Early prediction of preeclampsia in pregnancy with circulating, cell-free RNA

Mira N. Moufarrej[1], Ronald J. Wong[2], Ana A. Campos[2], Cecele C. Quaintance[2], Rene V. Sit[3], Michelle Tan[3], Norma F. Neff[3], Maurice L. Druzin[4], Virginia D. Winn[4], Gary M. Shaw[2], David K. Stevenson[2], Stephen R. Quake[1,3,5,*]

[1] Department of Bioengineering, Stanford University, Stanford, CA USA 94305
[2] Department of Pediatrics, Stanford University School of Medicine, Stanford, CA USA 94305
[3] Chan Zuckerberg Biohub, Stanford, CA USA 94305
[4] Department of Obstetrics and Gynecology, Stanford University School of Medicine, Stanford, CA USA 94305
[5] Department of Applied Physics, Stanford University, Stanford, CA USA 94305

[*] Corresponding author: steve@quake-lab.org

Abstract

Liquid biopsies that measure circulating, cell-free RNA (cfRNA) offer an unprecedented opportunity to noninvasively study the development of pregnancy-related complications and to bridge gaps in clinical care. Here, we identify cfRNA transcriptomic changes across gestation and at post-partum that are associated with preeclampsia (PE), a multi-organ syndrome diagnosed after 20 weeks of gestation, using 118 samples from 42 pregnant mothers (18 normotensive (NT), 24 with PE). We find that changes in cfRNA gene expression between NT and PE mothers are most striking and stable early in gestation, suggesting that the identified cfRNA signals may correlate with PE pathogenesis. These changes also reflect our understanding of the known biology of PE, as supported by gene ontology analysis. Finally, we identify and independently validate (8 NT, 8 with PE/gestational hypertension) 11 genes, which when measured between 5–16 weeks of gestation can form a liquid biopsy test with clinically relevant specificity (88% [55–99%]) and sensitivity (100% [74–100%]) (All reported as value, [95% confidence interval]). Taken together, these results show that cfRNA measurements can form the basis of a test that would predict PE early in pregnancy, which has been an important and until now unrealized objective for obstetric care, and can help characterize the pathogenesis of PE in real time.

Introduction

Advances in obstetrics and neonatology have significantly mitigated many of the adverse pregnancy outcomes related to preterm birth (PTB) and preeclampsia (PE) [1,2]. Nonetheless, the standards of care implemented today focus on how to treat a mother and child once a complication has been diagnosed, proving both insufficient and costly [3–6]: PE and related hypertensive disorders cause 14% of maternal deaths each year globally, second only to hemorrhage [7], and cost $2B in care in the first year following delivery [5]. Worse, 3 out of 5 maternal deaths in the US are preventable and often associated with a missed or delayed complication diagnosis [8]. Such outcomes highlight the need for novel tools that would aid in identifying which women are at risk for hypertensive diseases, such as PE, before clinical presentation. Indeed, early prediction of PE, which has not been achieved to date, may enable prevention or reduction of a pregnant mother's risk of developing PE [9,10] if coupled with appropriate treatment.

PE is a leading pregnancy-related complication that globally affects 4–5% of pregnancies [11–13] and is associated with a significant increase in adverse maternal (e.g., maternal death, heart attack, stroke, seizures, and hemorrhage) and perinatal (e.g., fetal growth restriction and PTB coupled with respiratory distress syndrome, intraventricular hemorrhage, cerebral palsy, and bronchopulmonary dysplasia) outcomes [14–18]. Long-term, PE presents an increased maternal risk for cardiovascular [19,20] and kidney [21,22] diseases. Formally defined as new-onset hypertension with proteinuria or other end-organ damage (e.g., liver, brain) occurring after 20 weeks of gestation [23], PE can clinically manifest anytime thereafter, including into the post-partum period [24]. Detection and diagnosis itself however can prove challenging as early signs like headaches and nausea can be easily confused with general pregnancy discomfort and because PE shares many signs and symptoms with other common complications like gestational thrombocytopenia and chronic hypertension.

To date, no recommended test exists that can predict the future onset of PE early in pregnancy [9], and investigational approaches that measure diverse biophysical and biochemical signals including the measurement of two angiogenic factors (soluble fms-like tyrosine kinase-1 (sFlt1), placental growth factor (PlGF)) in the second and third trimesters [25,26] have at best only moderate positive predictive values (8–33%) [27]. A test with good performance metrics could guide the prophylactic use of low-dose aspirin, which has been shown to reduce the risk of PE if initially administered before 16 weeks of gestation [10,28]. Liquid biopsies that measure plasma cell-free RNA (cfRNA) suggest a means to bridge this gap in clinical care; however, until recently, such work often failed to progress beyond initial discovery [29]. Recent efforts have instead either focused on confirmation of PE at clinical diagnosis [30] or on limited discovery (n = 8 with PE/gestational hypertension) earlier in pregnancy with encouraging but unvalidated results [31]. Consequently, the prediction of PE early in gestation (≤ 16 weeks), long before symptoms present when, such a test would be most useful to guide the prophylactic use of potential therapeutics (e.g., low-dose aspirin) remains a key objective to improve obstetric care [9].

Clinical care may also be improved by a better understanding of the pathogenesis of PE. PE is a disease specific to humans as it does not occur in other species [32], and elucidating its pathogenesis has proven especially challenging. Broadly, it is accepted that PE occurs in two stages – abnormal placentation occurring early in pregnancy followed by systemic endothelial dysfunction [15,32,33]. Because PE can clinically present any time after 20 weeks of gestation and with a diversity of symptoms, significant effort has been made to subclassify the disease based upon the timing of onset (i.e., early-onset at < 34 weeks of gestation vs late-onset thereafter) as a proxy for pathology [34,35]; however, debate over the significance of such subtypes is still ongoing [32,33,36,37]. Noninvasive methods like liquid biopsies present a means to indirectly observe pathogenesis in real time and identify biological changes associated with PE for all proposed subtypes both prior to and at diagnosis.

Here, we report that cfRNA transcriptomic changes can distinguish between normotensive (NT) and PE pregnancies before clinical diagnosis and across gestation: early (5–12 weeks), mid (13–18 weeks), and late (23–33 weeks), and even into the post-partum period (0–2 weeks after delivery) regardless of PE subtype. Interestingly, the majority of these cfRNA changes are most striking early in pregnancy suggesting that the identified cfRNA signal may correlate with PE pathogenesis, which is thought to also occur at this time. Indeed, gene ontology (GO) analysis

identified pathways that reflect known PE biology. Finally, we identified and independently validated (n = 8 NT, 8 PE/gestational hypertension) 11 genes that when measured between 5–16 weeks of gestation can form a predictive PE signature with a specificity of 88% [55–99%] and sensitivity of 100% [74–100%] (AUC=94%, [79–100%]) (All reported as value, [95% confidence interval]). Taken together, our results showed that cfRNA measurements can form the basis for a clinically relevant test that would predict PE well before clinical presentation and can help characterize the pathogenesis of PE in real time.

Results

Clinical study design

To identify changes associated with PE well before traditional diagnoses, we designed a prospective study and recruited pregnant mothers at their first clinical visit, between 5–12 weeks of gestation, of which 66 were included in this study (28 NT, 38 with PE). For each participant, we then collected 4 plasma samples corresponding to early (5–12 weeks), mid- (13–18 weeks), and late (23–33 weeks) gestation, and post-partum (0–2 weeks after delivery) (Fig 1A). We also recorded maternal pre-pregnancy and pregnancy characteristics (Table S1). We defined a pregnancy as NT if it was both uncomplicated and went to full-term (≥ 37 weeks) and as PE based on current guidelines (see Methods). For mothers who developed PE, all antenatal blood samples were collected prior to diagnosis. To verify our results, we then split this larger group into discovery (n = 49, [20 NT, 29 with PE]) and validation (n = 17, [8 NT, 9 with PE]) cohorts that were processed at separate times. Our final analysis included a subset of these samples that passed pre-defined quality metrics as discussed in the Supplement (Supp. note 1, Methods, Fig S1,2).

After confirming sample quality, 147 samples from 57 mothers (26 NT, 31 with PE) were included in the final analysis. Specifically, 118 samples from 42 participants (18 NT, 24 with PE) were sequenced (RNAseq) in discovery and 29 samples from 15 participants (8 NT, 7 with PE) were processed using reverse-transcription quantitative polymerase chain reaction (RT-qPCR) in validation.

In the discovery cohort, we analyzed cfRNA from samples from four timepoints (early, mid-, and late gestation, and post-partum); whereas in validation, we focused on samples collected at early and mid-gestation because of the aforementioned clinical need for high quality predictive tests for PE early in gestation (Fig 1A). Across gestational timepoints in both cohorts, we found no significant difference in sampling time between PE and NT groups ($p \geq 0.7$, 0.06). Known risk factors for PE, such as pre-pregnancy maternal body mass index (BMI), maternal age, and the number of prior pregnancies, followed expected trends, but were not significantly different between PE and NT groups ($p \geq 0.19$, 0.52) (Fig 1B).

In mothers who later developed PE, a small fraction, 4 (17%) and 1 (14%) had a history of PE in the discovery and validation cohorts, respectively. We also observed no significant difference ($p = 0.14$, 1.0) in gestational age at onset between those who experienced mild (n = 11, 4) as compared with severe (n = 13, 3) symptoms (Fig 1C). Furthermore, 10 mothers who developed PE also delivered preterm (n = 9, 1) as compared with no mothers in the NT group as reflected by significantly different gestational ages at delivery ($p = 10^{-5}$, 0.25, one-sided test) (Fig 1D), which was consistent both with our NT group inclusion criteria and with epidemiological evidence that

PE increases the risk of spontaneous or indicated preterm delivery [5,27] (values reported as discovery, validation, Mann-Whitney rank test).

Identification of cfRNA changes PE across gestation in mothers who developed PE

To identify gene changes associated with PE across pregnancy, we performed differential expression profiling using the discovery cohort and excluded post-partum samples, which instead served as a hold-out set, since they presented a clear break in gestational time and a marked change in physiology – specifically the absence of contributions from the fetus and placenta. This analysis identified 556 differentially expressed genes (DEGs) that differed across gestation between mothers who later developed PE and NT mothers who did not experience complications (adjusted $p \leq 0.05$).

To further understand when these changes occur during gestation, we estimated the log fold-change (logFC) for each gene by early, mid-, and late gestation as well as post-partum. We observed that these gene changes occurred most strikingly in early and mid-gestation as indicated by a clear bimodal distribution with two peaks centered around logFC of ±0.75 and broadly showed no difference at post-partum (mean logFC = 0.02) (Fig 2A). We then quantified the relative dispersion around the estimated logFC for each gene using an approximation for the coefficient of variation (CV) or the ratio between an error bound, ∂, and the estimated logFC (Fig 2B). After bootstrapping ∂, defined here as the error bound associated with the lower (or upper in the case of negative logFC values) 95% confidence interval (CI), we found that gene changes were most stable during early and mid-gestation where over 40% of genes had a CV < 1 as compared to 27% at late gestation and 21% at post-partum.

We then asked whether a subset of gene changes approximately proportional in number to total sample number (n = 27, 28, 31, 30) was sufficient to segregate PE (n = 13, 14, 19, 16) and NT (n = 14, 14, 12, 14) samples across gestation. We performed hierarchical clustering on a subset of all DEGs for each sampled time group based on two criteria designed to identify gene changes that were stable (CV < 0.5) and striking ($|\text{logFC}| \geq 0.75$). In total, 51, 38, 17, and 14 genes met these criteria in early, mid-, late gestation, and post-partum, respectively, of which only 11 genes met the specific cutoff criteria at 2 or more timepoints (File S1). This method separated PE and NT samples across gestation and at post-partum with good specificity (93 [71–99], 56 [32–80], 75 [47–92], and 86% [62–97%]) and sensitivity (85 [59–97], 79 [53–94], 74 [52–89], and 81% [57–94%]) (Fig 2C, values reported for early, mid-, and late gestation and post-partum, [95% CI]).

