooke J; Lai RP; Bundy JG; Jourdan F; Ebbels T, Pathway analysis in metabolomics: Recommendations for the use of over-representation analysis. PLoS Comput Biol 2021, 17 (9), e1009105. [PubMed: 34492007]

79. Pique-Regi R; Romero R; Tarca AL; Sendler ED; Xu Y; Garcia-Flores V; Leng Y; Luca F; Hassan SS; Gomez-Lopez N, Single cell transcriptional signatures of the human placenta in term and preterm parturition. Elife 2019, 8.

80. Breiman L, Random Forests. Machine Learning 2001, 45 (1), 5–32.

81. Friedman J; Hastie T; Tibshirani R, Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw 2010, 33 (1), 1–22. [PubMed: 20808728]

82. O'Gorman N; Wright D; Syngelaki A; Akolekar R; Wright A; Poon LC; Nicolaides KH, Competing risks model in screening for preeclampsia by maternal factors and biomarkers at 11–13 weeks gestation. Am J Obstet Gynecol 2016, 214 (1), 103 e1–103 e12.

83. Romero R; Chaemsaithong P; Docheva N; Korzeniewski SJ; Tarca AL; Bhatti G; Xu Z; Kusanovic JP; Dong Z; Chaiyasit N; Ahmed AI; Yoon BH; Hassan SS; Chaiworapongsa T; Yeo L, Clinical chorioamnionitis at term IV: the maternal plasma cytokine profile. J Perinat Med 2016, 44 (1), 77–98. [PubMed: 26352068]

84. Szarka A; Rigo J Jr.; Lazar L; Beko G; Molvarec A, Circulating cytokines, chemokines and adhesion molecules in normal pregnancy and preeclampsia determined by multiplex suspension array. BMC Immunol 2010, 11, 59. [PubMed: 21126355]

85. Jarmund AH; Giskeodegard GF; Ryssdal M; Steinkjer B; Stokkeland LMT; Madssen TS; Stafne SN; Stridsklev S; Moholdt T; Heimstad R; Vanky E; Iversen AC, Cytokine Patterns in Maternal Serum From First Trimester to Term and Beyond. Front Immunol 2021, 12, 752660. [PubMed: 34721426]

86. Shainker SA; Silver RM; Modest AM; Hacker MR; Hecht JL; Salahuddin S; Dillon ST; Ciampa EJ; D'Alton ME; Otu HH; Abuhamad AZ; Einerson BD; Branch DW; Wylie BJ; Libermann TA; Karumanchi SA, Placenta accreta spectrum: biomarker discovery using plasma proteomics. Am J Obstet Gynecol 2020, 223 (3), 433 e1–433 e14.

87. Alexander GR; Kogan MD; Himes JH, 1994–1996 U.S. singleton birth weight percentiles for gestational age by race, Hispanic origin, and gender. Matern Child Health J 1999, 3 (4), 225–31. [PubMed: 10791363]

88. Tin A; Yu B; Ma J; Masushita K; Daya N; Hoogeveen RC; Ballantyne CM; Couper D; Rebholz CM; Grams ME; Alonso A; Mosley T; Heiss G; Ganz P; Selvin E; Boerwinkle E; Coresh J, Reproducibility and Variability of Protein Analytes Measured Using a Multiplexed Modified Aptamer Assay. J Appl Lab Med 2019, 4 (1), 30–39. [PubMed: 31639705]

89. Rolnik DL; Nicolaides KH; Poon LC, Prevention of preeclampsia with aspirin. Am J Obstet Gynecol 2022, 226 (2S), S1108–S1119. [PubMed: 32835720]

90. Wright D; Wright A; Nicolaides KH, The competing risk approach for prediction of preeclampsia. Am J Obstet Gynecol 2020, 223 (1), 12–23 e7. [PubMed: 31733203]

91. Gomez-Lopez N; Romero R; Hassan SS; Bhatti G; Berry SM; Kusanovic JP; Pacora P; Tarca AL, The Cellular Transcriptome in the Maternal Circulation During Normal Pregnancy: A Longitudinal Study. Front Immunol 2019, 10, 2863. [PubMed: 31921132]

92. Gilgès D; Vinit MA; Callebaut I; Coulombel L; Cacheux V; Romeo PH; Vigon I, Polydom: a secreted protein with pentraxin, complement control protein, epidermal growth factor and von Willebrand factor A domains. Biochem J 2000, 352 Pt 1 (Pt 1), 49–59. [PubMed: 11062057]

93. Karpanen T; Padberg Y; van de Pavert SA; Dierkes C; Morooka N; Peterson-Maduro J; van de Hoek G; Adrian M; Mochizuki N; Sekiguchi K; Kiefer F; Schulte D; Schulte-Merker S, An Evolutionarily Conserved Role for Polydom/Svep1 During Lymphatic Vessel Formation. Circ Res 2017, 120 (8), 1263–1275. [PubMed: 28179432]

94. Morooka N; Futaki S; Sato-Nishiuchi R; Nishino M; Totani Y; Shimono C; Nakano I; Nakajima H; Mochizuki N; Sekiguchi K, Polydom Is an Extracellular Matrix Protein Involved in Lymphatic Vessel Remodeling. Circ Res 2017, 120 (8), 1276–1288. [PubMed: 28179430]

95. Siddiqui S; Gurung RL; Liu S; Ping Seet EC; Lim SC, Genetic Polymorphisms and Cytokine Profile of Different Ethnicities inSeptic Shock Patients and their Association with Mortality. Indian J Crit Care Med 2019, 23 (3), 135–138. [PubMed: 31097890]

96. Nakada TA; Russell JA; Boyd JH; Thair SA; Walley KR, Identification of a nonsynonymous polymorphism in the SVEP1 gene associated with altered clinical outcomes in septic shock. Crit Care Med 2015, 43 (1), 101–8. [PubMed: 25188548]

97. Schwanzer-Pfeiffer D; Rossmanith E; Schildberger A; Falkenhagen D, Characterization of SVEP1, KIAA, and SRPX2 in an in vitro cell culture model of endotoxemia. Cell Immunol 2010, 263 (1), 65–70. [PubMed: 20236627]

98. Jung IH; Elenbaas JS; Alisio A; Santana K; Young EP; Kang CJ; Kachroo P; Lavine KJ; Razani B; Mecham RP; Stitziel NO, SVEP1 is a human coronary artery disease locus that promotes atherosclerosis. Sci Transl Med 2021, 13 (586), eabe0357. [PubMed: 33762433]