Noting that only 11 genes were shared across timepoint specific DEG subsets after applying hard cutoffs, we then wondered whether changes associated with each subset persisted across timepoints but at reduced levels. Further analysis indicated that genes identified at one timepoint (i.e., early gestation) can moderately separate PE and NT samples from other timepoints (i.e., mid- and late gestation and post-partum), and that typically such separation is most pronounced at a timepoint closest to that in which the gene changes were identified (Fig S3). Finally, we validated 4 DEG changes in early and mid-gestation using RT-qPCR and validation dataset samples. For all but 1 gene at 1 timepoint (PLAC8, mid-gestation), which had a CV close to 1 indicating reduced confidence in the estimate, we found that the logFC estimated using this method were consistent with those estimated in discovery using RNAseq, an orthogonal method (Fig S4).

Next, we sought to understand these changes' longitudinal dynamics over gestation using k-means clustering. We observed that the 556 identified DEGs could be well categorized into 3 distinct trends (Fig 2D, S5). Resembling a valley or V-shape, the first trend (Group 1) described the longitudinal behavior of 66 genes (12%), for which measured levels were significantly reduced in PE samples (-2x to -4x) across gestation with a minimum in mid-gestation. Peaking in early gestation (2x), the second trend (Group 2) described the behavior of 212 genes (38%) that had significantly elevated levels in PE samples across early and mid-gestation and to a lesser extent in late gestation (1.4x). Finally, the third trend (Group 3) described the behavior of 278 (50%) genes with consistently lower levels (-1.4x) across gestation. In all cases, gene changes were far less evident post-partum and trended toward no difference between PE and NT, which may reflect a placental contribution to DEG levels. Separately, we confirmed that these patterns were not spurious (i.e., an artifact of the k-means clustering algorithm) by permuting the data thereby scrambling any time-related structure while preserving its overall distribution. Following permutation, we observed no longitudinal patterns, which were instead replaced by 3 nearly flat, uninformative trends (Fig S6).

Finally, we asked what pathways these genes participated in. Grouping genes by longitudinal behavior, GO analysis revealed clear associations with PE pathogenesis and progression (Fig 2E) (See File S2 for full list). Broadly, genes with decreased levels across gestation in Groups 1 and 3 corresponded to the immune response, cellular apoptosis (i.e., TNFR1-mediated ceramide production), and cellular metabolism and catabolism (i.e., heterocycle metabolic process, regulation of catabolic process). In contrast, genes in Group 2 that had increased levels across gestation were involved in endothelial function (i.e., platelet activation, signaling, and aggregation), cellular activation (i.e., signaling by Rho GTPases), cellular invasion (i.e., lamellipodium organization), and wound healing (i.e., hemostasis).

Development of a machine learning classifier to predict future PE onset on or before 16 weeks of gestation

Having established that we can observe gene changes associated with the development of PE across gestation, we sought to build a robust classifier that can identify mothers at risk of PE at or before 16 weeks of gestation. After training, the final model performed well on the discovery dataset with a near perfect AUC (94%, [87–100%]), perfect specificity (100%, [91–100%]) and positive predictive value (PPV) (100%, [89–100%]), and high sensitivity (88%, [70–96%]) and negative predictive value (NPV) (89%, [74–97%]) (Fig 3A, Table 1).

We then tested this model on a completely independent cohort recently reported on by Del Vecchio and colleagues [31] who isolated and sequenced cfRNA from 25 samples collected between 12–17 weeks of gestation from mothers who experienced diverse complications (n = 8 NT, 8 with PE/gestational hypertension, 7 with gestational diabetes, 2 with chronic hypertension). Across this entire dataset, the final model once again performed well with high AUC (88%. [78–97%]), perfect sensitivity (100%, [74–100%]) and NPV (100%, [83–100%]), and good specificity (76%, [53–91%]) and PPV (67%, [39–88%]) (Fig 3A, Table 1).

Upon further inspection of the 4 erroneously classified samples (false positives), 1 and 2 were from mothers who had chronic hypertension or later developed gestational diabetes, respectively – known risk factors for PE [38–44]. To match our discovery cohort's control group, we then focused

on comparing samples from PE or gestational hypertension to samples from NT participants alone and observed an AUC (94%, [79–100%]) identical to that measured in the discovery cohort, and equally high values for sensitivity (100%, [74–100%]), specificity (88%, [55–99%]), PPV (89%, [59–99%]), and NPV (100%, [71–100%]).

The final model also proved well-calibrated across both datasets with a calibration curve slope of about 1 (1.27, 1.06) and intercept of close to 0 (-0.06, -0.05) (Values reported as Discovery, Del Vecchio et al.). The probability of PE was also estimated as nearly 0% for most NT or otherwise complicated pregnancy samples and as almost 100% for PE or gestational hypertension samples. Far fewer samples were estimated to have a 50% probability of PE, the cutoff for classification (Fig 3B).

We then wondered whether this model provided any predictive power for samples collected later in gestation or at diagnosis between 17-38 weeks' gestation. To answer this, we measured model performance across three datasets: the discovery cohort (samples collected ≥ 16 weeks of gestation but not post-partum) and two independently collected cohorts [30]. For both independent cohorts (iPEC, PEARL-PEC), Munchel and colleagues isolated and sequenced cfRNA from 97 NT (n = 73, 24) and PE (n = 40, 24) samples collected at time of PE diagnosis or a matched timepoint for NT samples. As expected, we found a reduced, but still significant, signal to predict PE with moderate AUC (79% [64–93%], 60% [50–69%], 62% [49–77%]), sensitivity (73 [48–90], 68 [52–80], 62% [43–80%]), and specificity (85 [59–97], 52 [41–63], 62% [43–80%]) (Values reported as discovery, iPEC, PEARL-PEC) (Table S2, Fig 3C). At these later gestational timepoints shortly prior to or at diagnosis, the model remains moderately calibrated in discovery and poorly calibrated in testing with a calibration curve slope of less than 1 (0.7, 0.07, 0.12) and intercept of close to 0 (0.06, 0.28, 0.43) (Values reported as discovery, iPEC, PEARL-PEC). Nonetheless, the model still discriminated between PE and NT pregnancies with many samples still receiving probability estimates at either extreme (0 or 100%) but an increased sample number with poor scores around 50% (Fig 3D) (All reported as value, [95% CI]).

Finally, we inspected the 11 genes (Table S3) used by the model to yield probability estimates. We find that many of the trends (i.e., decreased or increased gene levels in PE) observed in the discovery dataset are upheld in Del Vecchio and colleagues' dataset, although not to the level of statistical significance after multiple hypothesis correction (adjusted $p \geq 0.40$, one-sided Mann-Whitney rank test) (Fig 3E). We also found that some models trained using a subset of the 11 initial genes can predict future PE onset with varying performance. Notably, performance improved across all metrics (sensitivity, specificity, PPV, NPV, and AUC) as we increased the number of genes included for model training with some subsets of sizes 4–10 achieving nearly equivalent performance to the full model, which further emphasizes the importance of the gene signature in aggregate as opposed to individually (Table S4, Fig S7).

Discussion

Noninvasive measurements of the cf-transcriptome present an opportunity to study human development and disease from any organ at a molecular scale. Here, we show that cfRNA measurements taken during pregnancy can clearly distinguish between PE and NT pregnancies with the most striking differences occurring early on and broadly showed no difference at post-partum, consistent with known PE biology, namely that the placenta drives the disease process

[32,33]. These findings are supported by orthogonal analyses including clustering and machine learning. Specifically, we find that the most striking and stable changes as defined by high |logFC| and low CV occur more frequently in early or mid-gestation as compared to later on and as broadly supported by longitudinal dynamics analysis. Our observation that cfRNA changes are less stable in late gestation may be consistent with what is known about fetal development over gestation, specifically that fetal variability increases with gestation as supported by ultrasound data [45]. Importantly, small subsets of the DEGs identified here (n = 11–51) can also be used to separate PE and NT pregnancies with the best sensitivity and specificity observed early in pregnancy as confirmed by both hierarchical clustering and machine learning.

The genes we identify and their longitudinal dynamics are also consistent with what we know about PE, namely that it likely develops very early in pregnancy with the resultant maternal syndrome manifesting later [32]. Following GO analysis, we find that the DEGs identified here both agree with prior evidence regarding PE's pathology and clinical presentation and provide new hypotheses to explore. Specifically, GO terms that relate to cellular invasion and organ perfusion (i.e., lamellipodium organization and hemostasis, Group 2), cellular apoptosis, catabolism and response to stress (i.e., TNFR1-mediated ceramide production, regulation of catabolic process, Groups 1 and 3), and endothelial dysfunction (i.e., platelet activation, signaling, and aggregation, Group 2) closely reflect what is known about the stages of the pathogenesis of PE: abnormal placentation as characterized by insufficient placental invasion of the uterine arteries and subsequent reduced placental perfusion, inadequate nutrient exchange, and systemic endothelial dysfunction [32].

Early in pregnancy during normal placentation, cytotrophoblasts invade the spiral arteries, which then undergo a transformation to accommodate larger blood volumes for increased fetal and placental nutrient requirements later in pregnancy, and peripheral villi regress via a mechanism involving oxidative stress. PE is thought to be a uniquely human disease because of the large nutritional exchange requirement imposed on the mother (3x higher than that for other mammals in the third trimester) [32]. Later on, PE presents in the mother as systemic endothelial dysfunction, where endothelial cells are closely implicated in platelet aggregation. Other terms like innate immune system process, type I interferon production, and positive regulation of T cell differentiation (Group 3) highlight the involvement of the innate and adaptive immune system, which is both consistent with previous work on immune system dynamics during early gestation in PE [46] and with known pregnancy biology about sustaining a successful pregnancy [47] and establishing maternal tolerance, the lack of which has been previously associated with PE [48].

PE is also a broad and complex syndrome. Because the complication can present clinically across more than 20 weeks with a diversity of symptoms, significant effort has been made to subclassify the disease based on timing of onset. Used as a proxy for hypothesized pathogenesis, the timing of PE onset subdivides the disease into 'placental' early-onset PE (occurring before 34 weeks of gestation) and 'maternal' late-onset PE (occurring on or after 34 weeks) [32,34,49]. However, early- and late-onset PE may represent a spectrum of disease severity that corresponds with timing of onset and may lead to additional pregnancy complications, such as intrauterine growth restriction (IUGR) [32,33,35,36]. Indeed, some evidence like transcriptional profiling of whole blood drawn at diagnosis suggests a common gene signature for both subtypes [37]. Our findings further corroborate

321 that timing of onset may lie on a spectrum, and a common signature can identify both early- and
322 late-onset PE of varied severity very early in gestation including as early as 5 weeks.
323
324 Finally, prediction of PE early in pregnancy has remained an unrealized objective for the field of
325 obstetrics with important health consequences. Our results here provide encouraging evidence that
326 it is possible both to predict PE risk before 16 weeks of gestation with high specificity and
327 sensitivity, and to differentiate PE from other pregnancy-related complications such as gestational
328 diabetes or underlying conditions such as chronic hypertension in a small, independent cohort. Of
329 the 7 gestational diabetes samples included in the independent cohort of Del Vecchio et al., only
330 2 were misclassified as PE. We note that identifying other adverse outcomes that increase one's
331 risk of PE may be an acceptable error mode whereas classifying an uncomplicated, NT pregnancy
332 as PE is not and would result in unnecessary, added stress. In our analysis, we found that only 1
333 NT pregnancy was misclassified across 25 discovery and 8 test samples. Given the small sample
334 size and limited ethnic and racial diversity in these cohorts, we emphasize that the performance
335 metrics although consistent and striking must still be validated in a larger and more diverse follow-
336 up study.
337
338 Notably, these predictions also relied on only 11 gene measurements – none of which overlapped
339 with previously reported genes that are not altered until much closer to diagnosis (sFlt1 and PlGF)
340 [25] or the cfRNA results in recent proof-of-concept work [30,31]. We also note that 11 genes can be
341 easily measured using RT-qPCR, highlighting its clinical relevance and presenting an inexpensive
342 alternative to RNAseq in future validation work. Given that both expense and efficacy are
343 important factors in addressing barriers to care, an inexpensive test with striking efficacy may
344 dramatically improve care and prevent missed or delayed diagnoses, which at the moment may
345 account for as many as 3 in 5 maternal deaths related to pregnancy in the US [8].
346
347 Tests such as this one that identify who is at risk of PE early in pregnancy have long been
348 acknowledged as an unrealized key objective in PE care [9]. They can also importantly be coupled
349 with the administration of low-dose aspirin, which if started early has been shown to reduce risk
350 of developing PE [10,28]. Although these results must be further confirmed in a blinded, larger
351 validation study, the findings reported here demonstrate that an 11 gene cfRNA-based classifier
352 can predict risk of PE with consistent, clinically relevant high sensitivity and specificity (88–100%
353 for both metrics) across 2 independent datasets that were collected and processed by completely
354 separate teams.
355
356 Taken together, we have shown that cfRNA measurement can form the basis for a robust liquid
357 biopsy test that predicts PE very early in gestation, to date an unrealized key objective for obstetric
358 care, and can be used to help characterize the pathogenesis of PE in real time.