99. Stitziel NO; Stirrups KE; Masca NG; Erdmann J; Ferrario PG; König IR; Weeke PE; Webb TR; Auer PL; Schick UM; Lu Y; Zhang H; Dube MP; Goel A; Farrall M; Peloso GM; Won HH; Do R; van Iperen E; Kanoni S; Kruppa J; Mahajan A; Scott RA; Willenberg C; Braund PS; van Capelleveen JC; Doney AS; Donnelly LA; Asselta R; Merlini PA; Duga S; Marziliano N; Denny JC; Shaffer CM; El-Mokhtari NE; Franke A; Gottesman O; Heilmann S; Hengstenberg C; Hoffman P; Holmen OL; Hveem K; Jansson JH; Jöckel KH; Kessler T; Kriebel J; Laugwitz KL; Marouli E; Martinelli N; McCarthy MI; Van Zuydam NR; Meisinger C; Esko T; Mihailov E; Escher SA; Alver M; Moebus S; Morris AD; Müller-Nurasyid M; Nikpay M; Olivieri O; Lemieux Perreault LP; AlQarawi A; Robertson NR; Akinsanya KO; Reilly DF; Vogt TF; Yin W; Asselbergs FW; Kooperberg C; Jackson RD; Stahl E; Strauch K; Varga TV; Waldenberger M; Zeng L; Kraja AT; Liu C; Ehret GB; Newton-Cheh C; Chasman DI; Chowdhury R; Ferrario M; Ford I; Jukema

JW; Kee F; Kuulasmaa K; Nordestgaard BG; Perola M; Saleheen D; Sattar N; Surendran P; Tregouet D; Young R; Howson JM; Butterworth AS; Danesh J; Ardissino D; Bottinger EP; Erbel R; Franks PW; Girelli D; Hall AS; Hovingh GK; Kastrati A; Lieb W; Meitinger T; Kraus WE; Shah SH; McPherson R; Orho-Melander M; Melander O; Metspalu A; Palmer CN; Peters A; Rader D; Reilly MP; Loos RJ; Reiner AP; Roden DM; Tardif JC; Thompson JR; Wareham NJ; Watkins H; Willer CJ; Kathiresan S; Deloukas P; Samani NJ; Schunkert H, Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. N Engl J Med 2016, 374 (12), 1134–44. [PubMed: 26934567]

100. Winkler MJ; Müller P; Sharifi AM; Wobst J; Winter H; Mokry M; Ma L; van der Laan SW; Pang S; Miritsch B; Hinterdobler J; Werner J; Stiller B; Güldener U; Webb TR; Asselbergs FW; Björkegren JLM; Maegdefessel L; Schunkert H; Sager HB; Kessler T, Functional investigation of the coronary artery disease gene SVEP1. Basic Res Cardiol 2020, 115 (6), 67. [PubMed: 33185739]
101. Liang XH; Zhao ZA; Deng WB; Tian Z; Lei W; Xu X; Zhang XH; Su RW; Yang ZM, Estrogen regulates amiloride-binding protein 1 through CCAAT/enhancer-binding protein-beta in mouse uterus during embryo implantation and decidualization. Endocrinology 2010, 151 (10), 5007–16. [PubMed: 20668027]
102. Metsalu T; Viltrop T; Tiirats A; Rajashekar B; Reimann E; Kõks S; Rull K; Milani L; Acharya G; Basnet P; Vilo J; Mägi R; Metspalu A; Peters M; Haller-Kikkatalo K; Salumets A, Using RNA sequencing for identifying gene imprinting and random monoallelic expression in human placenta. Epigenetics 2014, 9 (10), 1397–409. [PubMed: 25437054]
103. Southren AL; Weingold AB, Plasma diamine oxidase in pregnancy complicated by age and weight factors. Am J Obstet Gynecol 1970, 106 (4), 607–16. [PubMed: 4983921]
104. Maintz L; Schwarzer V; Bieber T; van der Ven K; Novak N, Effects of histamine and diamine oxidase activities on pregnancy: a critical review. Hum Reprod Update 2008, 14 (5), 485–95. [PubMed: 18499706]
105. Goyal R; Yellon SM; Longo LD; Mata-Greenwood E, Placental gene expression in a rat 'model' of placental insufficiency. Placenta 2010, 31 (7), 568–75. [PubMed: 20621762]
106. Weyer K; Glerup S, Placental regulation of peptide hormone and growth factor activity by proMBP. Biol Reprod 2011, 84 (6), 1077–86. [PubMed: 21270431]
107. Vernof KK; Benirschke K; Kephart GM; Wasmoen TL; Gleich GJ, Maternal floor infarction: relationship to X cells, major basic protein, and adverse perinatal outcome. Am J Obstet Gynecol 1992, 167 (5), 1355–63. [PubMed: 1442990]
108. Coulam CB; Wasmoen T; Creasy R; Siiteri P; Gleich G, Major basic protein as a predictor of preterm labor: a preliminary report. Am J Obstet Gynecol 1987, 156 (4), 790–6. [PubMed: 3578393]
109. MacDonald TM; Walker SP; Hiscock R; Cannon P; Harper A; Murray E; Hui L; Dane K; Middleton A; Kyritsis V; de Alwis N; Hannan NJ; Tong S; Kaitu'u-Lino TJ, Circulating Delta-like homolog 1 (DLK1) at 36 weeks is correlated with birthweight and is of placental origin. Placenta 2020, 91, 24–30. [PubMed: 32174303]
110. Jensen CH; Krogh TN; Højrup P; Clausen PP; Skjødt K; Larsson LI; Enghild JJ; Teisner B, Protein structure of fetal antigen 1 (FA1). A novel circulating human epidermal-growth-factor-like protein expressed in neuroendocrine tumors and its relation to the gene products of dlk and pG2. Eur J Biochem 1994, 225 (1), 83–92. [PubMed: 7925474]
111. Cleaton MA; Dent CL; Howard M; Corish JA; Gutteridge I; Sovio U; Gaccioli F; Takahashi N; Bauer SR; Charnock-Jones DS; Powell TL; Smith GC; Ferguson-Smith AC; Charalambous M, Fetus-derived DLK1 is required for maternal metabolic adaptations to pregnancy and is associated with fetal growth restriction. Nat Genet 2016, 48 (12), 1473–1480. [PubMed: 27776119]
112. Díaz M; Bassols J; Aragonés G; Mazarico E; López-Bermejo A; Ibáñez L, Decreased placental expression of pre-adipocyte factor-1 in children born small-for-gestational-age: association to early postnatal weight gain. Placenta 2013, 34 (4), 331–4. [PubMed: 23410724]
113. Tarca AL; Romero R; Pique-Regi R; Pacora P; Done B; Kacerovsky M; Bhatti G; Jaiman S; Hassan SS; Hsu CD; Gomez-Lopez N, Amniotic fluid cell-free transcriptome: a glimpse

into fetal development and placental cellular dynamics during normal pregnancy. BMC Med Genomics 2020, 13 (1), 25. [PubMed: 32050959]