Acknowledgements. We thank N. Aghaeepour and I. Maric from Stanford University for helpful discussion and data analysis suggestions, the many researchers and clinicians affiliated with the March of Dimes Prematurity Research Center at Stanford University for insightful feedback after oral presentations, and the participants in this study for their invaluable contribution.

Funding. This work was supported by the Chan Zuckerberg Biohub and the March of Dimes Foundation. The March of Dimes Foundation provided funding for sample collection and processing. The Chan Zuckerberg Biohub provided funding for RNA sequencing and cloud computing. M.N.M. is supported by the Stanford Bio-X Bowes Fellowship. V.D.W. is supported by the H&H Evergreen Fund.

Author contributions. M.N.M., G.M.S., D.K.S., and S.R.Q. conceptualized and designed this study. M.L.D., V.D.W., G.M.S., and D.K.S. designed the cohort study at Lucile Packard Children's Hospital. R.J.W., A.A.C., and C.C.Q. collected and processed whole blood samples and corresponding sample and participant metadata. M.N.M. developed experimental protocols for cfRNA extraction and data generation and processed all samples from cfRNA extraction up to and including RT-qPCR and library generation for RNA sequencing. R.V.S., M.T., and N.F.N. developed and executed experimental protocols for RNA sequencing and computational protocols for raw sequencing data transfer. M.N.M. conceptualized and developed computational analyses, analyzed the data, and wrote the initial manuscript draft in collaboration with S.R.Q.. All authors contributed to the writing and editing the manuscript.

Competing interests. The authors declare the following competing interests: M.N.M., G.M.S., D.K.S., and S.R.Q. are inventors on a patent application submitted by the Chan Zuckerberg Biohub and Stanford University that covers noninvasive early prediction of preeclampsia. S.R.Q. is a founder, consultant, and shareholder of Mirvie.

Materials & methods

Clinical study design

For this longitudinal, prospective study, we enrolled pregnant mothers (aged 18 years or older) receiving routine antenatal care on or prior to 12 weeks of gestation at Lucile Packard Children's Hospital at Stanford University, following study review and approval by the Institutional Review Board at Stanford University (#21956). All signed informed consent prior to enrollment. Whole blood samples for plasma isolation were then collected at three distinct timepoints during their pregnancy course and once post-partum. Mothers were defined as having PE (cases, n = 38) based upon current American College of Obstetrics and Gynecology (ACOG) guidelines (see below). Our control (NT) cohort (n = 28) consisted of women who had uncomplicated term pregnancies and either normal spontaneous vaginal or caesarean deliveries. For mothers who developed PE, all antenatal samples included in this study were collected prior to clinical diagnosis.

Definition of preeclampsia

Preeclampsia was defined per the ACOG guidelines [27] based on two diagnostic criteria: 1) new-onset hypertension developing on or after 20 weeks of gestation and 2) new-onset proteinuria or in its absence, thrombocytopenia, impaired liver function, renal insufficiency, pulmonary edema, or cerebral or visual disturbances.

New-onset hypertension was defined when the systolic and/or diastolic blood were at least 140 or 90 mmHg, respectively, on at least 2 separate occasions between 4 hours and 1 week apart. Proteinuria was defined when either 300 mg protein was present within a 24-hour urine collection or an individual urine sample contained a protein/creatinine ratio of 0.3 mg/dL, or if these were not available, a random urine specimen had more than 1 mg protein as measured by dipstick. Thrombocytopenia, impaired liver function, and renal insufficiency were defined as a platelet count of > 10,000/μL, liver transaminases ≥ 2x of normal, and serum creatinine > 1.1 mg/dL, respectively.

Symptoms were defined as severe per the ACOG guidelines. Specifically, PE is defined as severe if any of the following symptoms were present and diagnosed as described above: new-onset hypertension with systolic and/or diastolic blood pressure of at least 160 or 110 mmHg respectively, thrombocytopenia, impaired liver function, renal insufficiency, pulmonary edema, new-onset headache unresponsive to medication and unaccounted for otherwise, or visual disturbances.

Finally, a pregnant mother was considered to have early-onset PE if onset occurred before 34 weeks of gestation and late onset thereafter.

Sample preparation

Plasma processing

Blood samples were collected in either EDTA-coated (Cat No 368661, Becton-Dickinson) or Streck cfRNA BCT (Cat No 218976, Streck) tubes at early, mid, and late gestation, and post-partum for each participant. Within 30 minutes, tubes were then centrifuged at 1600g for 30 minutes at room temperature. Plasma was transferred to 2-mL microfuge tubes and centrifuged at 13000g for another 10 minutes in a microfuge. One milliliter aliquots were then transferred to 2-mL Sarsted screw cap microtubes (Cat No 50809242, Fisher Scientific) and stored at -80˚C until analysis.

cfRNA isolation

In 96-sample batches, cfRNA from 1 mL plasma samples was extracted in a semi-automated fashion using the Opentrons 1.0 system and Norgen Plasma/Serum Circulating and Exosomal RNA Purification 96-Well Kit (Slurry Format) (Cat No 29500, Norgen). Samples were subsequently treated with Baseline-ZERO DNAse (Cat No DB0715K, Lucigen) for 20 minutes at 37˚C. DNAse-treated cfRNA was then cleaned and concentrated into 12 μL using Zymo RNA Clean and Concentrator-96 kit (Cat No R1080, Zymo).

Following cfRNA extraction from plasma samples, isolated RNA concentrations were estimated for a randomly selected 11 samples per batch using Bioanalyzer RNA 6000 Pico Kit (Cat No 5067-1513, Agilent) per manufacturer instructions.

Sequencing library preparation

cfRNA sequencing libraries were prepared with SMARTer Stranded Total RNAseq Kit v2 - Pico Input Mammalian Components (Cat No 634419, Takara) from 4 μL of eluted cfRNA according to the manufacturer's instructions. Samples were barcoded using SMARTer RNA Unique Dual Index

Kit – 96U Set A (Cat No 634452, Takara), and then pooled in an equimolar fashion and sequenced on Illumina's NovaSeq platform (2x75 bp) to an average depth of 50 million reads per sample.

RT-qPCR assay

We performed RT-qPCR in two stages: reverse transcription and preamplification followed by Taqman probe-based qPCR per manufacturer instructions. All primers and controls were commercially designed and validated (Bio-Rad), and their unique assay IDs can be found in the supplement (Table S5). Specifically, we prepared cDNA using random hexamers from an initial volume of 4 µL cfRNA per sample (Bio-Rad Reliance Select cDNA Synthesis Kit, Cat No 12012801). Prior to reverse transcription (RT), we spiked in 1 µL RT control probe into each sample to measure the latter step's efficiency. After RT, we combined 5 µL of prepared cDNA with a primer pool, which was then preamplified for 12 cycles and subsequently diluted 1:5 (Tris EDTA) (Bio-Rad SsoAdvanced PreAmp Supermix, Cat No 1725160). Finally, we performed each qPCR assay in triplicate on the Bio-Rad CFX384 touch machine. Each replicate contained 1 µL of pre-amplified and diluted cDNA sample, 0.5 µL of the corresponding primers, and super-mix at the appropriate concentration (Bio-Rad SsoAdvanced Universal Probes Supermix, Cat No 1725285). Following qPCR, cycle threshold (C_t) values were extracted using CFX software.

Sequencing data analysis

Bioinformatic processing

For each sample, raw sequencing reads were trimmed using trimmomatic (v 0.36) and then mapped to the human reference genome (hg38) with STAR (v 2.7.3a). Duplicate reads were then removed by GATK's (v 4.1.1) MarkDuplicates tool. Finally, mapped reads were sorted and quantified using htseq-count (v 0.11.1). Read statistics were estimated using FastQC (v 0.11.8).

Across samples, the bioinformatic pipeline was managed using snakemake (v 5.8.1). Read and tool performance statistics were aggregated across samples and steps using MultiQC (v 1.7). Following sample quality filtering, all gene counts were adjusted to $\log_2$-transformed counts per million reads (CPM) with trimmed mean of M values (TMM) normalization [50].

Sample quality filtering

For every sequenced sample, we estimated three quality parameters as previously described [51,52]. To estimate RNA degradation in each sample, we first counted the number of reads per exon and then annotated each exon with its corresponding gene ID and exon number using htseq-count. Using these annotations, we measured the frequency of genes for which all reads mapped exclusively to the 3' most exon as compared to the total number of genes detected. RNA degradation for a given sample can then be approximated as the fraction of genes where all reads mapped to the 3' most exon. To estimate ribosomal read fraction, we compared the number of reads that mapped to the ribosome (Region GL00220.1:105424-118780, hg38) relative to the total number of reads (Samtools view). To estimate DNA contamination, we quantified the ratio of reads that mapped to intronic as compared to exonic regions of the genome.

After measuring these three metrics across nearly 700 samples, we empirically estimated each one's 95th percentile bound. We considered any given sample an outlier, low quality sample if its value for at least one of these metrics was greater than or equal to the 95th percentile bound.

Once values for each metric was estimated across the entire dataset, we visualized: (1) whether low-quality samples clustered separately using hierarchical clustering (average linkage, Euclidean distance metric) and (2) whether sample quality drove variance in gene measurements using principal component analysis (PCA). These analyses were performed in python (v 3.6) using scikit-learn for PCA (v 0.23.2), scipy for hierarchical clustering (v 1.5.1), and nheatmap for heatmap and clustering visualization (v 0.1.4).

Differential expression analysis

Differential expression analysis was performed in R using limma (v 3.38.3). To identify gene changes associated with PE across gestation, we used a mixed-effects model and included gestational age (continuous variable), case (binary variable, NT = 0 vs PE = 1), and the interaction between the two in the design matrix. We chose to model gestational age as a continuous variable, specifically a natural cubic spline with 4 degrees of freedom to account for the range across which samples were collected (1–3 months per collection period). We also blocked for participant identity (categorical variable), modeling it as a random effect to account for auto-correlation between samples from the same person. Post-partum samples were excluded from differential expression, instead serving as a hold-out dataset since they presented a clear break in gestational time.

Per the limma-voom guide, to account for sample auto-correlation over time, we ran the function voomWithQualityWeights twice. We first ran it without any blocking on participant identity, and used this base estimation to approximate sample auto-correlation based on participant identity using the function duplicateCorrelation. After estimating correlation, voomWithQualityWeights was run again, this time blocking for participant identity and including the estimated auto-correlation level. A linear model was then fit for each gene using lmFit and differential expression statistics were approximated using Empirical Bayes (eBayes) methods. Finally, differentially expressed genes were identified using the relevant design matrix coefficients and the function, topTable, with Benjamini-Hochberg multiple hypothesis correction.

log(Fold change) and coefficient of variation (CV) estimation

We define $\log_2$-transformed fold-change (logFC) as the difference between the median gene level (logCPM, see Bioinformatic processing section above) between PE and NT samples for a given sample collection period (i.e., early, mid, and late gestation, or post-partum). In the case where a given person had multiple samples included into a specific collection period, we only used the values associated with the first collected sample to avoid artificially reducing within-group (PE or NT) variance due to auto-correlation among samples from the same person.

We define the CV using an approximation. Specifically, we consider CV to be the ratio between an error bound, ∂, and the estimated logFC. We defined the error bound, ∂, as the one-sided error bound associated with the lower (or upper in the case of negative logFC values) 95% CI as estimated by bootstrapping. This definition of ∂ as a one-sided bound that approaches 0 (equivalent to no FC) allowed us explore how confident we could be in an estimated logFC. For instance, a CV = 1 would indicate that at the boundary of proposed values, the logFC for a given gene becomes effectively 0. Similarly, a CV > 1 would indicate even less confidence in a proposed average logFC and indicate that at the boundary, the estimated logFC changes signs (i.e., a negative logFC becomes a positive one or vice versa).