114. Tarca AL; Romero R; Erez O; Gudicha DW; Than NG; Benshalom-Tirosh N; Pacora P; Hsu CD; Chaiworapongsa T; Hassan SS; Gomez-Lopez N, Maternal whole blood mRNA signatures identify women at risk of early preeclampsia: a longitudinal study. J Matern Fetal Neonatal Med 2021, 34 (21), 3463–3474. [PubMed: 31900005]
115. Savitz DA; Terry JW Jr.; Dole N; Thorp JM Jr.; Siega-Riz AM; Herring AH, Comparison of pregnancy dating by last menstrual period, ultrasound scanning, and their combination. Am J Obstet Gynecol 2002, 187 (6), 1660–6. [PubMed: 12501080]
116. Koh W; Pan W; Gawad C; Fan HC; Kerchner GA; Wyss-Coray T; Blumenfeld YJ; El-Sayed YY; Quake SR, Noninvasive in vivo monitoring of tissue-specific global gene expression in humans. Proc Natl Acad Sci U S A 2014, 111 (20), 7361–6. [PubMed: 24799715]
117. Para R; Romero R; Miller D; Galaz J; Done B; Peyvandipour A; Gershater M; Tao L; Motomura K; Ruden DM; Isherwood J; Jung E; Kanninen T; Pique-Regi R; Tarca AL; Gomez-Lopez N, The Distinct Immune Nature of the Fetal Inflammatory Response Syndrome Type I and Type II. Immunohorizons 2021, 5 (9), 735–751. [PubMed: 34521696]
118. Bhatti G; Romero R; Gomez-Lopez N; Pique-Regi R; Pacora P; Jung E; Yeo L; Hsu CD; Kavdia M; Tarca AL, The amniotic fluid cell-free transcriptome in spontaneous preterm labor. Sci Rep 2021, 11 (1), 13481. [PubMed: 34188072]
119. Schanton M; Maymo JL; Perez-Perez A; Sanchez-Margalet V; Varone CL, Involvement of leptin in the molecular physiology of the placenta. Reproduction 2018, 155 (1), R1–R12. [PubMed: 29018059]
120. Perez-Perez A; Vilarino-Garcia T; Guadix P; Duenas JL; Sanchez-Margalet V, Leptin and Nutrition in Gestational Diabetes. Nutrients 2020, 12 (7).
121. Steinbrekera B; Colaizy TT; Vasilakos LK; Johnson KJ; Santillan DA; Haskell SE; Roghair RD, Origins of neonatal leptin deficiency in preterm infants. Pediatr Res 2019, 85 (7), 1016–1023. [PubMed: 30845123]
122. Tsiakkas A; Saiid Y; Wright A; Wright D; Nicolaides KH, Competing risks model in screening for preeclampsia by maternal factors and biomarkers at 30–34 weeks' gestation. Am J Obstet Gynecol 2016, 215 (1), 87 e1–87 e17.
123. Tan MY; Syngelaki A; Poon LC; Rolnik DL; O'Gorman N; Delgado JL; Akolekar R; Konstantinidou L; Tsavdaridou M; Galeva S; Ajdacka U; Molina FS; Persico N; Jani JC; Plasencia W; Greco E; Papaioannou G; Wright A; Wright D; Nicolaides KH, Screening for pre-eclampsia by maternal factors and biomarkers at 11–13 weeks' gestation. Ultrasound Obstet Gynecol 2018, 52 (2), 186–195. [PubMed: 29896812]
124. Gallo DM; Wright D; Casanova C; Campanero M; Nicolaides KH, Competing risks model in screening for preeclampsia by maternal factors and biomarkers at 19–24 weeks' gestation. Am J Obstet Gynecol 2016, 214 (5), 619 e1–619 e17.
125. Andrietti S; Silva M; Wright A; Wright D; Nicolaides KH, Competing-risks model in screening for pre-eclampsia by maternal factors and biomarkers at 35–37 weeks' gestation. Ultrasound Obstet Gynecol 2016, 48 (1), 72–9. [PubMed: 26566592]
126. Wright D; Syngelaki A; Akolekar R; Poon LC; Nicolaides KH, Competing risks model in screening for preeclampsia by maternal characteristics and medical history. Am J Obstet Gynecol 2015, 213 (1), 62 e1–62 e10.
127. Gudicha DW; Romero R; Kabiri D; Hernandez-Andrade E; Pacora P; Erez O; Kusanovic JP; Jung E; Paredes C; Berry SM; Yeo L; Hassan SS; Hsu CD; Tarca AL, Personalized assessment of cervical length improves prediction of spontaneous preterm birth: a standard and a percentile calculator. Am J Obstet Gynecol 2021, 224 (3), 288 e1–288 e17.