543
544 ***Hierarchical clustering analysis***
545 For each sample collection period, hierarchical clustering was performed using differentially
546 expressed genes with an |logFC| ≥ 0.75 and CV < 0.5. For each gene that passed these thresholds,
547 we calculated a z-score across all samples (at most 1 per subject, the earliest collected sample in a
548 given group) in each sample collection period using the function StandardScaler in scikit-learn (v
549 0.23.2), Average linkage hierarchical clustering with a Euclidean distance metric was then
550 performed for both rows (gene z-scores) and columns (samples in same collection group) for a
551 given matrix in python using scipy (v 1.5.1). Clustering and corresponding heatmaps were
552 visualized using nheatmap (v 0.1.4).
553
554 ***Longitudinal dynamics analysis***
555 To group gene changes by longitudinal behavior, we performed k-means clustering on a matrix
556 where each row was a differentially expressed gene and each column was the estimated logFC for
557 a given sample collection period (556 genes x 4 timepoints). We measured the sum of squared
558 distances for every 'k' between 1 and 16 (4^2) where 16 represents the maximum possible k (4
559 timepoints with 2 possibilities each, logFC > 0 or logFC < 0). We then identified the optimal k
560 clusters using the well-established elbow method to identify the smallest 'k' that can best explain
561 the data, visually described as the elbow (or knee) of a convex plot like that for the sum of squared
562 distances vs k (Fig S5). To do so, we used the function KneeLocator as implemented in the package
563 kneed (v 0.7.0). Having identified the optimal number of clusters, 'k', we labeled every
564 differentially expressed gene with its assigned cluster and visualized average behavior (mean) and
565 the 95% CI per cluster using Seaborn line plot (v 0.10.0).
566
567 To confirm that the identified patterns were not spurious (i.e., an artifact of the k-means clustering
568 algorithm), we permuted the data columns (logFC per timepoint) for each gene thereby scrambling
569 any time-related structure while preserving its overall distribution. We then visualized the result
570 using Seaborn line plot as described above.
571
572 ***Gene ontology (GO) analysis***
573 GO analysis was performed using the tool, GProfiler (v 1.0.0), for the following data sources,
574 Gene ontology: biological processes (GO:BP) and Reactome (REAC). In identifying GO terms,
575 electronic GO annotations (IEA) were included in the analysis. DEGs related to the ribosome were
576 excluded from GO analysis given their extensive annotation that can lead to spurious term
577 associations. To narrow this initial list (77 GO terms) to a smaller number for plotting purposes,
578 the initial GO table was then pruned to only include parent terms (as filtered by the column,
579 parents).
580
581 ***Logistic regression feature selection and training***
582 To build a robust classifier that can identify mothers at risk of PE at or before 16 weeks of
583 gestation, we first pre-selected features using the differential expression analysis (see section with
584 same name above) as a starting point. Next, to correct for batch effect, where we define batch as a
585 set of samples processed at the same time by a distinct group (e.g., Discovery cohort = batch, Del
586 Vecchio and colleagues' cohort = batch), we subtracted the mean logCPM per gene from all
587 cfRNA measurements in the batch (zero-centering). We note that all cohorts used in this work had
588 balanced batch-group design and therefore met the requirements for zero-centering to work as

589 expected [53]. Model training then used the entire discovery dataset and consisted of two stages –
590 further feature pre-selection based on three metrics followed by the construction of a logistic
591 regression model with an elastic net penalty.
592
593 For feature pre-selection, we focused on three practical metrics measured across all samples
594 collected on or before 16 weeks of gestation: gene change size (|logFC|), gene change stability
595 (CV), and median gene expression (median CPM). Filtering on median gene expression helped
596 ensure that any genes selected could be robustly detected in smaller initial plasma volumes like
597 those that might be drawn in the clinic. All model hyperparameters and feature pre-selection
598 cutoffs were tuned using accuracy as the outcome metric and leave-one-out cross validation.
599
600 To understand the importance of each gene feature, we trained a separate logistic regression model
601 for every possible feature subset (2046 total combinations of 1-10 genes). No feature pre-selection
602 was performed for this sub-analysis. All model hyperparameters were tuned as previously
603 described. We defined a gene subset as weakly predictive if the model yielded an AUC > 50% on
604 the test sets (dataset of Del Vecchio et al.).
605
606 In all cases, performance metrics were assessed as described below (see next section).
607
608 ***Calibration analysis***
609 We assessed model calibration using the calibration curve's slope and intercept as recommended
610 by Van Calster and colleagues [54] for small sample numbers (i.e., n < 200 for case group, n < 200
611 for control group). Specifically, we estimated the calibration curve using the function,
612 calibration_curve, from sklearn.calibration and 10 bins. We then estimated slope and intercept
613 using the function polyfit with 1 degree from numpy. As described by Van Calster and colleagues,
614 we considered a well-calibrated model to have a slope of approximately 1 and an intercept of
615 approximately 0. Curve slopes greater than 1 suggested moderate, underestimated risk scores
616 whereas slopes less than 1 suggested the opposite. Additionally, negative calibration curve
617 intercept values suggested underestimation whereas positive values suggested overestimation.
618
619 ***Performance metric analysis***
620 Model performance was assessed using several statistics including sensitivity, specificity, PPV,
621 NPV, and ROC AUC. Given a 2x2 confusion matrix where rows 1and 2 represent true negatives
622 and positives and columns 1 and 2 represent negative and positive predictions respectively, we can
623 define the value in row 1, column 1 as true negatives (TN), row 1, column 2 as false positives (FP),
624 row 2, column 1 as false negatives (FN), and row 2, column 2 as true positives (TP). We can then
625 define the following proportions: (1) Sensitivity = TP / (TP + FN) (2) Specificity = TN / (TN +
626 FP) (3) PPV = TP / (TP + FP) (4) NPV = TN / (TN + FN). For each proportion, we calculated 95%
627 CIs using Jeffrey's interval [55] and the function, proportion_confint, from
628 statsmodels.stats.proportion. We also approximated AUC and its corresponding 95% CI using the
629 scikit-learn function, roc_auc_score, and bootstrapping respectively.
630
631 **RT-qPCR data analysis**
632 ***Sample quality filtering***
633 RT-qPCR and sample quality were assessed using two controls. First, we confirmed the absence
634 of genomic DNA contamination using a commercially available control (Table S5). Any sample

where genomic DNA was detected as defined by the manufacturer would have been excluded from further analysis; however, in this work, no sample contained DNA contamination at a detectable level. Next, we measured reverse-transcription efficiency (RT) using a control probe spiked in at a known concentration prior to RT and subsequently measuring the probe's C_t following preamplification (no primers included for RT in preamplification) and qPCR as detailed above (see RT-qPCR assay). For inclusion in any further analysis, we required that the RT probe (at levels below the corresponding non-template control) be detected in a given sample well.

Fold-change estimation

We estimated fold-change from RT-qPCR data using the $\Delta\Delta C_t$ method. Specifically, for every gene of interest (G_i), we first estimated $\Delta C_t = C_t\ (G_i) - C_t\ (\text{RT control})$ to account for different reverse-transcription efficiencies across samples. For more details about the RT control, refer to the sections titled "Sample quality filtering" and "RT-qPCR assay". Having defined ΔC_t for every gene measured in each sample, we can then define $\Delta\Delta C_t$ for a given gene as the difference between the median ΔC_t in the PE group and in the NT group. Fold-change can then be defined as $2^{-\Delta\Delta C_t}$.

Statistical analysis

All p-values reported herein were calculated using the non-parametric Mann-Whitney rank test unless otherwise stated. One-sided tests were performed where required based on the hypothesis tested.

Code/data availability

All computational analyses were performed using Python 3.6 or R 3.5, and will be available on Github. Raw and processed sequencing and RT-qPCR data will be deposited with the SRA and GEO respectively.

Figure 1. Sample, maternal, and pregnancy characteristics with the exception of gestational age at delivery are matched across NT and PE groups. Panels illustrate matched sample collection time (weeks) in (**A**), matched maternal characteristics (Left to right: BMI, age, and previous pregnancies) in (**B**), matched gestational age at PE onset regardless of PE symptom severity in (**C**), and gestational age at delivery in (**D**). For (**A**), schematic depicts blood sampling across gestation and plasma isolation. X-axis represents gestational age during pregnancy and time post-delivery (weeks) thereafter (ns = not significant, ** = 10^{-5}).

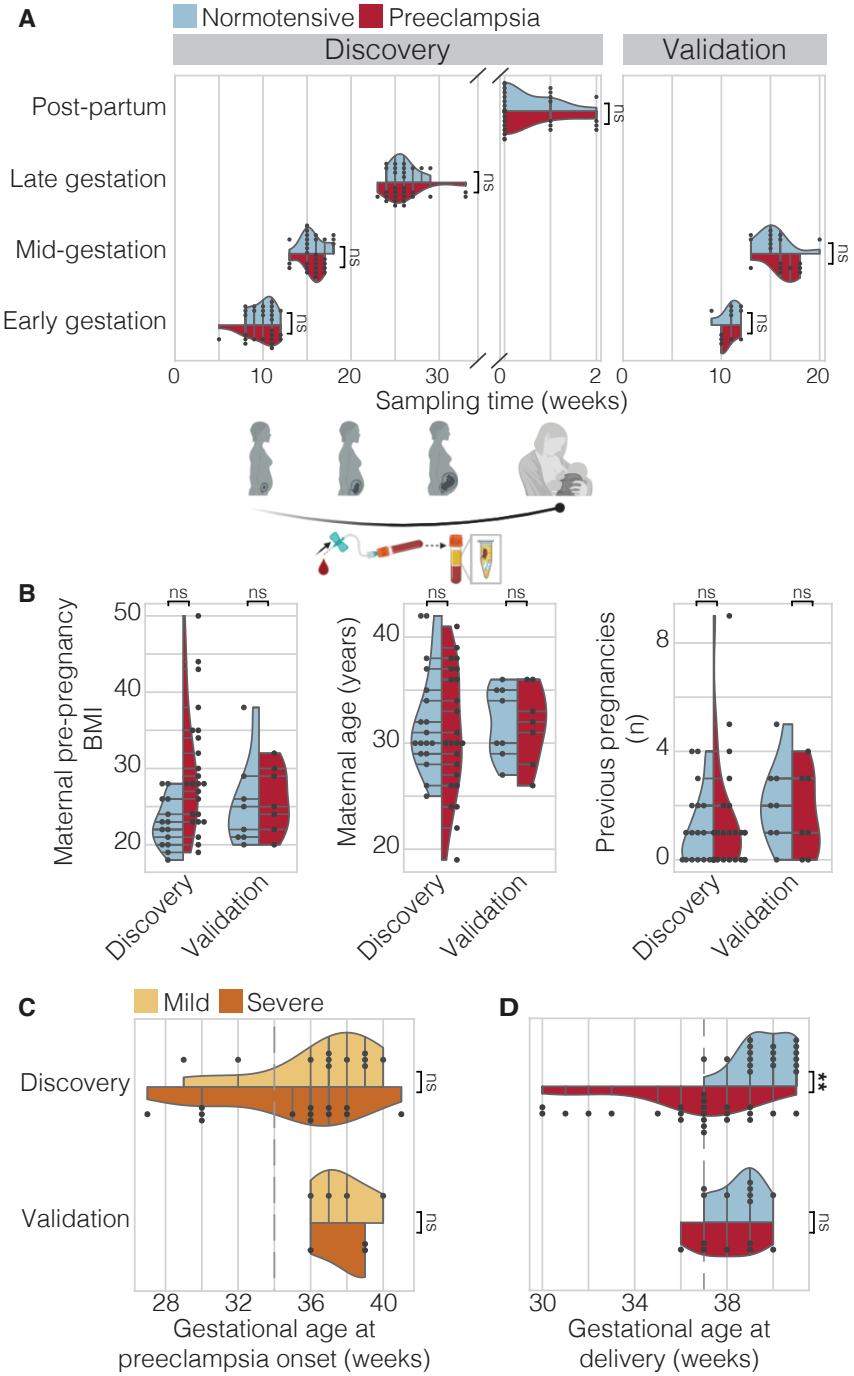
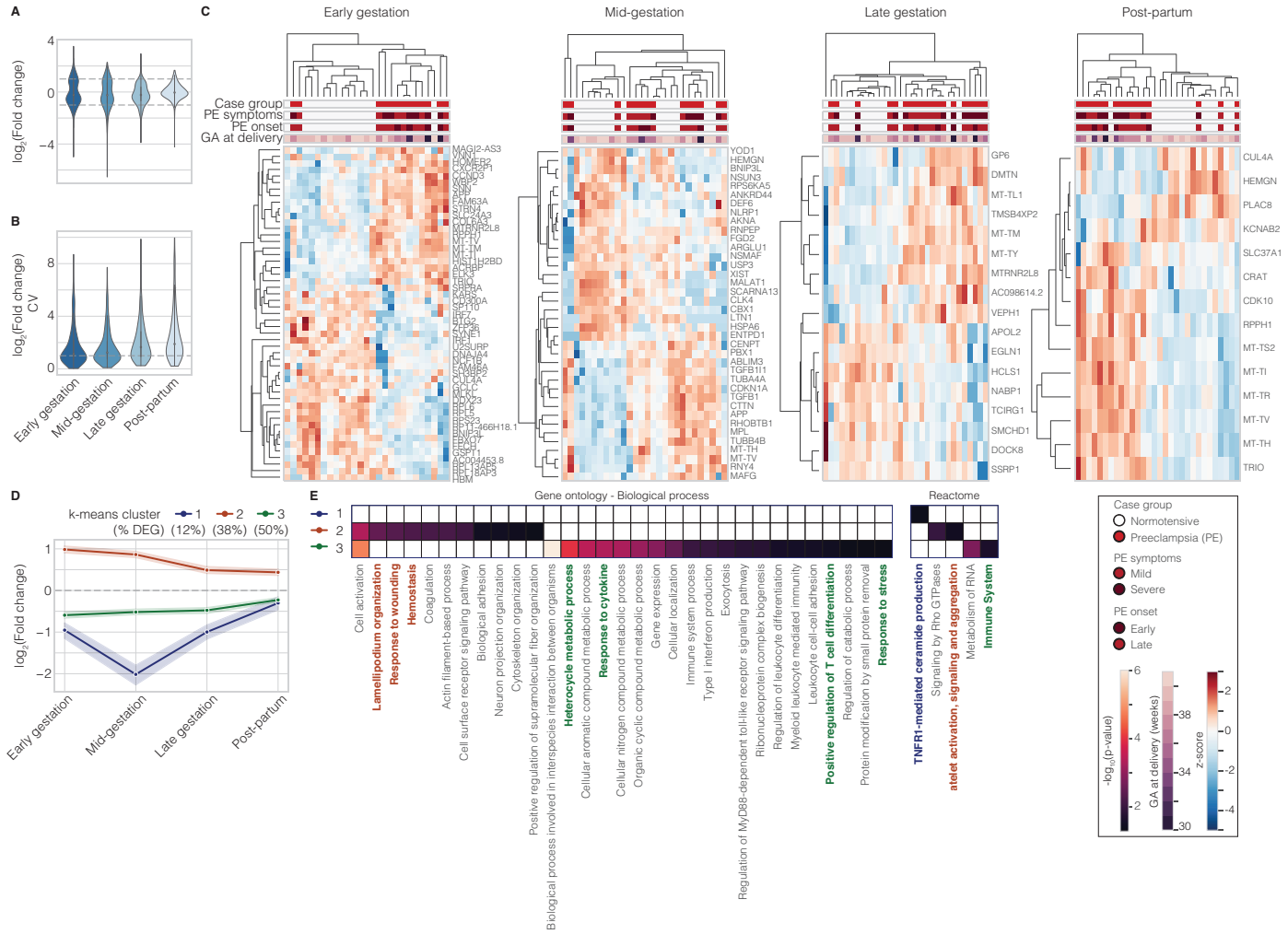
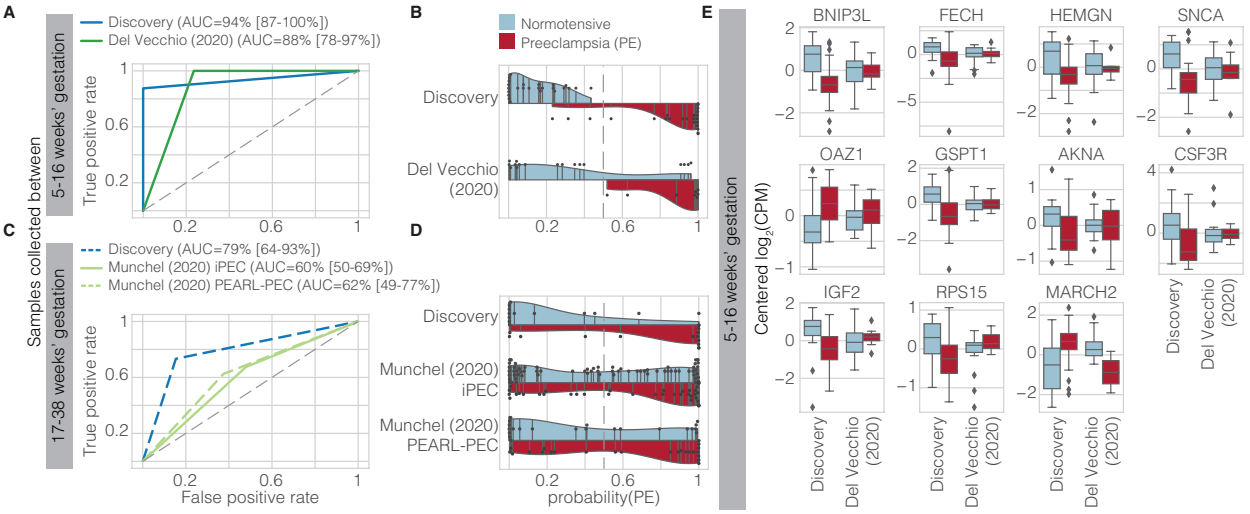


Figure 2. Across gestation and prior to diagnosis, changes in the cfRNA transcriptome segregate PE and NT samples and agree with known PE biology. (**A**) Distribution of log(Fold change) with dashed lines at log(Fold change) = ± 1 and (**B**) CVs with dashed line at CV = 1 for all differentially expressed genes between PE and NT samples across gestation. In (**B**), 2, 9, 5, and 10 genes with CV > 10 across early, mid-, late gestation and post-partum respectively are omitted. (**C**) Across gestation, in each sample collection period, a subset of differentially expressed genes can separate PE and NT samples despite differences in symptom severity, PE onset subtype, and gestational age (GA) at delivery. (**D**) Across gestation, differentially expressed genes can be described by 3 longitudinal trends as revealed by k-means clustering. (**E**) The genes in each longitudinal trend group reflect known PE biology as highlighted across two databases (GO biological processes and the reactome). Some PE associated terms are emphasized in bold, colored text that corresponds to group color. Heatmap only includes parent terms (see File S2 for full list).



683 **Figure 3. A subset of cfRNA changes can predict risk of PE early in gestation** (**A**, **C**) Classifier
684 performance as quantified by a receiver operating characteristic curve for samples collected in
685 early gestation between 5–16 weeks in (**A**) and samples collected later in gestation between 17–
686 38 weeks in (**C**). For each dataset, including 3 independent previously published datasets, legend
687 states the AUC and the corresponding 95% CI in square brackets. (**B**, **D**) Estimated probability of
688 PE as outputted by logistic regression for both PE and NT samples in each dataset shows that the
689 model is well-calibrated and fairly confident (probability(PE) = 0 or 1) across most predictions.
690 Dashed line at 0.5 indicates classifier cutoff where probability(PE) > 0.5 constitutes a sample
691 predicted as PE. (**E**) Prediction of PE incorporates cfRNA levels for 11 genes for which centered
692 $\log_2$(Fold change) trends hold across discovery and independent validation (Del Vecchio (2020))
693 datasets.
694



695

696 **Table 1. PE prediction performance metrics for samples collected early in gestation (between**
697 **5–16 weeks).** Control and case sample numbers are reported as the total sample number and in
698 parentheses, the number of samples misclassified. Case column also include the sample number
699 split between PE and gestational hypertension (GH) cases in square brackets. All other statistics
700 including sensitivity specificity, PPV, NPV, and AUC are reported as the estimated percentage
701 followed by the 95% CI in square brackets. In Del Vecchio[1], the control group is defined as
702 samples from any pregnant mother who did not develop PE or GH including those with other
703 underlying or pregnancy-related complications like chronic hypertension and gestational diabetes
704 respectively. In Del Vecchio[2], the control group is defined as samples strictly from NT pregnant
705 mothers who did not experience complications.
706

	Control (n)	**Case (n)**	**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**	**AUC (%)**
Discovery	25 (0)	24 (3) [All PE]	88 [70-96]	100 [91-100]	100 [89-100]	89 [74-97]	94 [87-100]
Del Vecchio[1] (2020)	17 (4)	8 (0) [5 PE, 3 GH]	100 [74-100]	76 [53-91]	67 [39-88]	100 [83-100]	88 [78-97]
Del Vecchio[2] (2020)	8 (1)	8 (0) [5 PE, 3 GH]	100 [74-100]	88 [55-99]	89 [59-99]	100 [71-100]	94 [79-100]

707

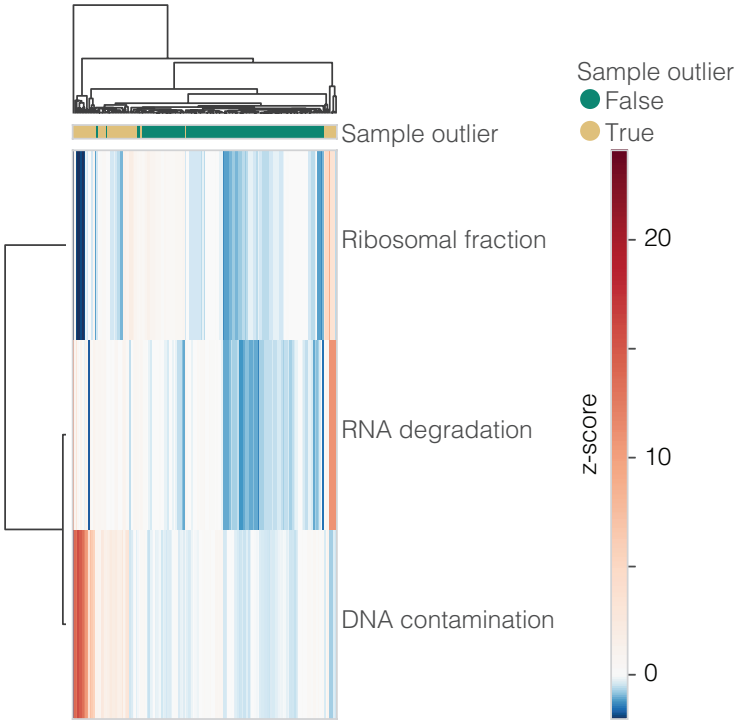
Supplementary text

Supplementary note 1: Establishing quality metrics to identify sample outliers

Because cfRNA measurements can be noisy [30,56], we have previously developed and reported on three quality metrics that can flag sequenced cfRNA samples with poor quality [51,52]. Specifically, these metrics aim to quantify unusually high levels of RNA degradation, DNA contamination, and/or ribosomal RNA by comparing a given sample's value for any of these metrics with what we expect empirically. We defined reasonable expected values for each metric based on the 95th percentile for ~700 previously sequenced samples across 3 cohorts.

We found that samples with outlier values for at least one of these three metrics both clustered separately (Fig S1) and served as leverage points in principal component analysis (PCA) (Fig S2A). To avoid introducing unwanted bias, we removed these low-quality samples from any further analysis. After removing outlier samples, we reran PCA and noticed that two samples continued to serve as leverage points (Fig S2B). We suspected that this may be due to genes that were poorly detected and consequently performed further filtering to identify well-detected genes across the entire cohort. Specifically, we used a basic cutoff that required a given gene be detected at a level of at least 0.5 counts per million reads (CPM) in at least 75% of samples after removing outlier samples. Following this step, we retain 7,041 genes for analysis. Upon inspection, we find that visualization using PCA is no longer driven by leverage points (Fig S2C).