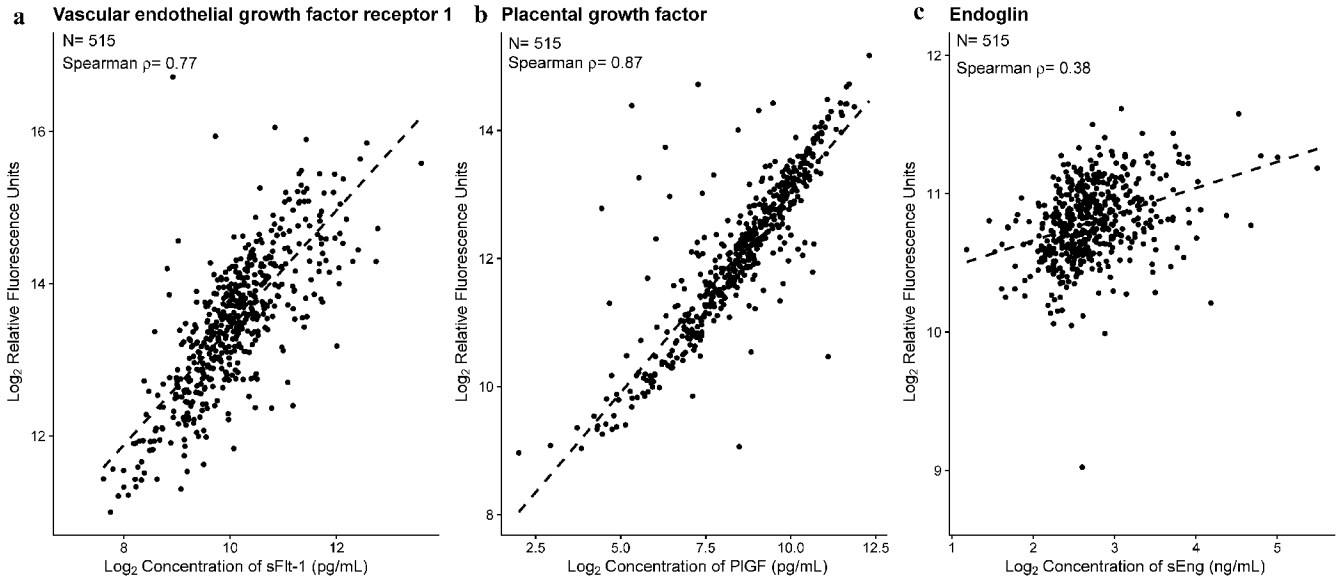


Figure 1: Agreement between ELISA and SomaScan® measurements for key angiogenic and anti-angiogenic proteins.

Protein abundance for 515 samples are shown, with one dot for each sample. The SomaScan relative fluorescence units in (log, base 2) (y-axis) is shown vs. ELISA based concentrations (log, base 2) (x-axis) for sFlt-1 (A), PlGF (B) and sEng (C). ρ is the Spearman's correlation coefficient. Correlation test $p<0.001$ for all three proteins.

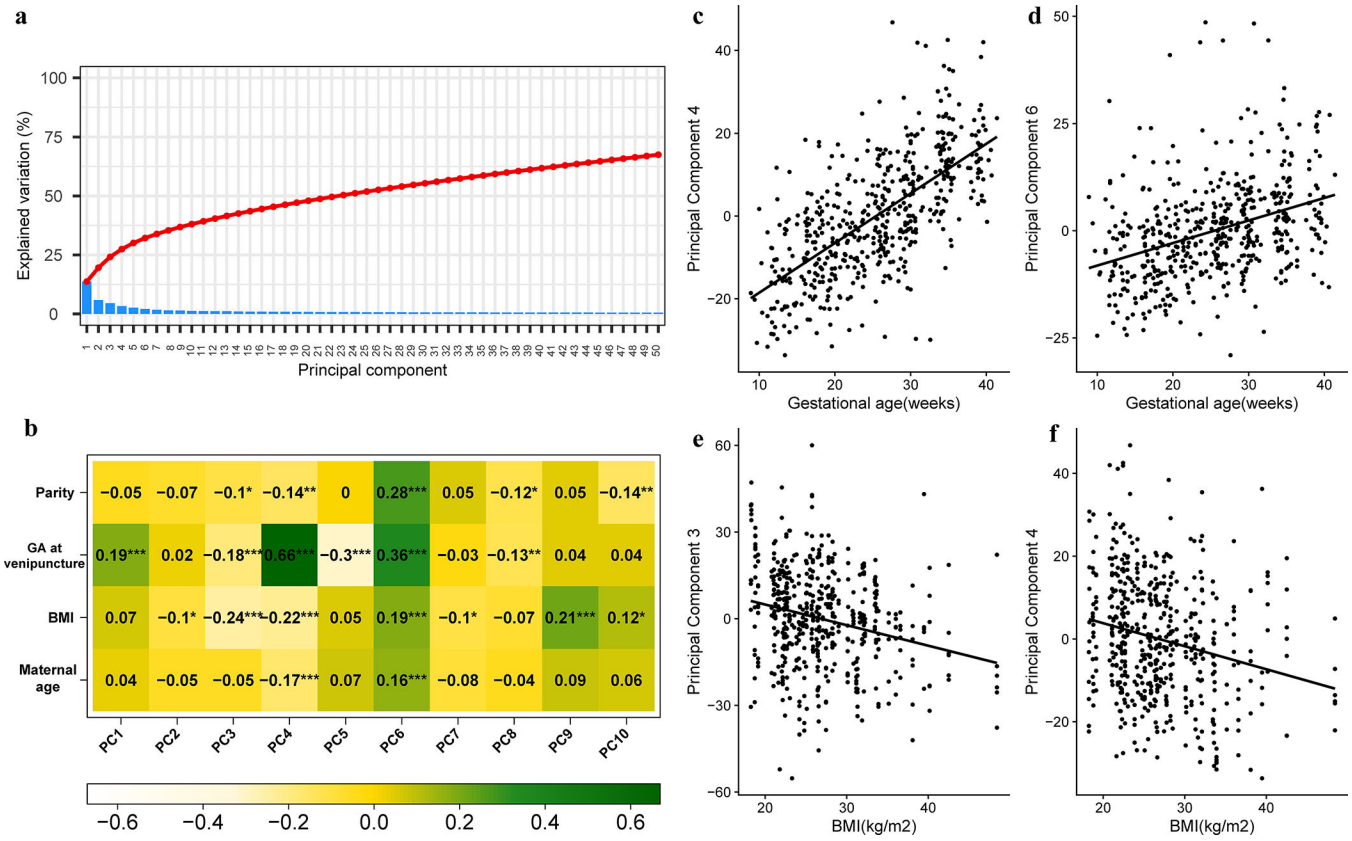


Figure 2: Principal component analysis of 6,277 proteins and correlation with gestational age and maternal characteristics.

Protein abundance in relative fluorescence units (log, base 2) was analyzed using principal components (PC) analysis. The resulting principal components can be understood as meta-proteins. A) The % of variance explained by each principal component is shown as a scree plot. (B) The top 10 principal components were correlated with maternal age, parity, gestational age (GA) and body mass index (BMI). The heatmap shows Pearson correlation coefficients between PC and covariates (significance levels: *<0.05, **<0.01, ***<0.001). The correlation between PC4 (C) and PC6 (D) with gestational age is also shown, with each dot representing one sample. Similar correlations are shown for PC3 (E) and PC4 (F) in relation with BMI.