728 **Fig S1. Samples with outlier values for at least one of QC metric cluster separately from most**
729 **non-outlier samples.**



730

Fig S2. Sample outliers and poorly detected genes drive principal component analysis (PCA) and serve as leverage points. Visualization of the top two principal components when performed using all samples and all genes (**A**) or only samples that pass QC metrics (**B**) reveals that certain samples can act as leverage points. Once sample outliers and lowly detected genes are removed from the cfRNA gene matrix (**C**), the top two principal components reflect natural variance in the data and are no longer driven by a few leverage points.

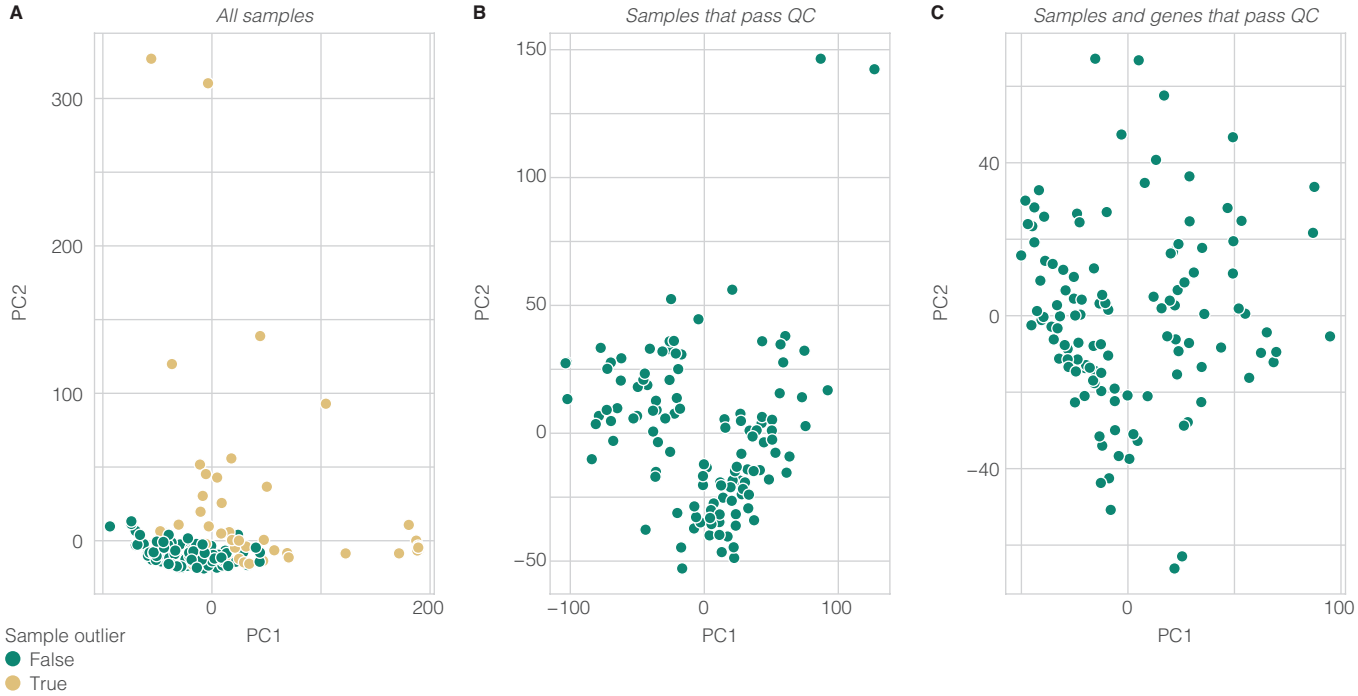
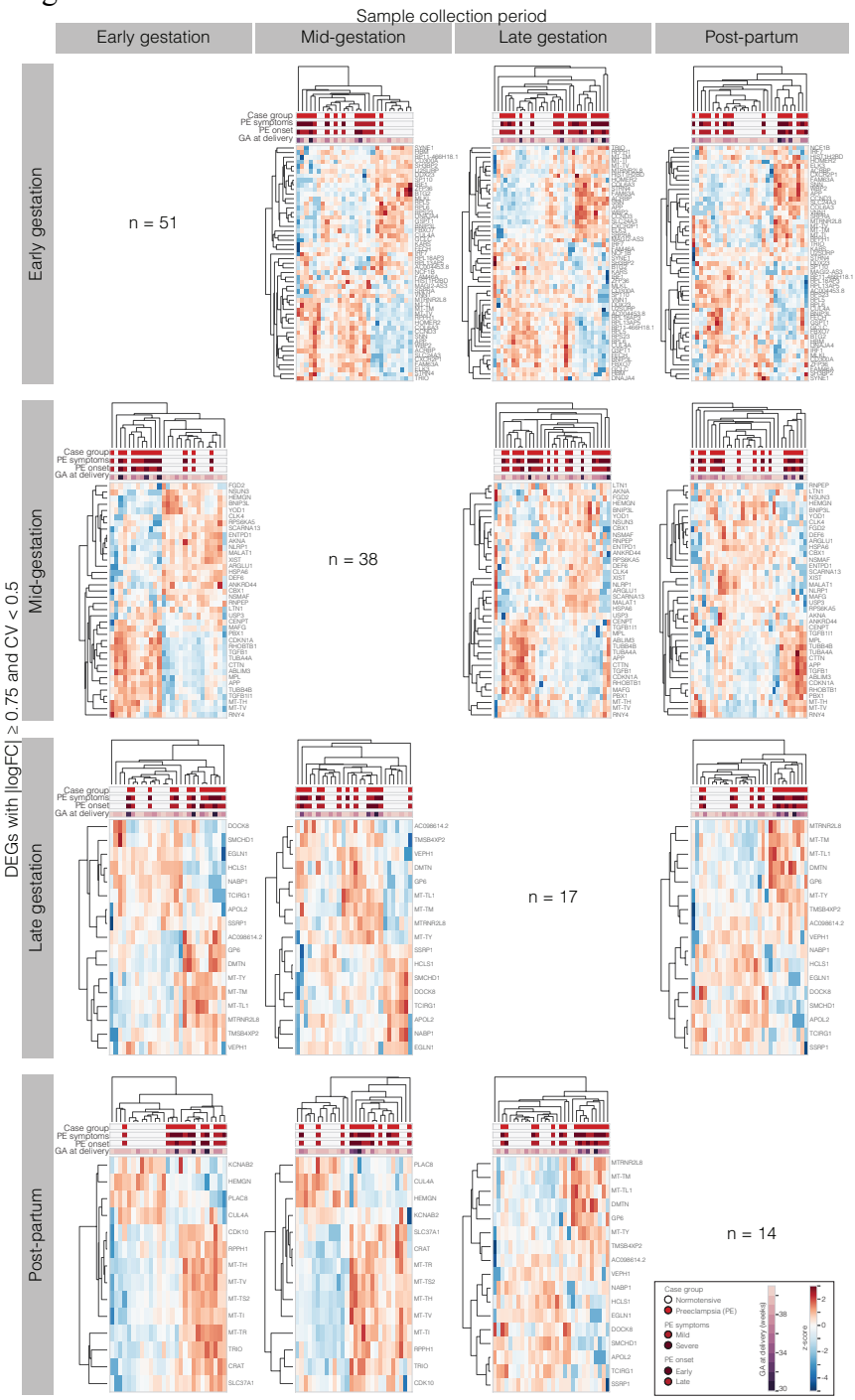
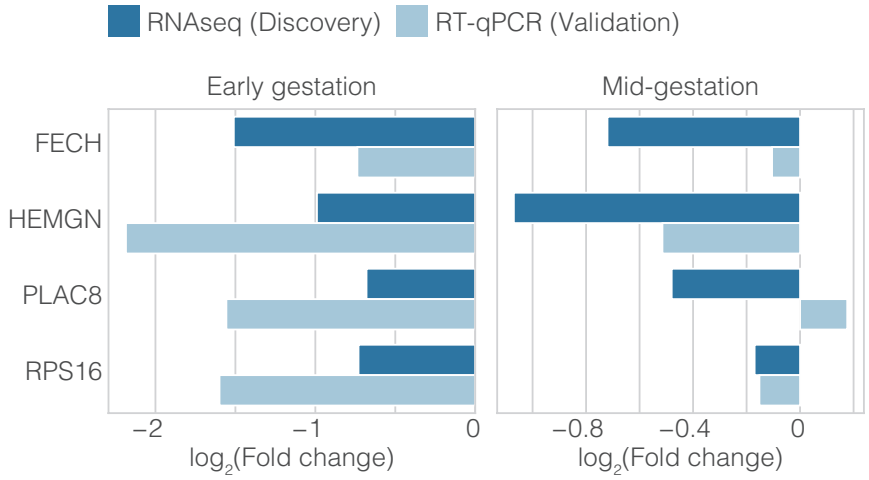


Fig S3. Across gestation and prior to diagnosis, changes in the cfRNA transcriptome identified at one timepoint can moderately segregate PE and NT samples at other timepoints. Differentially expressed genes (DEGs) with |logFC| ≥ 0.75 and CV < 0.5 were identified at each timepoint across gestation. Each row visualizes how well a specific DEG subset from a given sample collection period can separate PE and NT samples in all other collection periods (columns). The number of genes identified per sample collection period is highlighted along the main diagonal.



746 **Fig S4. Log(Fold change) as estimated by RNAseq and RT-qPCR across two cohorts**
747 **(discovery and validation) broadly agree with the exception of PLAC8 at mid-gestation.**



748

Fig S5. Longitudinal dynamics across gestation can be best described in 3 clusters. The optimal k clusters (dashed line) or elbow of this convex plot comparing a performance metric, the sum of squared distances, and values of k occurs at k = 3.

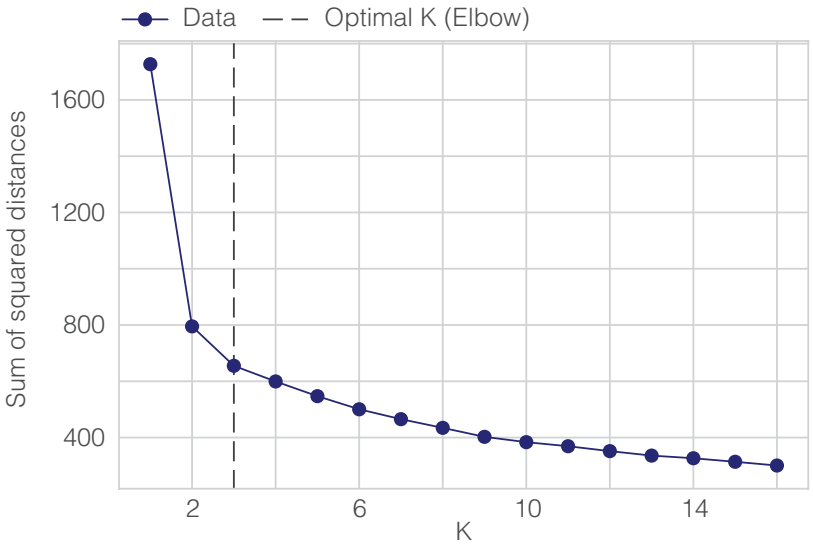
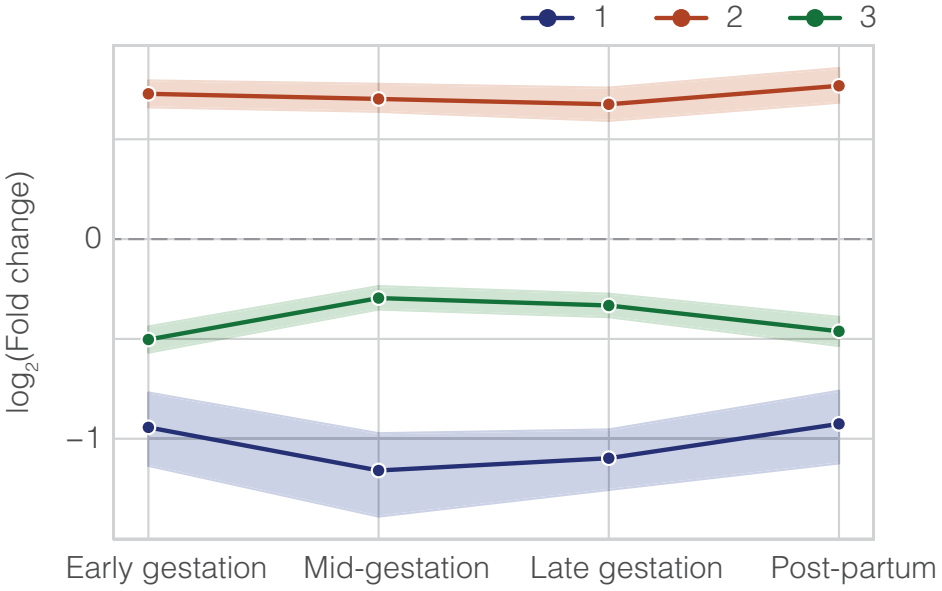
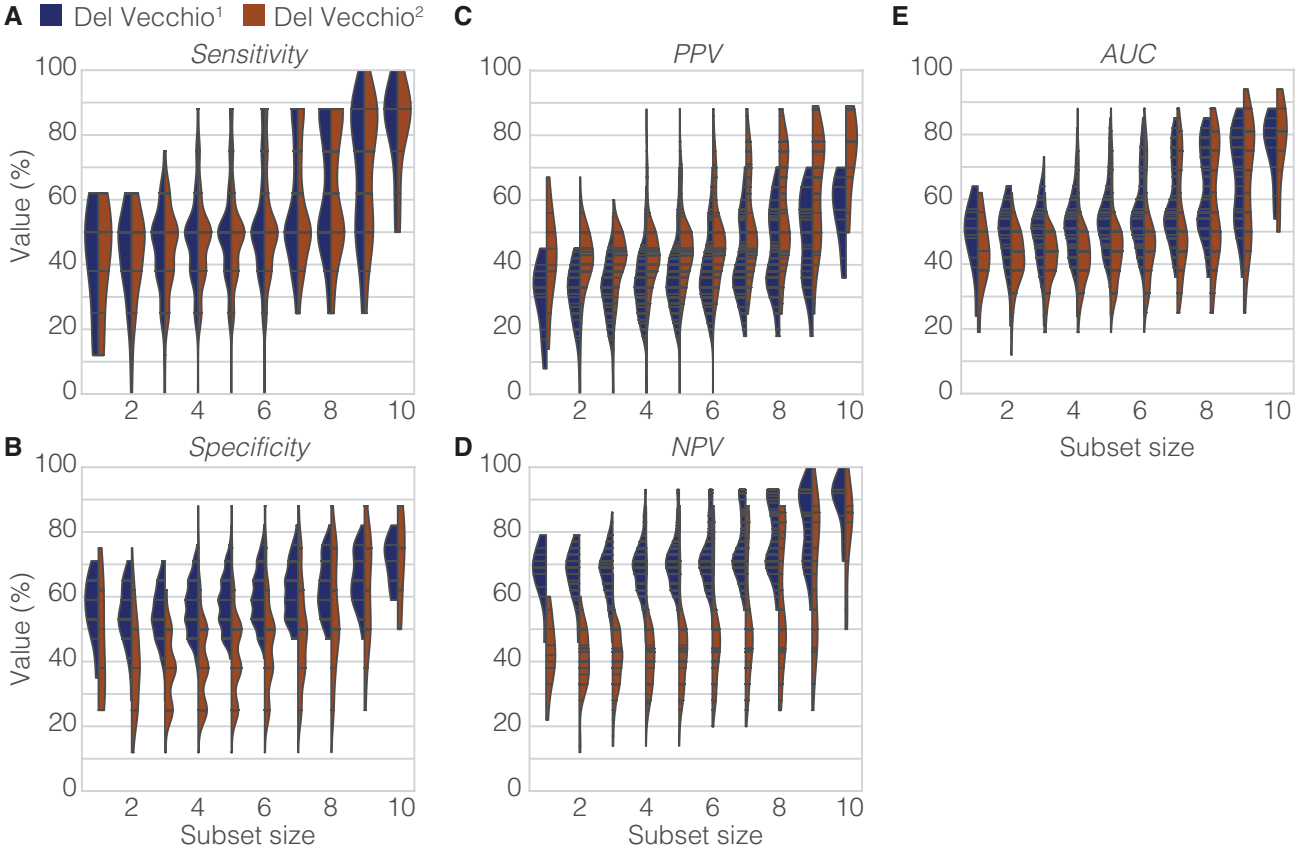


Fig S6. K-means clustering reveals meaningful longitudinal patterns. Following permutation of the data matrix prior to k-means clustering, longitudinal changes over gestation are replaced by 3 flat lines.



757 **Fig S7. Logistic regression models trained on subsets of 1–10 genes of the initial 11 genes can**
758 **moderately predict future PE onset** with improving performance as subset size increases and as
759 characterized by sensitivity (**A**), specificity (**B**), PPV (**C**), NPV (**D**), and AUC (**E**). In Del
760 Vecchio[1], the control group is defined as samples from any pregnant mother who did not develop
761 PE or gestational hypertension including those with other underlying or pregnancy-related
762 complications like chronic hypertension and gestational diabetes respectively. In Del Vecchio[2],
763 the control group is defined as samples strictly from NT pregnant mothers who did not experience
764 complications.
765



766

767 **Table S1. Participant, pregnancy, and PE characteristics across both discovery and**
768 **validation datasets.** Maternal age and BMI, gestational age (GA) at delivery, fetal weight, and
769 GA at PE onset are reported mean ± standard deviation. All other values are reported as
770 percentages. Small for GA (SGA) was defined as an infant with a birthweight below the 10th centile
771 for their GA at delivery.
772

Maternal pre-pregnancy characteristics

		Age (years)	BMI (kg/m^2)	Nulliparous (%)
Discovery	NT	32 ± 5	23 ± 3	44
	PE	31 ± 6	30 ± 8	38
Validation	NT	32 ± 3	25 ± 6	12
	PE	32 ± 4	26 ± 4	29

Maternal ethnicity/race

		Ethnicity	Race			
		Hispanic (%)	White (%)	Black (%)	Asian (%)	Unknown (%)
Discovery	NT	0	100	0	0	0
	PE	42	50	4	25	21
Validation	NT	38	100	0	0	0
	PE	57	86	0	14	0

Pregnancy characteristics

		GA at delivery (weeks)	Mode of delivery (% vaginal)	Multi-gestation (%)	Preterm delivery (%)	Fetal sex (% male)	Fetal weight (kg)	SGA (%)
Discovery	NT	40 ± 1	61	0	0	56	3.5 ± 0.4	0
	PE	36 ± 3	46	12	38	54	2.8 ± 0.7	4
Validation	NT	39 ± 1	62	0	0	25	3.6 ± 0.6	0
	PE	38 ± 1	57	0	14	57	3.4 ± 0.3	0

PE characteristics

		GA at onset (weeks)	Onset type (% early)	Symptom severity (% severe)	History of PE (%)
Discovery	PE	36 ± 4	25	54	17
Validation	PE	38 ± 2	0	43	14

773

774 **Table S2. PE prediction performance metrics for samples collected late in gestation or at PE**
775 **diagnosis (between 17–38 weeks).** Control and PE sample numbers are reported as the total
776 sample number and in parentheses, the number of samples misclassified. All other statistics
777 including sensitivity specificity, PPV, NPV, and AUC are reported as the estimated percentage
778 followed by the 95% CI in square brackets.
779

	NT (n)	**PE (n)**	**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**	**AUC (%)**
Discovery	13 (2)	15 (4)	73 [48-90]	85 [59-97]	85 [59-97]	73 [48-90]	79 [64-93]
Munchel (2020) iPEC	73 (35)	40 (13)	68 [52-80]	52 [41-63]	44 [32-56]	75 [61-85]	60 [50-69]
Munchel (2020) PEARL-PEC	24 (9)	24 (9)	62 [43-80]	62 [43-80]	62 [43-80]	62 [43-80]	62 [49-77]

780

781 **Table S3. PE prediction relies on the cfRNA levels of 11 genes.** For every gene, symbol,
782 ENSEMBL ID, full name, odds ratio (OR) based on the logistic regression coefficient, and a subset
783 of GO biological processes and molecular functions are reported from left to right.
784

Gene	ENSEMBL	Name	OR	Biological Process [GO]	Molecular Function [GO]
GSPT1	ENSG00000103342	G1 to S phase transition 1	1.69	regulation of translational termination, protein methylation, G1/S transition of mitotic cell cycle	GTPase activity, RNA binding, GTP binding, translation release factor activity, protein binding
BNIP3L	ENSG00000104765	BCL2 interacting protein 3 like	0	regulation of protein targeting to mitochondrion, defense response to virus, regulation of apoptotic process	lamin binding, protein homodimerization activity, identical protein binding, protein binding
MARCH2	ENSG00000099785	membrane associated ring-CH-type finger 2	0.02	protein ubiquitination, endocytosis	protein binding, ubiquitin protein ligase activity
IGF2	ENSG00000167244	insulin like growth factor 2	0.12	in utero embryonic development, platelet degranulation, regulation of gene expression by genetic imprinting	growth factor activity, integrin binding, insulin receptor binding, protein serine/threonine kinase activator activity
HEMGN	ENSG00000136929	hemogen	15.83	multicellular organism development, regulation of osteoblast differentiation, cell differentiation	protein binding
OAZ1	ENSG00000104904	ornithine decarboxylase antizyme 1	51.6	positive regulation of intracellular protein transport, negative regulation of polyamine transmembrane transport	ornithine decarboxylase inhibitor activity, protein binding
CSF3R	ENSG00000119535	colony stimulating factor 3 receptor	0.19	regulation of myeloid cell differentiation, cytokine-mediated signaling pathway, cell	signaling receptor activity, cytokine receptor activity

				adhesion, defense response	
RPS15	ENSG00000115268	ribosomal protein S15	0.76	osteoblast differentiation, translation, rRNA processing	RNA binding, structural constituent of ribosome, protein binding
AKNA	ENSG00000106948	AT-hook transcription factor	0.08	neuroblast division in subventricular zone, regulation of inflammatory response, epithelial to mesenchymal transition	DNA binding, protein binding
SNCA	ENSG00000145335	synuclein alpha	2.09	calcium ion homeostasis, synaptic vesicle exocytosis, receptor internalization, regulation of norepinephrine uptake	actin binding, tau protein binding, SNARE binding, phosphoprotein binding, protein kinase inhibitor activity
FECH	ENSG00000066926	ferrochelatase	1.98	response to drug, heme biosynthetic process	ferrochelatase activity, 2 iron

785

Table S4. Logistic regression models trained on some subsets of 4–10 genes of the initial 11 genes can predict future PE onset with nearly equivalent performance metrics. The associated performance metrics for each data split and some high-performing gene subsets is reported including sensitivity (Sens), specificity (Spec), PPV, NPV, and AUC, which are reported as the estimated percentage. In Del Vecchio[1], the control group is defined as samples from any pregnant mother who did not develop PE or gestational hypertension including those with other underlying or pregnancy-related complications like chronic hypertension and gestational diabetes respectively. In Del Vecchio[2], the control group is defined as samples strictly from NT pregnant mothers who did not experience complications.

Subset	*Size*	**Del Vecchio et al (1)**					**Del Vecchio et al (2)**				
		PPV	*NPV*	*Sens*	*Spec*	*AUC*	*PPV*	*NPV*	*Sens*	*Spec*	*AUC*
MARCH2, OAZ1, CSF3R, SNCA	4	64	93	88	76	82	88	88	88	88	88
GSPT1, MARCH2, OAZ1, CSF3R, SNCA	5	70	93	88	82	85	88	88	88	88	88
MARCH2, OAZ1, CSF3R, RPS15, AKNA, SNCA	6	70	93	88	82	85	88	88	88	88	88
MARCH2, HEMGN, OAZ1, CSF3R, RPS15, SNCA, FECH	7	70	93	88	82	85	88	88	88	88	88
GSPT1, MARCH2, HEMGN, OAZ1, CSF3R, RPS15, SNCA, FECH	8	70	93	88	82	85	88	88	88	88	88
BNIP3L, MARCH2, IGF2, HEMGN, OAZ1, CSF3R, RPS15, AKNA, FECH	9	67	100	100	76	88	89	100	100	88	94
BNIP3L, MARCH2, IGF2, HEMGN, OAZ1, CSF3R, RPS15, AKNA, SNCA, FECH	10	67	100	100	76	88	89	100	100	88	94

798 **Table S5. RT-qPCR Taqman probe details**. Unique AssayID corresponds to Biorad assay
799 identification codes.