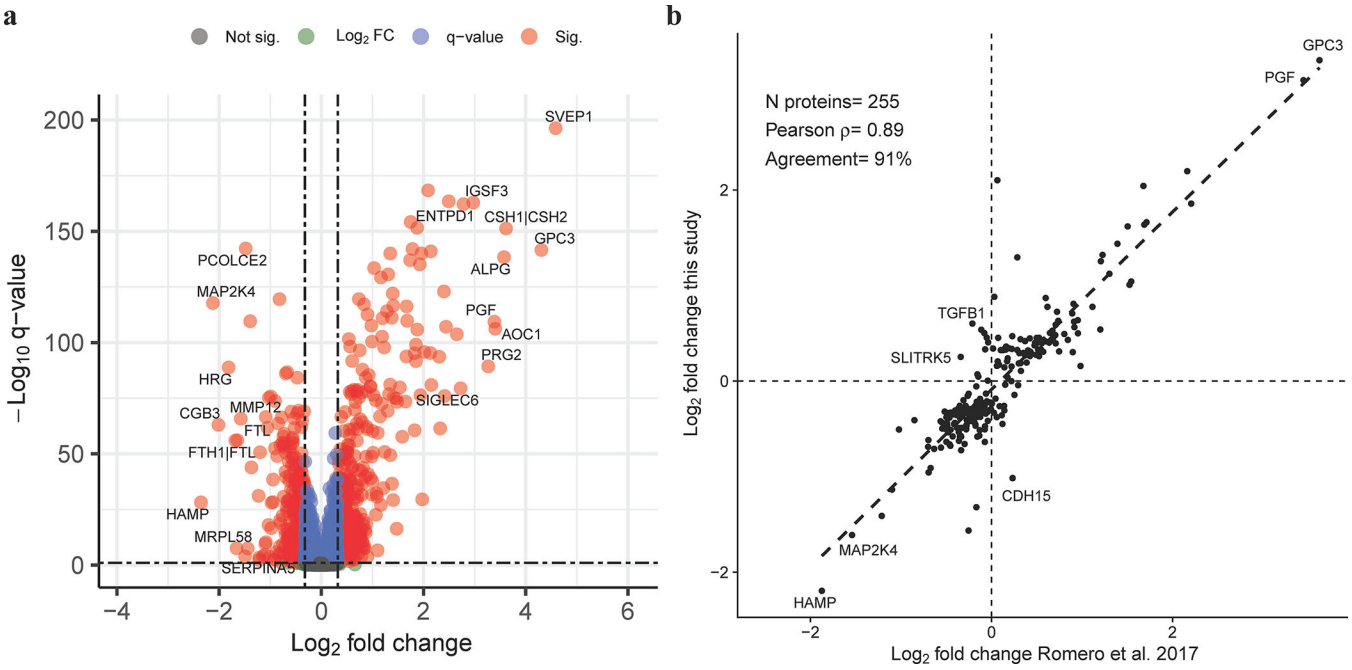


Figure 3: Proteomic changes with advancing gestation in normal pregnancy.

The volcano plot (A) shows the significance (y-axis) vs. magnitude of change (x-axis) for each protein. Protein with significant modulation (adjusted p-value, q <0.1 and fold change >1.25, N=953 proteins) are shown in red. The names of a select set of most significant proteins are also displayed. The correlation between fold changes (log, base 2) from 10 weeks to 32 weeks of gestation and similar results based on Romero et al. 2017 study is shown for 255 proteins deemed significant in this study and profiled in Romero et al. 2017 using SomaScan® platform v3. The gene symbols of the top increased and decreased proteins with concordant direction of change between studies are listed in the figure. The same is true for top three proteins with the most discordant fold change between studies.

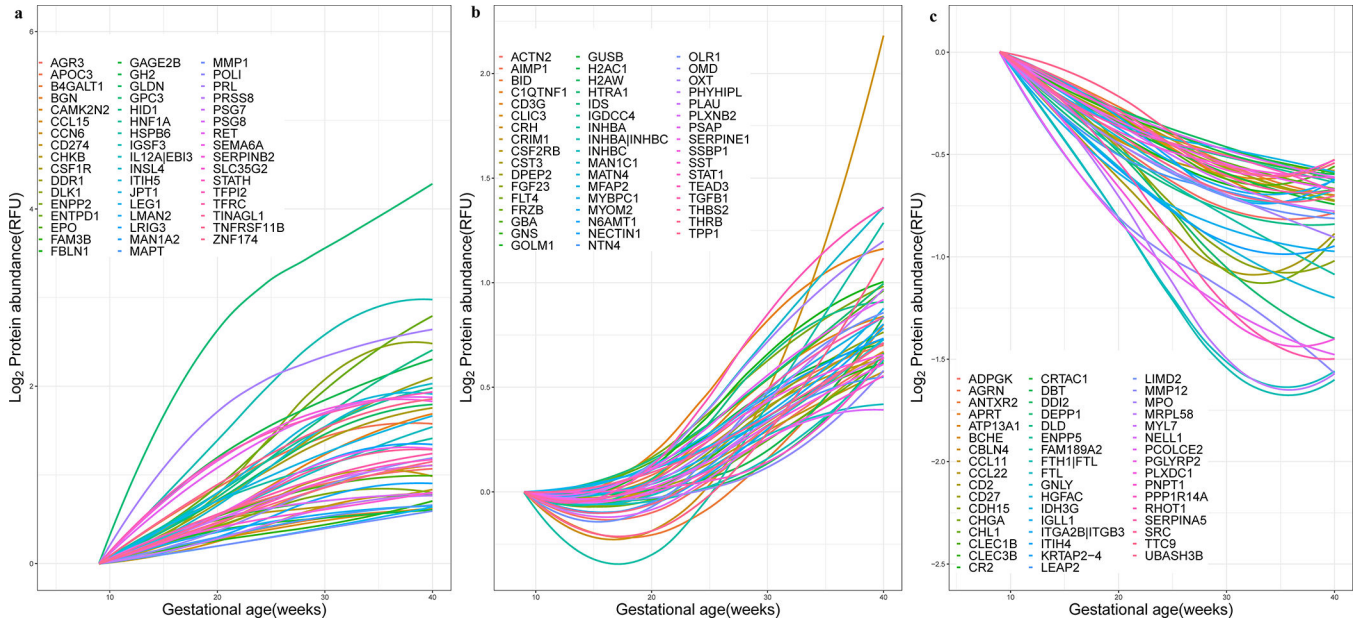


Figure 4: Maternal plasma proteomic trajectories throughout gestation.