	Fluorophore	Unique AssayID
FECH	FAM	qHsaCEP0049507
HEMGN	FAM	qHsaCEP0051994
PLAC8	FAM	qHsaCEP0052516
RPS16	FAM	qHsaCEP0050314
DNA contamination control	FAM	qHsaCtlP0001004
RT control	FAM	qHsaCtlP0001001

800
801 **File S1. Table of 108 differentially expressed genes included in Fig 2C heatmaps.** For every
802 gene, the symbol, ENSEMBL ID, sample collection groups for which the gene passed cutoff
803 thresholds, full name, ENSEMBL gene type, and a subset of GO biological processes and
804 molecular functions are reported.
805
806 **File S2. Table of all significant GO terms.**
807

Bibliography

1. Centers for Disease Control and Prevention (CDC). Healthier mothers and babies. *MMWR Morb. Mortal. Wkly. Rep.* **48,** 849–858 (1999).
2. Basso, O. *et al.* Trends in fetal and infant survival following preeclampsia. *JAMA* **296,** 1357–1362 (2006).
3. Behrman, R. E., Butler, A. S. & Institute of Medicine (US) Committee on Understanding Premature Birth and Assuring Healthy Outcomes. Societal Costs of Preterm Birth. (2007).
4. Blencowe, H. *et al.* Born too soon: the global epidemiology of 15 million preterm births. *Reprod Health* **10 Suppl 1,** S2 (2013).
5. Stevens, W. *et al.* Short-term costs of preeclampsia to the United States health care system. *Am. J. Obstet. Gynecol.* **217,** 237–248.e16 (2017).
6. Beam, A. L. *et al.* Estimates of healthcare spending for preterm and low-birthweight infants in a commercially insured population: 2008-2016. *J. Perinatol.* **40,** 1091–1099 (2020).
7. Say, L. *et al.* Global causes of maternal death: a WHO systematic analysis. *Lancet Glob. Health* **2,** e323-33 (2014).
8. Petersen, E. E. *et al.* Vital Signs: Pregnancy-Related Deaths, United States, 2011-2015, and Strategies for Prevention, 13 States, 2013-2017. *MMWR Morb. Mortal. Wkly. Rep.* **68,** 423–429 (2019).
9. McCarthy, F. P., Ryan, R. M. & Chappell, L. C. Prospective biomarkers in preterm preeclampsia: A review. *Pregnancy hypertension* **14,** 72–78 (2018).
10. Roberge, S. *et al.* The role of aspirin dose on the prevention of preeclampsia and fetal growth restriction: systematic review and meta-analysis. *Am. J. Obstet. Gynecol.* **216,** 110–120.e6 (2017).
11. Duley, L. The global impact of pre-eclampsia and eclampsia. *Semin Perinatol* **33,** 130–137 (2009).
12. Hutcheon, J. A., Lisonkova, S. & Joseph, K. S. Epidemiology of pre-eclampsia and the other hypertensive disorders of pregnancy. *Best Pract Res Clin Obstet Gynaecol* **25,** 391–403 (2011).
13. Abalos, E., Cuesta, C., Grosso, A. L., Chou, D. & Say, L. Global and regional estimates of preeclampsia and eclampsia: a systematic review. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **170,** 1–7 (2013).
14. Zhang, J., Meikle, S. & Trumble, A. Severe maternal morbidity associated with hypertensive disorders in pregnancy in the United States. *Hypertens Pregnancy* **22,** 203–212 (2003).
15. Steegers, E. A. P., von Dadelszen, P., Duvekot, J. J. & Pijnenborg, R. Pre-eclampsia. *Lancet* **376,** 631–644 (2010).
16. Li, X. *et al.* Hypertensive disorders of pregnancy and risks of adverse pregnancy outcomes: a retrospective cohort study of 2368 patients. *J Hum Hypertens* (2020). doi:10.1038/s41371-020-0312-x
17. Hansen, A. R., Barnés, C. M., Folkman, J. & McElrath, T. F. Maternal preeclampsia predicts the development of bronchopulmonary dysplasia. *J. Pediatr.* **156,** 532–536 (2010).
18. Wang, A. *et al.* Circulating anti-angiogenic factors during hypertensive pregnancy and increased risk of respiratory distress syndrome in preterm neonates. *J. Matern. Fetal Neonatal Med.* **25,** 1447–1452 (2012).
19. Bellamy, L., Casas, J.-P., Hingorani, A. D. & Williams, D. J. Pre-eclampsia and risk of

854 cardiovascular disease and cancer in later life: systematic review and meta-analysis. *BMJ*
855 **335,** 974 (2007).
856 20. Ahmed, R., Dunford, J., Mehran, R., Robson, S. & Kunadian, V. Pre-eclampsia and future
857 cardiovascular risk among women: a review. *J. Am. Coll. Cardiol.* **63,** 1815–1822 (2014).
858 21. Vikse, B. E., Irgens, L. M., Leivestad, T., Skjaerven, R. & Iversen, B. M. Preeclampsia and
859 the risk of end-stage renal disease. *N. Engl. J. Med.* **359,** 800–809 (2008).
860 22. McDonald, S. D., Han, Z., Walsh, M. W., Gerstein, H. C. & Devereaux, P. J. Kidney disease
861 after preeclampsia: a systematic review and meta-analysis. *Am. J. Kidney Dis.* **55,** 1026–
862 1039 (2010).
863 23. Hypertension in pregnancy. Report of the American College of Obstetricians and
864 Gynecologists' Task Force on Hypertension in Pregnancy. *Obstet. Gynecol.* **122,** 1122–1131
865 (2013).
866 24. Goel, A. *et al.* Epidemiology and mechanisms of de novo and persistent hypertension in the
867 postpartum period. *Circulation* **132,** 1726–1733 (2015).
868 25. Zeisler, H. *et al.* Predictive Value of the sFlt-1:PlGF Ratio in Women with Suspected
869 Preeclampsia. *N. Engl. J. Med.* **374,** 13–22 (2016).
870 26. Sovio, U. *et al.* Prediction of Preeclampsia Using the Soluble fms-Like Tyrosine Kinase 1 to
871 Placental Growth Factor Ratio: A Prospective Cohort Study of Unselected Nulliparous
872 Women. *Hypertension* **69,** 731–738 (2017).
873 27. Gestational hypertension and preeclampsia: ACOG practice bulletin, number 222. *Obstet.*
874 *Gynecol.* **135,** e237–e260 (2020).
875 28. Roberge, S. *et al.* Early administration of low-dose aspirin for the prevention of preterm and
876 term preeclampsia: a systematic review and meta-analysis. *Fetal Diagn Ther* **31,** 141–146
877 (2012).
878 29. Whitehead, C. L., Walker, S. P. & Tong, S. Measuring circulating placental RNAs to non-
879 invasively assess the placental transcriptome and to predict pregnancy complications.
880 *Prenat. Diagn.* **36,** 997–1008 (2016).
881 30. Munchel, S. *et al.* Circulating transcripts in maternal blood reflect a molecular signature of
882 early-onset preeclampsia. *Sci. Transl. Med.* **12,** (2020).
883 31. Del Vecchio, G. *et al.* Cell-free DNA Methylation and Transcriptomic Signature Prediction
884 of Pregnancies with Adverse Outcomes. *Epigenetics* 1–20 (2020).
885 doi:10.1080/15592294.2020.1816774
886 32. Phipps, E. A., Thadhani, R., Benzing, T. & Karumanchi, S. A. Pre-eclampsia: pathogenesis,
887 novel diagnostics and therapies. *Nat. Rev. Nephrol.* **15,** 275–289 (2019).
888 33. Burton, G. J., Redman, C. W., Roberts, J. M. & Moffett, A. Pre-eclampsia: pathophysiology
889 and clinical implications. *BMJ* **366,** l2381 (2019).
890 34. von Dadelszen, P., Magee, L. A. & Roberts, J. M. Subclassification of preeclampsia.
891 *Hypertens Pregnancy* **22,** 143–148 (2003).
892 35. Huppertz, B. Placental origins of preeclampsia: challenging the current hypothesis.
893 *Hypertension* **51,** 970–975 (2008).
894 36. Raymond, D. & Peterson, E. A critical review of early-onset and late-onset preeclampsia.
895 *Obstet Gynecol Surv* **66,** 497–506 (2011).
896 37. Chaiworapongsa, T. *et al.* Differences and similarities in the transcriptional profile of
897 peripheral whole blood in early and late-onset preeclampsia: insights into the molecular
898 basis of the phenotype of preeclampsiaa. *J Perinat Med* **41,** 485–504 (2013).
899 38. Ostlund, I., Haglund, B. & Hanson, U. Gestational diabetes and preeclampsia. *Eur. J.*

900 *Obstet. Gynecol. Reprod. Biol.* **113,** 12–16 (2004).
901 39. Schneider, S., Freerksen, N., Röhrig, S., Hoeft, B. & Maul, H. Gestational diabetes and
902 preeclampsia--similar risk factor profiles? *Early Hum. Dev.* **88,** 179–184 (2012).
903 40. Nerenberg, K. A. *et al.* Risks of gestational diabetes and preeclampsia over the last decade
904 in a cohort of Alberta women. *J Obstet Gynaecol Can* **35,** 986–994 (2013).
905 41. Weissgerber, T. L. & Mudd, L. M. Preeclampsia and diabetes. *Curr Diab Rep* **15,** 9 (2015).
906 42. Rey, E. & Couturier, A. The prognosis of pregnancy in women with chronic hypertension.
907 *Am. J. Obstet. Gynecol.* **171,** 410–416 (1994).
908 43. McCowan, L. M., Buist, R. G., North, R. A. & Gamble, G. Perinatal morbidity in chronic
909 hypertension. *Br J Obstet Gynaecol* **103,** 123–129 (1996).
910 44. Sibai, B. M. *et al.* The impact of prior preeclampsia on the risk of superimposed
911 preeclampsia and other adverse pregnancy outcomes in patients with chronic hypertension.
912 *Am. J. Obstet. Gynecol.* **204,** 345.e1-6 (2011).
913 45. Verburg, B. O. *et al.* New charts for ultrasound dating of pregnancy and assessment of fetal
914 growth: longitudinal data from a population-based cohort study. *Ultrasound Obstet.*
915 *Gynecol.* **31,** 388–396 (2008).
916 46. Han, X. *et al.* Differential dynamics of the maternal immune system in healthy pregnancy
917 and preeclampsia. *Front. Immunol.* **10,** 1305 (2019).
918 47. Ander, S. E., Diamond, M. S. & Coyne, C. B. Immune responses at the maternal-fetal
919 interface. *Sci. Immunol.* **4,** (2019).
920 48. Szarka, A., Rigó, J., Lázár, L., Beko, G. & Molvarec, A. Circulating cytokines, chemokines
921 and adhesion molecules in normal pregnancy and preeclampsia determined by multiplex
922 suspension array. *BMC Immunol.* **11,** 59 (2010).
923 49. Lisonkova, S. & Joseph, K. S. Incidence of preeclampsia: risk factors and outcomes
924 associated with early- versus late-onset disease. *Am. J. Obstet. Gynecol.* **209,** 544.e1-
925 544.e12 (2013).
926 50. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression
927 analysis of RNA-seq data. *Genome Biol.* **11,** R25 (2010).
928 51. Pan, W. Development of diagnostic methods using cell-free nucleic acids. (Stanford
929 University, 2016).
930 52. Moufarrej, M. N., Wong, R. J., Shaw, G. M., Stevenson, D. K. & Quake, S. R. Investigating
931 Pregnancy and Its Complications Using Circulating Cell-Free RNA in Women's Blood
932 During Gestation. *Front. Pediatr.* **8,** 605219 (2020).
933 53. Goh, W. W. B., Wang, W. & Wong, L. Why batch effects matter in omics data, and how to
934 avoid them. *Trends Biotechnol.* **35,** 498–507 (2017).
935 54. Van Calster, B. *et al.* Calibration: the Achilles heel of predictive analytics. *BMC Med.* **17,**
936 230 (2019).
937 55. Brown, L. D., Cai, T. T. & DasGupta, A. Interval Estimation for a Binomial Proportion. *Stat*
938 *Sci* **16,** 101–133 (2001).
939 56. Ibarra, A. *et al.* Non-invasive characterization of human bone marrow stimulation and
940 reconstitution by cell-free messenger RNA sequencing. *Nat. Commun.* **11,** 400 (2020).
941