The figure shows three clusters of proteins with steady increase (A), slow increase or decrease early in pregnancy followed by an increase later in pregnancy (B), and overall decreasing trend (C). The 50 proteins most representative of each cluster are shown, with names displayed representing corresponding gene symbols. See Supplementary File 2 for a depiction of the raw data used to derive the protein trajectories for selected proteins.

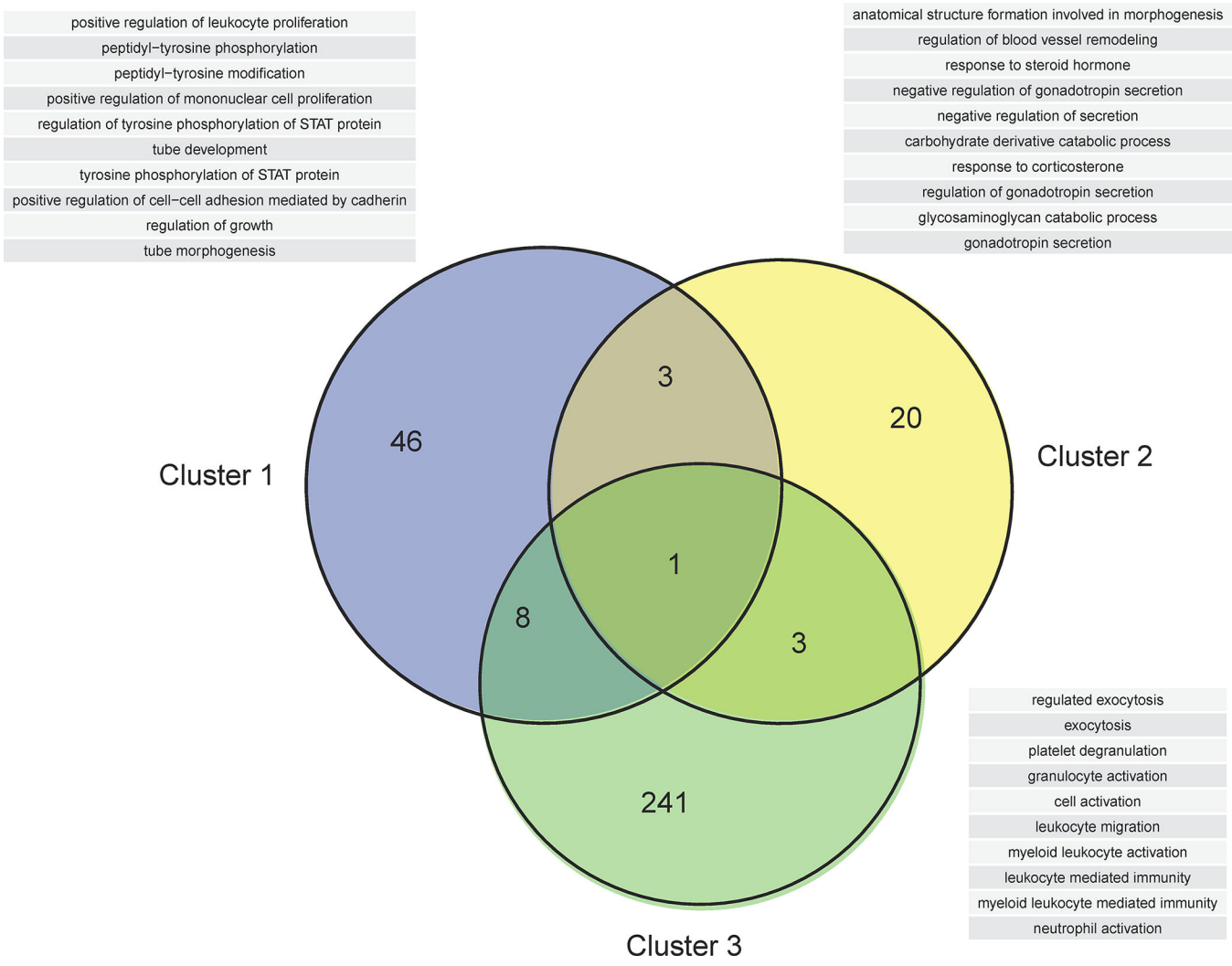


Figure 5: Biological processes associated with maternal plasma protein modulation with gestational age.

The Venn diagram shows the overlap in biological processes significantly associated with differentially modulated proteins for each cluster illustrated in Figure 4. See also Supplementary File 3 for the full list of biological processes associated with gestational age modulation. The list of the top 10 biological processes (ranked by enrichment p-value) for each Venn diagram category is shown in a table.

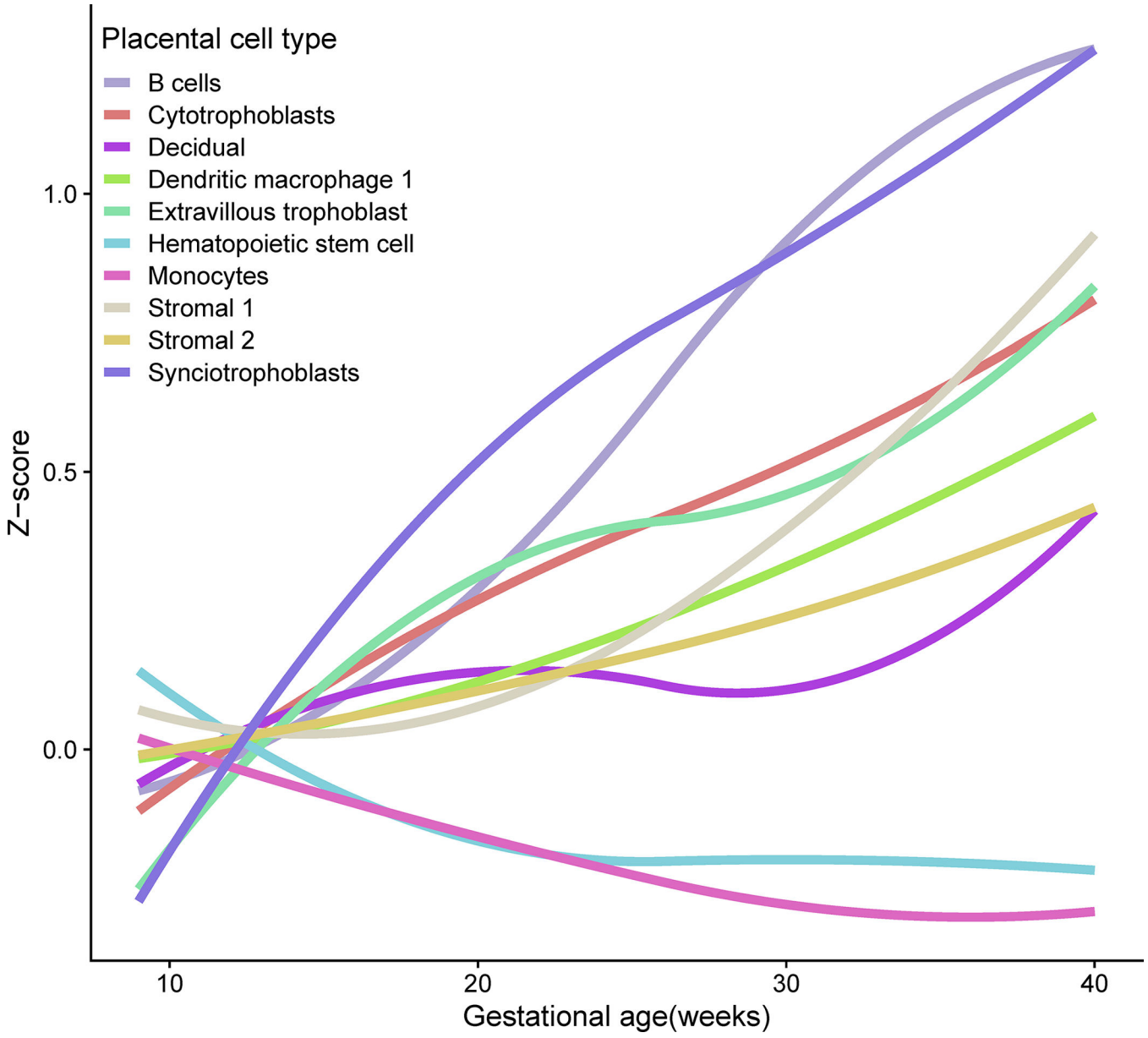


Figure 6: Maternal plasma proteomic trajectories of single-cell signatures throughout gestation.

For each placental single-cell signature the average Z-score of member proteins is shown as a function of gestational age. The gene symbols corresponding to each signature are: *Synciotrophoblasts* (KISS1, CSH1, TFPI2, CGA, GH2, PSG3, PSG2, PSG1, HOPX, CRH, GDF15, S100P, PSG11), *B cells* (CD79A, CD74, RPS5), *Extravillous trophoblast* (AOC1, PRG2, IGF2, NOTUM, FSTL3, FLT1, EBI3, PAPPA2, HPGD, HLA-G, PAPPA, ITM2B, KRT19, SERPINE2, MFAP5, HEXB, QSOX1, TPM1, TNFSF10), *Stromal 1* (TIMP1, DLK1, COL3A1, COL1A1, TGFBI, IL1RL1, COL6A2, IGFBP3, DCN, COMP, SERPINE2, COL6A1), *Stromal 2* (CXCL14, EGFL6, PTGDS, APOD, TCF21, DLK1, IGFBP3, COL3A1, PLA2G2A, COL1A1, C7, GPC3, LUM, CTHRC1, SERPINF1, RARRES2), *hematopoietic stem cell* (SPARCL1, ENPP2, EDN1, IGFBP7, CRIP2, A2M, SOCS3, ID1), *Monocytes* (S100A8, LYZ, S100A9, IL1B, S100A12, CXCL2,

BCL2A1, CCL3, CCL20, CXCL3, G0S2, PLAUR, FCN1, SOD2, C15orf48, EREG, IL1RN), *Decidual* (IGFBP1, LUM, DKK1, IGFBP2, DCN, RBP1, IGFBP4, PRL, IGFBP5, HSD11B1, IGFBP6, CD248, TIMP3, CFD), *Dendritic/Macrophage 1* (APOE, APOC1, CCL18, CD74, SPP1, C1QC, FTL, RNASE1, CXCL3, CTSZ), *Cytotrophoblasts* (PAGE4, CGA, TINAGL1, SPINT1, SPINT2, LDHB).

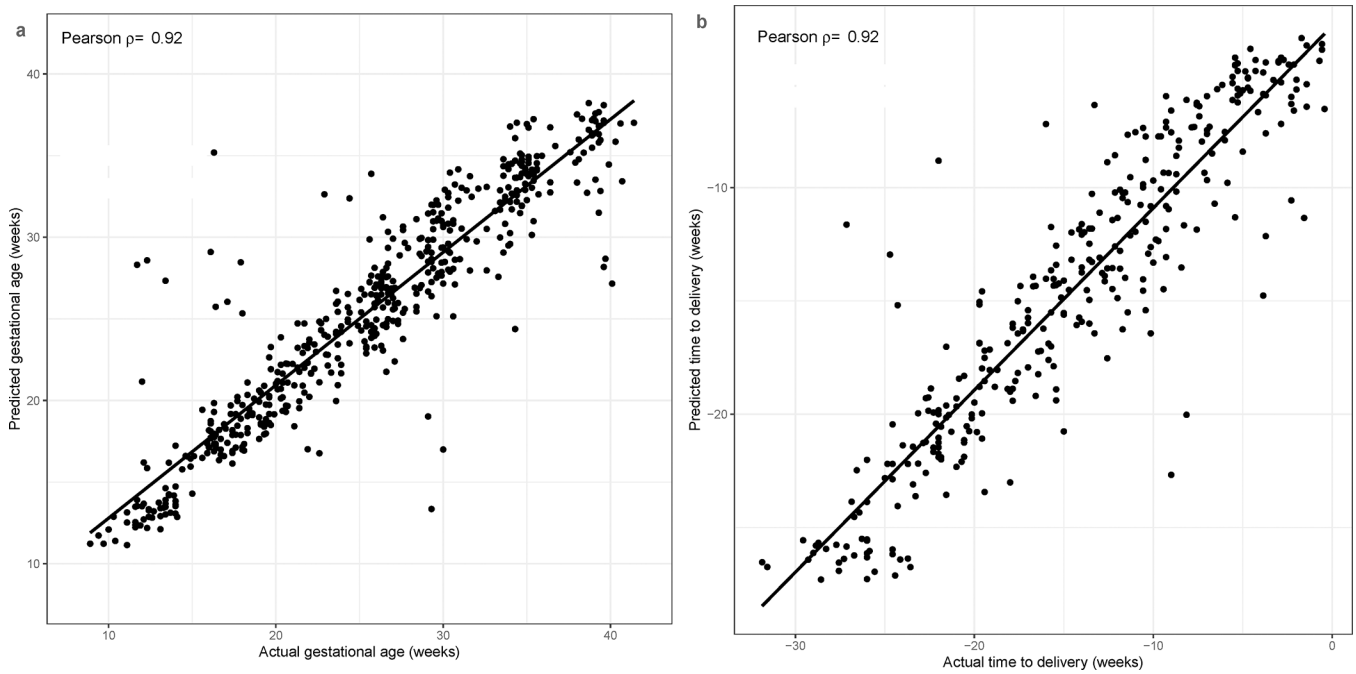
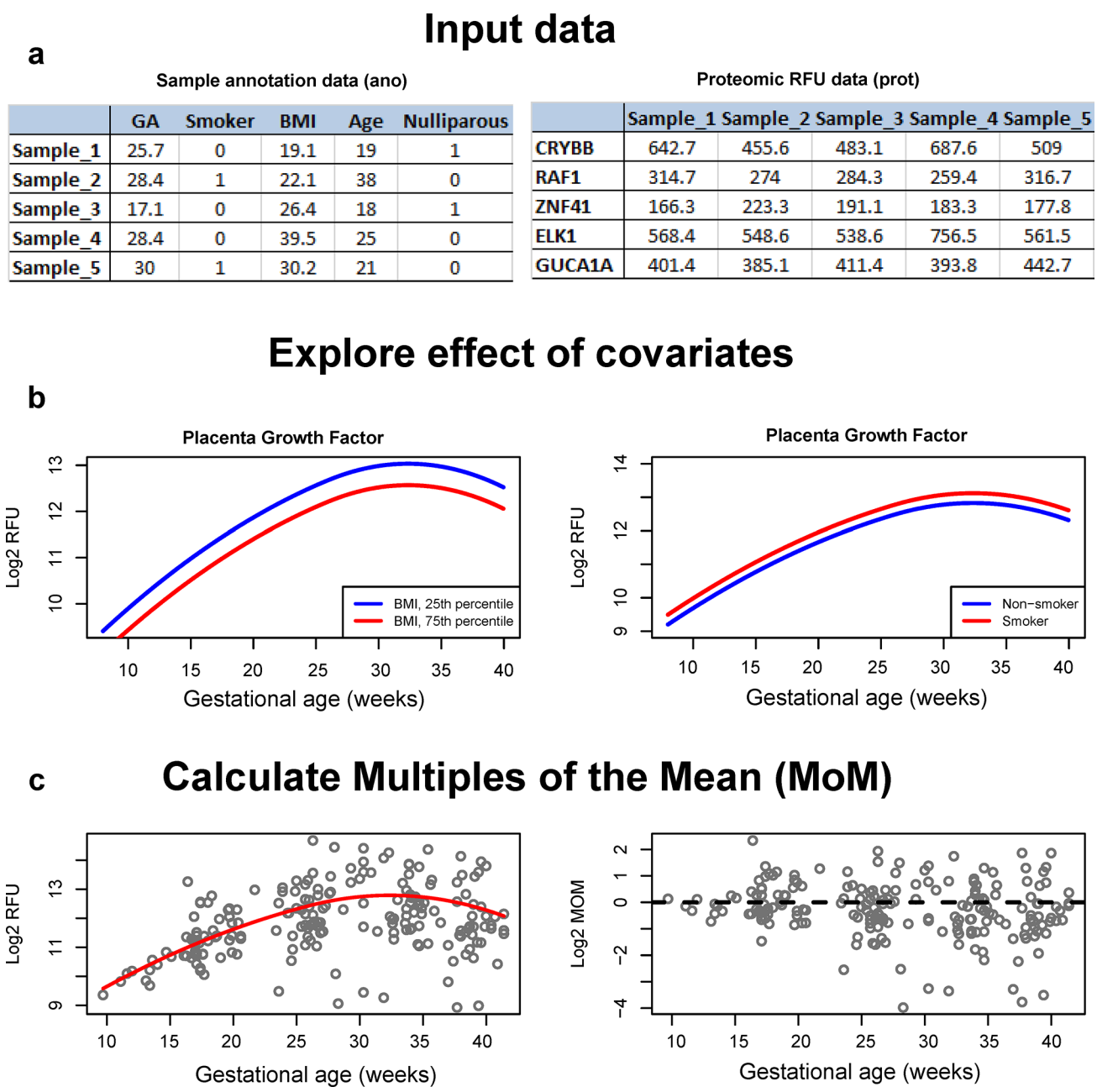


Figure 7: Prediction of gestational age and of time from sample to delivery using proteomic data.

Prediction of gestational age (A) and of time from sample to a spontaneous term delivery (B) is shown. Each dot corresponds to a sample. Random forest predictions are obtained via cross-validation, in which, data from all samples of a given patient are left out when selecting predictor proteins and training the model. ρ: correlation coefficient.



	GA	Smoker	BMI	Age	Nulliparous
Sample_1	25.7	0	19.1	19	1
Sample_2	28.4	1	22.1	38	0
Sample_3	17.1	0	26.4	18	1
Sample_4	28.4	0	39.5	25	0
Sample_5	30	1	30.2	21	0

	Sample_1	Sample_2	Sample_3	Sample_4	Sample_5
CRYBB	642.7	455.6	483.1	687.6	509
RAF1	314.7	274	284.3	259.4	316.7
ZNF41	166.3	223.3	191.1	183.3	177.8
ELK1	568.4	548.6	538.6	756.5	561.5
GUCA1A	401.4	385.1	411.4	393.8	442.7

Figure 8: Functionality implemented in the SomaPreg package.

Sample annotation data paired with proteomic data (RFU) (5 proteins of 7288 are shown) are used as input (A) to determine the expected proteomic abundance (B) and to calculate multiple of the mean values (MoM) for gestational age and maternal characteristics (C).